

University of Nottingham
UK | CHINA | MALAYSIA

Essays on Applied Labour Economics:

Personality Traits, Artificial Intelligence and Formalisation Policies

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

Matías Alejandro Golman

Student ID: 20303841

Supervisors:

Trudy Owens

Richard Upward
María García-Vega

School of Economics

Faculty of Social Sciences
University of Nottingham

April 2026

Acknowledgements

This process of learning and creation would not have been possible without the support of many people and institutions, not only during the doctoral years, but from the very beginning.

I want to thank my siblings, Vanina and Ariel, for their guidance, patience and love. For the conversations and hugs we've had and will have. For being the mirrors that show me it's always possible to rethink and expand oneself. To my parents, Cynthia and Mario, for knowing how to accompany me through uncharted waters with trust and love. For nurturing in me a passion for curiosity and depth. For caring for me and teaching me to care for others. To Anne, for the symbiosis, the adventures, the laughter, and the embraces. For sharing the passion and commitment to what moves us each. To my supervisors, Richard, Trudy and María, for their invaluable guidance and support, both intellectual and personal. For accompanying me through uncertainty, trusting my processes and encouraging me to move forward. For embodying the kind of mentorship I hope to offer others. I will be forever grateful. To my lifelong friends, Pedro, Fabri and Pablito, for the countless conversations. For years of dialectic and hermeneutics. For the exchange of ideas, the discussions and the shared learning. For always being there. To Roxana and Nacho, my mentors who have become foundational, for having encouraged me to undertake this challenge. For teaching me by example. For their affection. To my friends Ire, Juli, Nai, Milu, Yami and Regi, for their diverse paths and perspectives, from which I learn so much. For their enduring care. For the shared experiences and those yet to come. To my friends Lau, Aldi, Fran and Euge, for growing together. For the reflections and encounters. To my friends Nico and Maru, for the adventures on the island that did me so much good. To Marcos, Marcelo, and Maren, friends I found during these doctoral years, for the companionship and the constant support. You made the journey in Beeston more beautiful to travel. To many more friends, family members and colleagues, who accompanied me in adventures, distractions or reflections. Who chose to share moments of their lives with me. To the educational institutions and professors who shaped me and inspired me. Above all, to the University of Buenos Aires, for its intellectual depth and freedom, and for the ingrained sense of belonging. To Argentina and Bariloche, for their people and forests, which impel me every day.

To my grandparents, who accompany me always.

Matías

April, 2026

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Chapter 3 is co-authored with Verónica Escudero and Elva López Mourelo. I conducted the systematic literature review, including the identification, screening, and selection of studies. I designed and constructed the meta-database, extracting and coding the paper's characteristics and their estimates, following established protocols. I developed and adapted the analytical framework, including all meta-regression specifications and publication bias correction methods. I performed all statistical analyses, produced all visualisations and tables, and wrote the complete manuscript draft, including the introduction, theoretical framework, methodology, results, discussion and appendices International Labour Reviews. Dr. Verónica Escudero provided supervision of the data collection process, contributed to the initial methodology draft, and guidance on analytical approaches. Dr. Elva López Mourelo supervised the data collection quality assessment and assisted in defining intervention categories and inclusion criteria.

Matías Alejandro Golman
April, 2026

Introduction

Labour markets worldwide face both persistent and emerging challenges. In Latin America, informality excludes millions of workers from social protection, job security, and labour rights. In developed economies, rapid changes in how firms recruit and what they value in workers raise new questions about skill demands and hiring practices. Understanding these challenges requires rigorous empirical evidence. This thesis contributes through three independent essays in applied labour economics. The first two chapters examine how employers communicate requirements and value characteristics in the UK labour market using online job vacancy data. The third provides a comprehensive assessment of formalisation policies in Latin America. Although the chapters differ in geographic focus and research questions, they share a methodological commitment to using novel data sources and careful empirical strategies to address policy-relevant questions.

Research has shown that personality traits predict wages, job stability, and career trajectories (Alderotti et al., 2023; Flinn et al., 2025; Gensowski, 2018). At the same time, the spread of artificial intelligence raises questions about which human attributes employers will value (Acemoglu et al., 2022a; Alekseeva et al., 2021). Understanding how employers communicate personality requirements in job advertisements, how these demands relate to wage offers, and how they change when firms adopt AI-related skills provides evidence relevant for workers making career decisions and policymakers designing education and early training programmes.

In developing countries, informality remains a central challenge. A large share of workers in Latin America lack access to social protection, job security, and labour rights. Governments have implemented diverse policies to encourage formalisation, from training programmes and wage subsidies to tax reductions and information campaigns (Escudero et al., 2019; Jessen and Kluve, 2021). Yet evidence on effectiveness remains scattered across studies using different methodologies, outcome definitions, and time horizons. Synthesising this evidence is essential for guiding policy decisions in a region where informality affects millions of workers.

Chapter 1: The Effect of AI Adoption on Personality Traits Demand. *Pages 9-124*

The first essay examines the demand for Big Five personality traits in UK job advertisements and how this demand relates to AI adoption. Using Boolean retrieval methods on 11.7 million online job vacancies from Adzuna between 2017 and 2022, I identify personality trait requirements through dictionaries compiled from established psychology and economics research (Goldberg, 1990; Costa Jr and McCrae, 1992; McCrae and John, 1992;

Heckman and Kautz, 2012). Working with unstructured job advertisement text requires building custom dictionaries and implementing validation procedures to transform free text into job characteristics and requirements. This involved manual validation through multiple rounds of randomly selected advertisements and over 1,700 specific replacements to ensure accuracy.

The analysis reveals that personality traits are highly demanded in UK job advertisements. Seventy-two per cent of vacancies mention at least one personality trait, exceeding the proportion mentioning education requirements (11 per cent) or soft skills (62 per cent). Extraversion is the most frequently demanded trait, appearing in 44 per cent of advertisements, followed by Conscientiousness (34 per cent), Openness (30 per cent), Emotional Stability (14 per cent), and Agreeableness (13 per cent). These demands vary systematically across occupations, with customer-facing and administrative roles showing the highest requirements.

After accounting for occupational changes, personality demands remain mostly stable over time. I use two approaches to examine how AI relates to personality demands. At the vacancy level, fixed effects models show that jobs requiring AI skills are more likely to demand Openness (7.2 percentage points) but less likely to demand Extraversion (2.7 percentage points), Emotional Stability (2.2 percentage points), and Conscientiousness (1.7 percentage points). At the firm level, I implement a staggered difference-in-differences design following Callaway and Sant'Anna (2021). Using entropy balancing to ensure comparability between treated and control firms (Hainmueller, 2012), I find that firms increase their demand for Openness (3.6 percentage points) across all subsequent hiring after posting AI positions for the first time. These effects persist when excluding AI positions from the analysis, showing that the relationship extends across firms' entire workforce.

The consistency of the Openness effect across both empirical approaches suggests that when firms adopt AI-related skills, they increasingly value traits related to creativity and adaptability. This contributes to research on how new technologies affect labour demand by showing that AI adoption relates not only to cognitive and technical skill requirements but also to personality characteristics.

Chapter 2: The Wage Penalty Paradox. *Pages 125-233*

The second essay investigates a puzzling pattern: while individual-level studies show that personality traits increase wages, job advertisements mentioning these traits are associated with lower wage offers (Acemoglu et al., 2022b; Deming and Kahn, 2018; Brenčič and McGee, 2023). Using the same dataset of 11.7 million UK job postings, I document this paradox and test four potential explanations.

The baseline analysis confirms that advertisements mentioning personality traits offer 4.6 per cent lower wages on average. The penalty is largest for Extraversion (3.2 per cent) and Emotional Stability (3.1 per cent), followed by Conscientiousness (2.0 per cent) and Agreeableness (1.4 per cent). Only Openness shows a small positive effect (0.7 per cent).

First, I examine whether personality trait mentions correlate with occupation or job characteristics that independently determine wages. Occupation fixed effects produce the largest reduction in trait penalties, but significant penalties persist. Trait penalties are smaller in occupations with higher proportions of women and migrants but larger in occupations with more degree holders and higher mean wages. This pattern contradicts simple productivity-based explanations.

Second, following the theory of equalising differences (Rosen, 1986), I test whether workers accept lower wages in exchange for non-monetary benefits. Simple additive controls for amenities had minimal impact. However, interaction analysis reveals that when personality traits are combined with traditional benefits, workplace culture signals, and quality amenities, trait penalties are reduced or eliminated. Some amenity combinations amplified penalties, particularly remote work, showing strategic use of personality language depending on compensation context.

Third, using TF-IDF text analysis, I examine whether trait mention intensity affects wage implications. The relationship is highly non-linear: low-intensity mentions carry minimal penalties, while high-intensity mentions show substantial penalties. This suggests employers use trait language strategically for different purposes depending on emphasis level.

Fourth, I train a BERT classifier to identify implicit gender targeting and examine whether trait penalties differ across gender categories. The evidence does not support demographic targeting as the primary driver. Most traits show larger penalties in male-targeted advertisements, and both categories show similar intensity patterns.

These findings indicate that the personality trait wage penalty reflects multiple overlapping employer strategies. The relationship between trait mentions and wages depends on context, linguistic presentation, and organisational characteristics rather than simply reflecting productivity differences.

Chapter 3: The Effectiveness of Formalisation Policies in Latin America. *Pages 234-303*

The third essay, co-authored with Verónica Escudero and Elva López Mourelo, presents the first comprehensive meta-analysis of formalisation policies in Latin America. We analyse 79 impact evaluations with 527 estimates from interventions implemented between

1990 and 2020. Unlike previous meta-analyses (Card et al., 2010; Floridi et al., 2020), we include Spanish-language studies and unpublished government reports, examine interventions targeting both worker and firm formalisation, and calculate standardised mean differences to enable systematic comparison across studies.

Our analysis addresses publication bias using FAT-PET-PEESE methods. We find that bias correction substantially reduces effect sizes. Uncorrected pooled estimates overstate true effects, with formality estimates falling by approximately one-third and earnings estimates by half or more after correction.

After correcting for publication bias, active labour market policies show consistent positive effects on formality outcomes (0.238 standard deviations), corresponding to increases of 12 to 14 percentage points in the probability of formal employment. Information campaigns (0.226 standard deviations) and labour inspections (0.320 standard deviations) also show positive effects. For earnings outcomes, social protection programmes with activation components show the largest effects (1.258 standard deviations) but negligible formality impact. This suggests these programmes primarily affect income through direct transfers rather than facilitating transitions to formal employment.

Programme benefits strengthen over time, with long-term effects exceeding short-term impacts by approximately 0.10 standard deviations for formality and 0.32 standard deviations for earnings. Multi-component interventions outperform single-component policies, indicating potential complementarities between different policy tools.

Contribution and Structure

Together, these essays contribute to understanding how labour markets function and how policy can improve outcomes for workers. Chapters 1 and 2 provide evidence on how employers communicate requirements in job advertisements, developing methods to measure employer demands using unstructured text data. Chapter 3 offers guidance for policymakers seeking to reduce informality, emphasising realistic expectations about impact magnitudes and the value of sustained interventions.

Each chapter is written as a self-contained essay with its own introduction, literature review, methodology, results, conclusion and references. Chapters 1 and 2 use the same database, so Chapter 2 refers to the former when describing the cleaning process of the job vacancy data. Chapter 3 can be read independently. The remainder of the thesis proceeds as follows. Chapter 1 presents the analysis of AI adoption and personality trait demand. Chapter 2 examines the wage penalty paradox. Chapter 3 provides the meta-analysis of formalisation policies. The thesis concludes with a discussion of implications for research and policy.

References

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2022a). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, 40(S1):S293–S340.

Acemoglu, D., Hazell, J., Restrepo, P., et al. (2022b). AI and jobs: evidence from online vacancies. *Journal of Labor Economics*, 40(S1).

Alderotti, G., Rapallini, C., and Traverso, S. (2023). The big five personality traits and earnings: A meta-analysis. *Journal of Economic Psychology*, 94:102570.

Alekseeva, L., Azar, J., Gine, M., Samila, S., and Taska, B. (2021). The demand for ai skills in the labor market. *Labour economics*, 71:102002.

Brenčič, V. and McGee, A. (2023). Employers' demand for personality traits. Discussion Paper 16083, Institute of Labor Economics (IZA).

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Card, D., Kluve, J., and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal*, 120(548):452–477.

Costa Jr, P. T. and McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, 6(4):343–359.

Deming and Kahn (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, 36(S1):S337–S369.

Escudero, V., Kluve, J., López Mourelo, E., and Pignatti, C. (2019). Active labour market programmes in Latin America and the Caribbean: Evidence from a meta analysis. *The Journal of Development Studies*, 55(12):2644–2661.

Flinn, C. J., Todd, P. E., and Zhang, W. (2025). Labor market returns to personality: A job search approach to understanding gender gaps. *Journal of Political Economy*, 133(4):000–000.

Floridi, A., Demena, B. A., and Wagner, N. (2020). Shedding light on the shadows of informality: A meta-analysis of formalization interventions targeted at informal firms. *Labour Economics*, 67(101925).

Gensowski, M. (2018). Personality, IQ, and lifetime earnings. *Labour Economics*, 51:170–183.

Goldberg, L. R. (1990). An alternative "description of personality": the big-five factor structure. *Journal of Personality and Social Psychology*, 59(6):1216—1229.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1):25–46.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4):451–464.

Jessen, J. and Kluve, J. (2021). The effectiveness of interventions to reduce informality in low- and middle-income countries. *World Development*, 138(105256).

McCrae, R. R. and John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):175–215.

Rosen, S. (1986). The theory of equalizing differences. In *Handbook of labor economics*, volume 1, pages 641–692. Elsevier.

The effect of AI adoption on Personality Traits demand: evidence from Online Job Vacancies

Matías Golman*

Abstract

This study examines the demand for Big Five personality traits in the UK labour market and their interaction with artificial intelligence (AI) adoption. Using text retrieval methods on 11.7 million online job vacancies (2017-2022), I find that 72 per cent of vacancies require at least one personality trait. Extraversion is most common (44 per cent), followed by Conscientiousness (34 per cent) and Openness (30 per cent). After accounting for occupational changes, personality demands remain mostly stable over time, with only Conscientiousness showing a decline (-2.1 percentage points) and Extraversion a modest increase (1.7 percentage points) by 2022. I use two approaches to study how AI affects personality demands. At the vacancy level, jobs requiring AI skills are more likely to demand Openness (7.2 percentage points) and less likely to demand other traits, even within the same firm and occupation. At the firm level, after companies adopt AI for the first time, they increase their demand for Openness (3.6 percentage points) across all subsequent hiring. These findings suggest that as firms adopt AI technologies, they increasingly value workers with traits associated with creativity and adaptability, both for AI-demanding positions and across their broader workforce.

Keywords: personality traits, artificial intelligence, Big Five, online job postings, labour demand, AI adoption, text analysis, UK labour market, staggered difference-in-differences
JEL Codes: J23, J24, O33, C55, M51

*University of Nottingham. Email: matias.golman@gmail.com.

Contents

1 Introduction **11**

2 Literature Review **14**

2.1 Measuring Personality Demands in Labour Markets 16

2.2 Artificial Intelligence and Labour Market Transformation 17

2.3 Research Gap and Study Contribution 18

3 Data **18**

4 Measurement **20**

4.1 Personality traits retrieval 21

4.1.1 Examples of Classification 23

4.2 AI demand retrieval 25

5 Validation **26**

6 Empirical design **31**

7 Results **35**

7.1 Personality Traits demand and their evolution 38

7.1.1 Complementarities in Trait Demands 38

7.1.2 Relationship with Other Job Requirements 39

7.1.3 Changes over time 42

7.2 AI demand and Personality Traits 45

7.2.1 Robustness: Distinguishing AI from General Computer Skills . . . 52

7.3 From Vacancy-Level to Firm-Level Effects 53

7.4 Staggered Difference-in-Differences Analysis 57

8 Conclusions **62**

9 Appendices **70**

1 Introduction

In the contemporary labour market, personality traits are recognised as important factors in employment outcomes. Research from the British Workplace Employment Relations Study shows that most employers use personality tests to complement performance or competency tests in their hiring processes (McGee & McGee, 2020). This practice suggests that employers value candidates' personality traits alongside their technical skills and competencies. The demand for personality traits has been linked to several factors, including the proliferation of information and communication technologies, the implementation of innovative work practices, and the transition from manufacturing to service-based employment (Borghans et al., 2014). However, the extent to which demand for specific personality traits has changed over time and how this demand varies between different occupations and firms remains an open question in the literature.

Moreover, the last decade has seen rapid advances in artificial intelligence (AI) and the availability of large datasets. This encompasses computational technologies such as machine learning algorithms, computer vision systems, natural language processing tools, and automated decision-making systems. These are distinct from the generative AI models, such as large language models, that became widely available from late 2022 onwards. The analysis here covers the period 2017–2022, prior to their widespread diffusion. These AI technologies are increasingly being adopted in several industries, potentially reshaping labour markets and skill demands (Acemoglu et al., 2022; Babina et al., 2024; Webb, 2019). Alekseeva et al. (2021) document a substantial increase in AI-related job postings between 2010 and 2019, growing from 0.02 per cent to 0.16 per cent of all online job postings. Similarly, using employee resume data, Babina et al. (2024) find AI investments grew more than seven-fold from 2010-2018 across sectors, with AI-investing firms experiencing significant growth in sales and employment. This growth has been particularly pronounced in high tech sectors and firms with higher cash reserves, higher markups, and more intensive R&D activities. AI adoption is expected to significantly impact labour markets, potentially changing the skills and traits that employers value most.

Despite the recognised importance of personality traits and the rapid adoption of AI, little is known about how these two trends connect. Studies have separately examined the demand for personality traits and the growth of AI-related jobs, but there is a significant gap in our understanding of how AI adoption might influence the demand for specific personality traits across different occupations and industries. This gap is particularly relevant given AI's potential to automate or enhance various tasks, which could change the personality traits that employers require. As Johnson and Acemoglu (2023) observe:

> It is therefore not surprising that despite the spread of AI technologies, many companies are increasingly seeking workers with social, rather than mathematical, or technical skills. At the root of this growing demand for social skills is the reality that neither traditional digital technologies nor AI can perform essential tasks that involve social interaction, adaptation, flexibility, and communication. (p.314)

This study addresses three key research questions: (1) What is the nature and extent of personality trait demands in the UK labour market, and how do these align with occupational characteristics? (2) How have personality trait demands evolved over time? (3) How does AI adoption relate to firms' demands for specific personality traits? The aim is to provide a comprehensive analysis of the relationship between AI adoption and demand for personality traits, offering insights into how the labour market is adapting to technological change. By examining these questions, I contribute to the ongoing discussion on the future of work and provide evidence-based insights for workforce development strategies.

To answer these questions, I use data from 11.7 million online job postings in the United Kingdom from 2017 to 2022, provided by Adzuna, a job vacancy aggregator. Unlike other vacancy aggregators, Adzuna includes complete unstructured job descriptions written by employers. I use advanced text analysis to identify personality trait requirements in job postings, along with measures of AI demand at both vacancy and firm levels. I complement this with occupational-level information from the UK Household Longitudinal Study (UKHLS) on self-assessed personality traits and firm-level details from Companies House and the Financial Analysis Made Easy (FAME) databases.

I find that 72 per cent of vacancies mention at least one word related to personality traits. Extraversion is the most frequently mentioned trait (44 per cent), followed by Conscientiousness (34 per cent), Openness (30 per cent), Emotional Stability (14 per cent), and Agreeableness (13 per cent)[1]. Notably, personality traits appear more often in job postings than soft skills (interpersonal competencies like communication and leadership), hard skills (technical capabilities such as data analysis and engineering), experience, seniority, and educational qualifications. I also find substantial variation in the demand for personality traits across occupations, with a higher predominance in Sales, Customer Service, Caring and Leisure-related roles.

When analysing time trends, I find that while raw personality trait demands have increased, most of these changes reflect shifts in the types of vacancies being advertised. After controlling for occupational and firm composition, only one significant trend emerges: a steady decline in Conscientiousness (-2.1 percentage points by 2022). I also find a modest

[1] These are defined in Section 4.

increase in Extraversion (1.7 percentage points by 2022), which is significant in all years after 2017 except 2022.

I use two empirical approaches: a fixed effects model at the vacancy level and a staggered difference-in-differences analysis at the firm level. The analysis of AI adoption reveals significant effects on personality trait demands, using both fixed effects and staggered difference-in-differences approaches. The fixed effects analysis shows that jobs requiring AI skills are significantly more likely to demand Openness (7.2 percentage points) but less likely to demand other traits, particularly Extraversion (-2.7 percentage points), Emotional Stability (-2.2 percentage points), and Conscientiousness (-1.7 percentage points). These findings align with theoretical predictions about AI complementing creative and adaptive capabilities while potentially replacing routine-oriented and interpersonal tasks (Johnson & Acemoglu, 2023).

The staggered difference-in-differences analysis compares firms that adopt AI at different points in time to examine how personality trait demands change following AI adoption. This approach tracks the same firms before and after they first adopt AI technologies, using firms that have not yet adopted AI as a control group. To ensure that treated firms (those adopting AI) and control firms are comparable in their observable characteristics, I employ entropy balancing, matching on pre-treatment personality trait demands, firm characteristics, and occupational composition. The results show modest but significant effects on overall personality trait demands, with firms increasing their demand for personality traits by 2.7 percentage points after AI adoption. The effect is strongest for Openness, with a 3.6 percentage point increase in demand following AI adoption, while Emotional Stability shows a decline of 2.6 percentage points. Importantly, these effects persist when excluding AI positions from the analysis, showing that AI adoption leads to changes in personality requirements across firms' entire workforce rather than only in AI-intensive roles. The consistency of the Openness effect across both empirical approaches strengthens confidence in AI's role in increasing demand for creative and adaptive traits.

This study contributes to the existing literature in several ways. First, it provides a comprehensive analysis of the demand for personality traits in the UK labour market using a large dataset of online job postings. This allows for a detailed understanding of how employers value different personality traits across occupations and industries, while validating these demands through assortative matching with workers' self-reported characteristics. Second, this study is unique in its temporal scope, analysing personality trait demands over 5 years (2017-2022), showing how they have evolved after accounting for changes in occupational composition. Third, my findings challenge the notion that formal qualifications are the primary focus of job advertisements, showing that personality traits appear in 72 per cent of vacancies compared to just 11 per cent mentioning educational

requirements.

Fourth, this study makes new contributions to our understanding of how AI adoption affects labour demand by examining its relationship with personality traits, an aspect previously unexplored. While existing research has documented AI's impact on technical skill requirements (Acemoglu et al., 2022) and cognitive tasks (Babina et al., 2024), this study is the first to systematically analyse how AI adoption influences demands for widely-known measures of personality traits. Fifth, by using both fixed effects and staggered difference-in-differences analyses, this research provides causal evidence on how AI adoption shapes personality requirements, finding a strong complementarity between AI and traits related to creativity and adaptability. This suggests that as firms adopt AI technologies, they increasingly value workers' capacity for innovation and flexible thinking.

These insights have important implications for job seekers, educators and policy makers. The results suggest the need to consider personality development, particularly traits related to creativity and adaptability, alongside traditional skill acquisition in workforce preparation strategies. This is especially relevant as firms continue to adopt AI technologies, which appear to increase demand for these characteristics.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature on personality traits in the labour market and recent developments in AI adoption. Sections 3 and 4 describe the data sources and methodology in detail. Section 5 examines the assortative matching between job requirements and worker characteristics. Section 6 presents the empirical framework. Section 7 presents the main results for the demand for personality traits and their interaction with AI adoption. Finally, Section 8 concludes and suggests directions for future research.

2 Literature Review

Personality traits, character attributes, soft skills, and social skills are often used interchangeably in the literature to reflect personal characteristics outside cognitive attributes. However, these terms connote different properties and levels of malleability (Heckman & Kautz, 2012; Woessmann, 2024). Soft skills and social skills represent dynamic attributes related to interpersonal and social competencies such as leadership, negotiation, teamwork, and communication. These are defined by the employer, locally relevant, generally considered learnable and modifiable through investments in training and work experience (Grugulis & Vincent, 2009; Matteson et al., 2016). Employers may select these during hiring or develop them in staff over time.

In contrast, personality traits and character attributes are more stable individual characteristics that remain relatively fixed after early adulthood (Cobb-Clark & Tan, 2011; Heckman & Kautz, 2012). They are also largely uncorrelated with cognitive skills

(McCrae & John, 1992). Due to this stability, employers have stronger incentives to hire individuals who already possess required personality traits.

Economists have increasingly incorporated personality traits into their human capital analyses over recent decades. It is recognised that personality traits significantly explain differences in economic outcomes (Borghans et al., 2008; Heckman & Rubinstein, 2001). The "Big Five" personality traits—Openness, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability—have become a standard framework.

This approach is grounded in the lexical hypothesis, which holds that personality characteristics important for human social interaction leave a systematic trace in natural language: speakers develop words to describe such characteristics, and those words persist in everyday use (Goldberg, 1990). Consequently, the breadth and frequency of trait-descriptive words in a language reflects the social importance of the underlying trait. John (1990) and McCrae and John (1992) formalised this intuition by collecting large inventories of trait-descriptive adjectives from dictionaries and grouping them through factor analysis, arriving at the five broad dimensions that constitute the Big Five. Because the framework is derived from natural language rather than imposed by researchers, it has been validated across languages and cultures (Borghans et al., 2008; King et al., 2005). Applied to job advertisements, the hypothesis implies that employers who value a trait will use the words associated with it when writing vacancy descriptions.

Many studies show the impact of personality traits on labour market outcomes, including unemployment (Fletcher, 2013; Lindqvist & Vestman, 2011; Viinikainen & Kokko, 2012), occupational sorting (Borghans et al., 2014; Edin et al., 2022; Todd & Zhang, 2020) and career success (Judge et al., 1999; Seibert & Kraimer, 2001). Research suggests that personality traits affect workers' productivity and therefore wages (Woessmann, 2024), although these effects vary by gender and across the wage distribution (Collischon, 2020; Flinn et al., 2025; Gensowski, 2018; Heineck & Anger, 2010; Mueller & Plug, 2006; Nyhus & Pons, 2005). For example, emotional stability and agreeableness are linked to higher wages for women, while extraversion correlates with higher wages for men (Mueller & Plug, 2006; Nyhus & Pons, 2005).

Conscientiousness has emerged as a particularly strong predictor of job performance across various occupations (Barrick & Mount, 1991; Judge & Cable, 1997). Certain traits appear more valuable in specific job types; for instance, extraversion correlates strongly with performance in occupations requiring interpersonal attributes, such as management or sales (Barrick & Mount, 1991). Gensowski (2018) found that personality traits, particularly openness to experience and extraversion, have significant effects on lifetime earnings, with these effects varying between different levels of educational attainment. Collischon (2020) used unconditional quantile regressions to show that the returns to personality traits

vary across the wage distribution, with larger effects observed at higher wage quantiles. Rohrbach-Schmidt et al. (2023) found that the relationship between personality traits and wages varies significantly with job task requirements, with extraversion and emotional stability associated with higher wages, while agreeableness and openness tend to correlate with lower wages within occupations.

2.1 Measuring Personality Demands in Labour Markets

Traditionally, personality traits have been measured through surveys and psychological tests administered to individuals. These assessment instruments are constructed based on the lexical hypothesis described earlier, where trait-descriptive adjectives from natural language are systematically analysed and grouped through factor analysis to identify underlying personality dimensions. Questionnaires like the NEO-PI, the Big Five, and similar instruments present respondents with standardised statements derived from this lexical approach, asking them to rate how accurately these statements describe their typical behaviours and feelings.

Recent advances in big data and text analysis have opened new ways to measure personality traits at scale. Deming and Kahn (2018) pioneered the use of online job postings to measure skill demands, including social skills that closely relate to certain personality traits. This approach, while offering unprecedented scale and detail, also presents challenges in terms of representativeness and interpretation (Hershbein & Kahn, 2018). These challenges include potential selection bias in which jobs appear in online platforms, with casual, freelance, or certain industry-specific roles being underrepresented. More fundamentally, personality trait mentions in job advertisements reflect employer demands rather than worker characteristics—capturing what employers believe they need or what they strategically emphasise to attract certain applicants (Abraham et al., 2024).

Building on this methodological foundation, researchers have developed sophisticated methods to extract personality-related information from job advertisements, ranging from keyword-based approaches to advanced machine learning techniques (Khaouja et al., 2021; Lovaglio et al., 2018; Turrell et al., 2019). Brenčič and McGee (2023a, 2023b) employed a knowledge engineering method to identify personality-related terms in online job advertisements, providing insights into employers' demand for specific traits across different occupations and tasks. Their findings using a two-week database from Monster.com reveal that in the US, 54% of job advertisements make at least one reference to personality traits, with employers primarily demanding workers who are extroverted (31%), conscientious (26%), and open-to-experience (21%).

Studies examining temporal changes in personality trait demands have revealed interesting trends. Evidence suggests that the labour market has increasingly rewarded

character skills over time, particularly related with extraversion and agreeableness (Deming, 2017; Weinberger, 2014). This trend has been attributed to technological change and the growing importance of teamwork in the workplace. Analysis of skill demands in the UK labour market has shown an increasing use of both analytical and interpersonal skills over time, with changes occurring mainly within occupations rather than between them (Dickerson & Morris, 2019). More productive firms tend to have higher skill requirements in their job postings, including those related to social skills and personality traits (Deming & Kahn, 2018), suggesting that personality trait demands may vary systematically between firms and labour markets.

2.2 Artificial Intelligence and Labour Market Transformation

Artificial intelligence (AI) has emerged as a transformative force in the labour market, with potential far-reaching implications for skill demands and job characteristics. Recent years have seen a significant increase in AI-related job postings, particularly in high-tech sectors and firms with higher cash reserves, higher markups, and more intensive R&D activities (Acemoglu et al., 2022; Alekseeva et al., 2021; Schmidt et al., 2024). This trend extends beyond the tech sectors, with AI adoption spreading through major industries including finance, manufacturing, retail and healthcare services (Babina et al., 2024; Choi & Leigh, 2024).

The impact of AI on the labour market is complex. Research shows that high-wage occupations and those with higher educational requirements tend to be more exposed to AI than low-wage occupations (Webb, 2019). However, this exposure does not straightforwardly translate into negative employment effects. While AI may substitute for certain tasks and potentially displace workers, research suggests it can create new tasks that complement human skills (Acemoglu & Restrepo, 2019). Supporting this complementarity hypothesis, Babina et al. (2024) find that AI-investing firms experience faster growth in both revenue and employment. The terminology of AI used in all these studies, as well as here, refers to non-generative AI: predictive and task-specific systems designed to automate or augment well-defined tasks. These technologies coexist with, but predate, newer generative models[2].

Studies have found that exposure to AI is associated with a shift in job postings toward non-routine analytical and interpersonal skills (Alekseeva et al., 2021), suggesting

[2]The study focuses on the period 2017–2022, prior to the widespread diffusion of generative AI and large language models. After generative AI became publicly available in late 2022, the AI-related vocabulary appearing in job advertisements changed substantially. Many firms began adopting AI-related language to signal technological orientation rather than to describe specific technical requirements, making it difficult to distinguish genuine AI skill demands from broader "AI hype" in job posting language. Restricting the sample to the pre-generative period ensures that dictionary-based retrieval captures real AI skill requirements rather than the more diffuse use of the terminology that followed.

that as AI takes over more routine and codifiable tasks, there may be increasing demand for uniquely human skills that are harder to automate. Analysis of AI-related job postings has shown that in addition to technical skills, there is a growing emphasis on skills related to communication, problem solving, creativity, and teamwork (Squicciarini & Nachtigall, 2021). These skills are closely associated with certain personality traits, suggesting that as AI adoption increases, there may be growing demand for personality traits that complement AI capabilities.

The increasing importance of "character" skills in the labour market, as documented by several studies (Deming, 2017; Deming & Kahn, 2018), may be partially driven by the need for workers who can interact effectively with AI systems and bridge the gap between technical capabilities and human needs. This trend aligns with observations that neither traditional digital technologies nor AI can fully replace essential tasks involving social interaction, adaptation, flexibility, and communication (Johnson & Acemoglu, 2023).

2.3 Research Gap and Study Contribution

The literature highlights the importance of personality traits in determining labour market outcomes and the growing impact of artificial intelligence on skill demands. However, there remains a significant gap in our understanding of how AI adoption influences the demand for specific personality traits across different occupations and firms. While existing research has provided valuable insight into the relationship between personality traits and labour market outcomes, there is a need for more comprehensive studies that examine how this relationship may be changing over time, particularly in response to technological advancements like AI.

This study aims to bridge this gap by using a large-scale dataset of online job postings to examine the evolving demand for personality traits in the context of increasing AI adoption. By analysing the demand for personality traits over a five-year period and investigating its interaction with AI adoption, this study seeks to provide a more comprehensive understanding of the evolving role of personality in the labour market.

3 Data

This paper uses data from Adzuna, a job vacancy aggregator[3]. Adzuna's algorithms collect information on open vacancies posted across different online platforms. The UK Office of National Statistics (ONS) uses Adzuna data to produce a weekly vacancy indicator.

Adzuna's data limitations are common to other online vacancy aggregators: the scope of online job advertisements does not fully capture all job adverts in the UK. Representative

[3]I am grateful to Urban Big Data Centre for granting the licence to use the Adzuna database under Data Sharing Agreement number 678175.

biases could arise from specific jobs that are not posted in online marketplaces. Casual or freelance work may be advertised using different methods, such as word of mouth, newspapers, local advertising, or in-shop windows. Comparing the 2021 data with the 2021 Census exhibits occupational composition differences at the 1-digit SOC level. Associate Professional & Technical occupations and Caring & Leisure roles have greater representation in online vacancies, while Administrative, Managerial, and Process & Machine Operative roles have lower representation. Table A.1 in Appendix A compares Adzuna's 2021 occupational distribution with the 2021 UK Census at 3-digit SOC level.

Adzuna provided a database without duplicate advertisements for the same week from the same online source, such as job.uk, jobsincare.co.uk or cv-library.co.uk[4]. According to Adzuna's representatives, their point-in-time estimates are more comprehensive than *Burning Glass Technologies (BGT)*, the most widely cited online vacancy aggregator. While Adzuna collects data from job boards, BGT's collection focuses on individual company websites, resulting in fewer vacancies.

Studies using BGT, such as Deming and Kahn (2018), Hershbein and Kahn (2018), Acemoglu et al. (2022), and Schmidt et al. (2024), highlight its structured information as an advantage. By contrast, Adzuna's unprocessed data, while more difficult to manage, offers valuable benefits. The raw job description text allows for flexible dictionary construction, which is essential for identifying personality-trait wording. It also enables more thorough data preprocessing, allowing the extraction of skills, education, and experience criteria directly from job descriptions. Access to the complete advertisement text makes it possible to study trait mentions both within and across advertisements.

When collecting job advertisements, Adzuna stores information from the web page and creates several key variables, including job title, offered salary, posting date and site, contract type (permanent or temporary), location, and whether the job is part- or full-time. I added information extracted from the job descriptions. Specifically, I cleaned the wage variable, job schedule, and contract type, and recoded job locations to the two-digit *Nomenclature des Unités territoriales statistiques* (NUTS) level. I also created indicators for soft and hard skills requirements[5], education, experience, seniority, and bonus availability. Adzuna's vacancy identifier allows tracking of advertisements over time, regardless of posting site.

Since Adzuna lacks standard occupational codes, I implemented a machine learning

[4]I further cleaned potential duplicate vacancies published in other sources or with different vacancy IDs but identical company, location and descriptive text. A step-by-step explanation of this procedure is in Appendix B.

[5]For soft skills, these include *communication skills, critical thinking, language skills, leadership, interpersonal skills, negotiation skills*, among others. For hard skills, I coded *administrative skills, data analysis, design skills, engineering, legal analysis, sales skills, customer service*, among others. The complete list is in Appendix B.

algorithm created by Turrell et al. (2019) to classify advertisements into 3-digit Standard Occupation Codes (SOC). The algorithm categorises advertisements based on their title, description, and industry. Using Companies House data, a public dataset of UK-registered companies, I matched Adzuna's companies to obtain Standard Industrial Codes (SIC), keeping only private companies whose names uniquely matched.Advertisements from recruitment agencies or non-private institutions lack industry codes, which are needed for the Turrell et al. (2019) algorithm and were therefore removed (reducing the sample by 45 per cent) and those without SIC codes (further reducing the sample by 56 per cent).

Finally, this study only uses vacancies with posted wages. A common criticism of wage data from online job advertisements is that they often target lower-skilled workers or positions that need to be filled urgently (Banfi & Villena-Roldan, 2019; Brenčič, 2012). Batra et al. (2023)'s meta-analysis highlights the scarcity of wage information, which is often only available in broad ranges, limiting its reliability as a proxy for actual wages. This scarcity is more evident in high-wage occupations and large private firms. After removing recruitment agencies and advertisements not matching Companies House, the dataset contained wage information for 46 per cent of vacancies. This is twice the proportion found in studies using job boards that don't aggregate sources (24.8 per cent in *monster.com* (Brenčič, 2012); 13.3 per cent in *trabajando.com* (Banfi & Villena-Roldan, 2019); 20 per cent in *careerbuilder.com* (Marinescu & Wolthoff, 2020); and 24.3 per cent in *xmrc.com* (Kuhn & Shen, 2023). The final database contains 11.7 million vacancies from 128,559 companies.

I complemented the job vacancy data with information at both occupational and firm levels. For occupations, I obtained data on gender and migrant proportions, average education, seniority, age, and tenure from the UK Household Longitudinal Study (UKHLS), a household panel survey. The UKHLS also asks individuals to self-assess their personality using the Big Five framework, allowing me to compare personality trait demands in job vacancies with workers' self-reported traits in different occupations. Appendix B provides detailed explanations of the additional data sources, cleaning process, variable construction, refinement of Adzuna's categories, and the Turrell et al. (2019) algorithm. At the firm level, I matched Adzuna data with the FAME database to identify publicly listed companies.

4 Measurement

The most widely used framework for studying personality traits in psychology is the Big Five. These five dimensions are Conscientiousness, Openness to experience, Extraversion, Agreeableness, and Emotional Stability[6]. This framework is based on the lexical hypothesis,

[6]Emotional Stability is commonly described as the opposite of Neuroticism.

which suggests that important personality characteristics become encoded in language over time. The words people use to describe themselves and others reflect important personality dimensions. In job advertisements, this is especially relevant as employers likely use words associated with valued personality traits, and job seekers recognise and respond to these trait-related terms when applying.

Researchers built and refined the taxonomy using factor analysis to group the many adjectives describing personality in natural language into these five traits (John, 1990). The Big Five follows the structure proposed by Goldberg (1990), McCrae and John (1992), and John (1990), who developed the initial work of Cattell (1943).

The Big Five structure has been validated across different languages and cultures (Borghans et al., 2008; King et al., 2005). This cross-cultural consistency has led to its widespread use in economics research, following key papers linking personality types to economic outcomes such as educational attainment, job selection, productivity and earnings (Almlund et al., 2011; Borghans et al., 2008; Heckman, 2011).

Each of the Big Five traits includes specific characteristics (Costa Jr & McCrae, 1992; Goldberg, 1990; Stanek & Ones, 2023):

- *Conscientiousness* relates to being organised, responsible, and diligent, with self-discipline and achievement drive.
- *Openness-to-experience* refers to an interest in new aesthetic, cultural, or intellectual experiences, showing a desire for cognitive exploration.
- *Extraversion* involves an outward orientation towards social interactions and activities, associated with sociability and positive attitude.
- *Agreeableness* appears in cooperative and unselfish behaviour that promotes harmonious interactions.
- *Emotional Stability* refers to consistency in emotional responses without rapid mood changes.

In this study, I use the Big Five framework to analyse personality trait demand in the UK labour market. By identifying words linked to each trait in job advertisements, I measure employers' demands for specific personality characteristics and examine how these vary across occupations and time.

4.1 Personality traits retrieval

As Lovaglio et al. (2018) outline, there are two main approaches to identify job advertisement requirements: *knowledge engineering* and *machine learning techniques*. Knowledge

engineering involves manually defining rules to classify text, while machine learning uses an inductive process to automatically build classifiers. This study uses the Knowledge Engineering method, which aligns with the theoretical foundations of the Big Five personality taxonomy.

I use established personality psychology research to identify relevant trait descriptors. When analysing job advertisements, this method offers two key advantages. First, it keeps trait identification true to the Big Five framework, avoiding semantically related but conceptually different terms that might be incorrectly included by open-vocabulary methods like machine learning. This reduces the risk of identifying traits through broad synonyms with ambiguous meanings. Second, it provides a clear and replicable method for identifying trait-related language in job postings, which helps interpret results in the context of labour markets.

My methodology builds on previous research using similar approaches. Deming and Kahn (2018) and Alekseeva et al. (2021) used Knowledge Engineering to search for various skills and character attributes in job postings. Brenčič and McGee (2023b) and Brenčič and McGee (2023a) also used knowledge engineering to identify personality-related terms in online job advertisements. While Brenčič and McGee (2023b) focused mainly on personality-trait words in a small dataset, and Brenčič and McGee (2023a) expanded this by examining traits across different job tasks, my research examines the relationship with AI demand and provides deeper analysis of job, occupation, and firm attributes.

To identify personality traits, I use a Boolean retrieval approach. A job posting is labelled as requiring one of the Big Five personality traits if it includes at least one keyword from a dictionary related to traits in its job requirements or job title[7]. The demand for each trait is then measured at an aggregate level as the percentage of job postings containing trait-related words, similar to Alekseeva et al. (2021) for AI skills and Draca et al. (2022) for remote work.

To mitigate the risk of false positives in trait retrieval (Brenčič & McGee, 2023b), I used a two-step approach. First, I compiled dictionaries from authoritative works in the field, including Goldberg (1990), Costa Jr and McCrae (1992), Goldberg (1992), McCrae and John (1992), John (1990), Saucier (1994) and Heckman and Kautz (2012). For four of the Big Five traits, I used only adjectives from these established dictionaries. However, for Emotional Stability, I supplemented the dictionary with additional job-relevant terms[8].

[7]Marinescu and Wolthoff (2020) found that job titles have crucial information and play a more important role than wages in explaining the number and quality of applicants. Titles often indicate hierarchy, experience and specialisation.

[8]Psychology literature predominantly uses neuroticism-related language, which resulted in fewer suitable terms for measuring Emotional Stability in job contexts. I therefore added job vacancy terms closely related to Emotional Stability requirements, including *fortitude, pressure, resilience, unflappability, patience, conflict, and tense situations.* Terms like *fortitude, resilience, unflappable and patience* refer to

Second, I refined these dictionaries through manual analysis to ensure that, in job vacancy contexts, the words actually referred to employee or job characteristics.

I analysed the context of each dictionary word using natural language processing tools. For each word, I examined up to eight rounds of 100 randomly selected sentences from job advertisements containing the word. I analysed the 10 words before and after each trait-related term. This process identified words with multiple meanings, such as *flexible* or *trust*, which could refer to work schedules rather than personality traits. I modified terms in contexts not referring to worker or team characteristics before dictionary matching. For example, I replaced phrases like "flexible hours" with "flx hours" and "national trust" with "national tst" to avoid false positives. I repeated this process of checking contexts and redefining false matches until no spurious matches appeared in samples of 100 advertisements.

I removed words that, regardless of context, did not refer to personality traits in job vacancies, such as *modest, conventional, deliberate, order, aesthetic, deep, relaxed* or *comfortable*. These terms were excluded because they typically described job conditions, physical environments, or general preferences rather than specific personality characteristics that employers seek in candidates—for instance, "comfortable salary" or "relaxed atmosphere" do not indicate personality requirements but rather job attributes. Table 1 contains the terms used for each of the Big Five and the source they were retrieved from.

4.1.1 Examples of Classification

Table 2 shows two job vacancies and their classification using the Boolean retrieval approach.

positive behavioural traits, while *pressure, conflict, and tense situations* appear in phrases where applicants must work well *under pressure*, in *conflictive* environments, or handle *tense situations*, demonstrating emotional stability requirements.

Table 1: **Dictionary of Personality Traits**

Agreeableness	**Extraversion**	**Conscientiousness**	**Emotional Stability**	**Openness**
affection (4, 5)	adventurous (3, 5, 7)	ambitious (2, 7)	at ease (3)	artistic (3 - 7)
agreeable (all)	assertive (1, 3 - 7)	careful (6)	calm (2, 3, 5)	bright (3, 6)
altruist (5, 7)	bold (3, 6)	competence (5, 7)	emotional (all)	clever (5)
amiable (1)	candour (1)	conscientious (all)	hostil (5, 7)	creative (1 - 3, 6)
considerate (3, 6)	confident (7)	decisive (1)	fearful (6)	curious (1, 2 - 5, 7)
cooperative (1, 3, 5 - 7)	courage (1)	dependable (1, 5)	tempered (2, 5, 6)	ideas (5, 7)
courteous (1)	daring (3, 6)	dutiful (7)	fortitude (ads)	imaginative (2 - 7)
empathy (1)	energetic (1, 3 - 7)	efficient (1, 4 - 7)	pressure (ads)	independent-mind (5)
forgiving (4, 5, 7)	enthusiastic (4, 5, 7)	hardwork (2, 3)	resilience (ads)	ingenious (5)
gentle (5)	expressive (1)	logic (1)	tense (4, 5)	innovative (3, 6)
good-nature (2, 5)	extraverted (all)	neat (3, 7)	unflappable (ads)	insightful (1, 4, 5)
helpful (3, 5, 6)	forceful (5)	organised (1, 3 - 7)	patience (ads)	intellectual (1, 5, 6)
kind (3 - 6)	friendly (7)	painstaking (5)	pressure (ads)	intelligent (1, 3, 5)
naturalness (1)	fun (2)	persevering (2)	-	introspective (3, 6)
pleasant (3, 5, 6)	gregarious (1, 5, 7)	persistence (1)	-	inventive (5)
sensitive (5)	humour (1)	planful (4, 5)	-	openness (all)
straight-forward (5, 7)	optimist (1)	precise (1, 5)	-	original (2, 4)
sympathetic (3 - 7)	outgoing (4, 5)	prompt (3, 6)	-	prefer variety (2)
trust (2, 3 - 7)	outspoken (5)	punctual (1, 2)	-	reflective (3)
unselfish (3, 5, 7)	passionate (2)	reliable (4, 5)	-	sharp (5)
warm (1, 3, 5 - 7)	playful (1)	responsible (3 - 5)	-	unconventional (5, 7)
	pro-active (2 - 7)	striving (5, 7)	-	wide interests (4, 5, 7)
	sociable (5, 7)	systematic (3, 7)	-	witted (5)
	spontaneous (1)	thorough (3 - 5, 7)	-	-
	talkative (1 - 6)		- -	-

Notes: The numbers denote which dictionaries the words are mentioned in. The numbers correspond to the following dictionaries: (1) Goldberg (1990); (2) Costa Jr and McCrae (1992); (3) Goldberg (1992); (4) McCrae and John (1992); (5) O. P. John et al. (1999); (6) Saucier (1994); (7) Heckman and Kautz (2012)

Table 2: **Two examples of job vacancies and the boolean indicator.**

Vacancy 1

Our comprehensive training will give you the skills and knowledge you need, so our main focus is that you are professional, ***kind*** *and* ***empathetic*** *to the needs of our service users. You will need to be* ***even-tempered****, willing to carry out personal hygiene tasks and be dedicated to providing outstanding levels of care.*

Vacancy 2

This chef apprenticeship offers a fantastic opportunity to learn and develop your skills and knowledge in a busy commercial kitchen. You will need to have a ***passion*** *for food and a good work ethic. You will need to be* ***organised*** *and have the ability to work under pressure. You must maintain a* ***pro active*** *attitude, seek to improve your skills and achieve personal development.*

Boolean categorization

Personality Trait	Vacancy 1	Vacancy 2
Agreeableness	1	0
Conscientiousness	0	1
Extraversion	0	1
Openness	0	0
Emotional Stability	1	0

4.2 AI demand retrieval

To measure AI skills demand in job postings, I use text analysis based on work by Acemoglu et al. (2022), Babina et al. (2024) and Choi and Leigh (2024). These studies use different methods to identify AI-related terms. Babina et al. (2024) and Choi and Leigh (2024) employ a co-occurrence approach using pre-classified skills in BGT data[9]. Acemoglu et al. (2022) uses a dictionary-based method. I expand on the dictionary-based approach, which works better for raw text analysis.

I extended the AI skills list from Acemoglu et al. (2022) and Babina et al. (2024) to include more AI-related terms. I grouped similar technology terms to ensure comprehensive retrieval of AI-related content in job postings. For example, when searching for *machine learning* requirements, I also searched for related specific techniques such as *cross-domain learning, federated learning,* and *reinforcement learning*, treating all these terms as indicators of machine learning skill demands. To expand the term list, I manually

[9] Babina et al. (2024) and Choi and Leigh (2024) identify AI skills by calculating how often skills appear alongside core AI terms ("Machine Learning", "Artificial Intelligence", "Natural Language Processing", and "Computer Vision") in BGT job postings. Skills frequently appearing with these core terms are classified as AI-related. This method uses BGT's pre-classified skill categories, unavailable in raw job posting text.

examined 100 advertisements for each term from Acemoglu et al. (2022) and Babina et al. (2024), analysing surrounding AI-related skills or tools. The complete list of skills from each source and text replacements is in Appendix B.

This manual review process, though more time-consuming than the co-occurrence method used by Babina et al. (2024) or Choi and Leigh (2024), serves a similar purpose of identifying related AI terms in context. Since Adzuna provides raw text rather than listed skills, a co-occurrence approach would consider all vacancy words rather than just skills, making it less reliable for identifying AI terms. By examining over 2,000 vacancies with AI-related content, my approach identified additional AI tools and software packages that might not be included in structured databases like BGT. I carefully ensured these terms were specific to AI and excluded broader IT functions.

5 Validation

To validate my text-based approach for identifying personality trait demands, I first test whether these demands align with workers' actual characteristics in the same occupations. If job advertisements effectively signal personality requirements, it would be expected that assortative matching—occupations with higher demands for specific traits would employ workers who score higher on those same traits in surveys. This analysis also helps confirm that personality-related language in job postings reflects genuine trait requirements rather than just differences in advertising styles.

I compare trait requirements in job advertisements with workers' self-reported traits from the UK Household Longitudinal Study (UKHLS)[10].

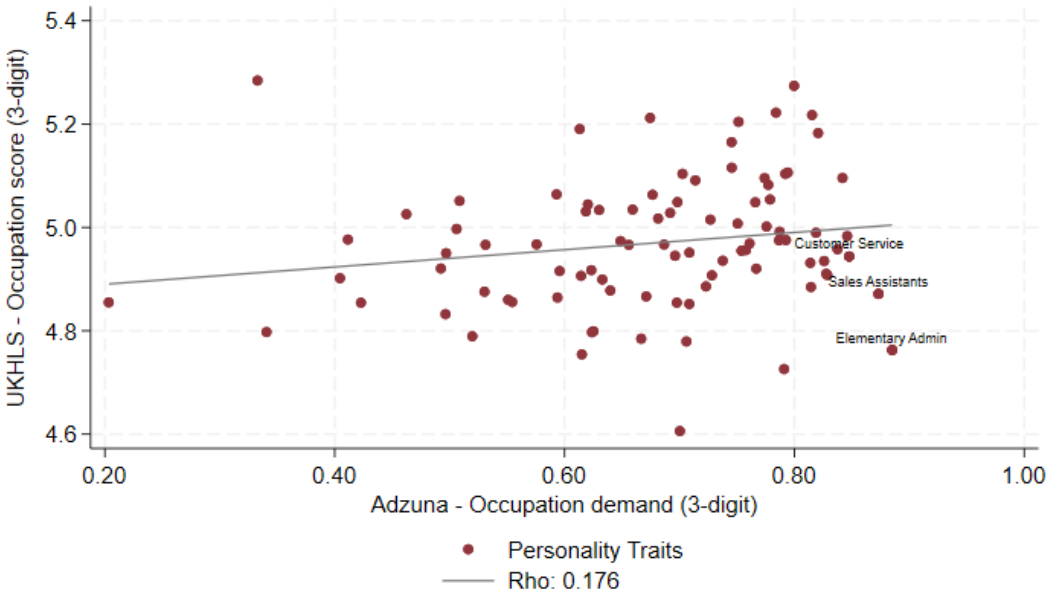
The scatter plots in Figure 1 show positive assortative matching for most personality traits: occupations with high trait demands tend to employ workers who report higher scores in those same traits. This correlation is particularly strong for Agreeableness (ρ = 0.464). The top 10 occupations with the highest self-reported Agreeableness scores demand this trait in 18 per cent of their vacancies on average, while the bottom 10 occupations demand it in only 5 per cent. The occupation with the highest trait demand (39.1 per cent of vacancies) also ranks among the top quartile for self-reported scores, exemplifying this strong assortative matching between job advertisement requirements and workers' self-reported traits[11].

[10]These range from 1 to 7, where 1 means *does not apply to me at all* and 7 means *applies to me perfectly*. Workers' self-assessments are averaged to create occupation-level mean scores for each trait.

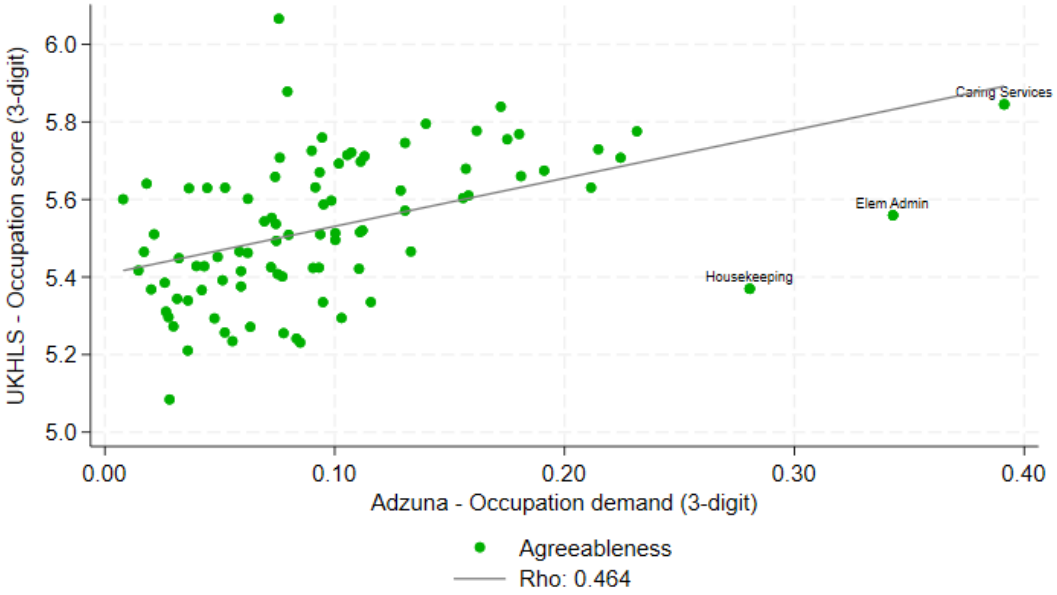
[11]Linear probability regressions without controls show that one standard deviation increase in the average occupation score for overall personality traits raises the probability of an advertisement mentioning any personality trait by 2.3 per cent, Agreeableness by 8.9 per cent, Conscientiousness by 0.1 per cent, Emotional Stability by -4.4 per cent, Extraversion by 8.9 per cent and Openness by 4.5 per cent. The regression output is in Table A.4, Appendix A.

Figure 1: Scatter-plot between mean trait's standardized scores in UKHLS and share of vacancies that mention Personality traits in Adzuna.

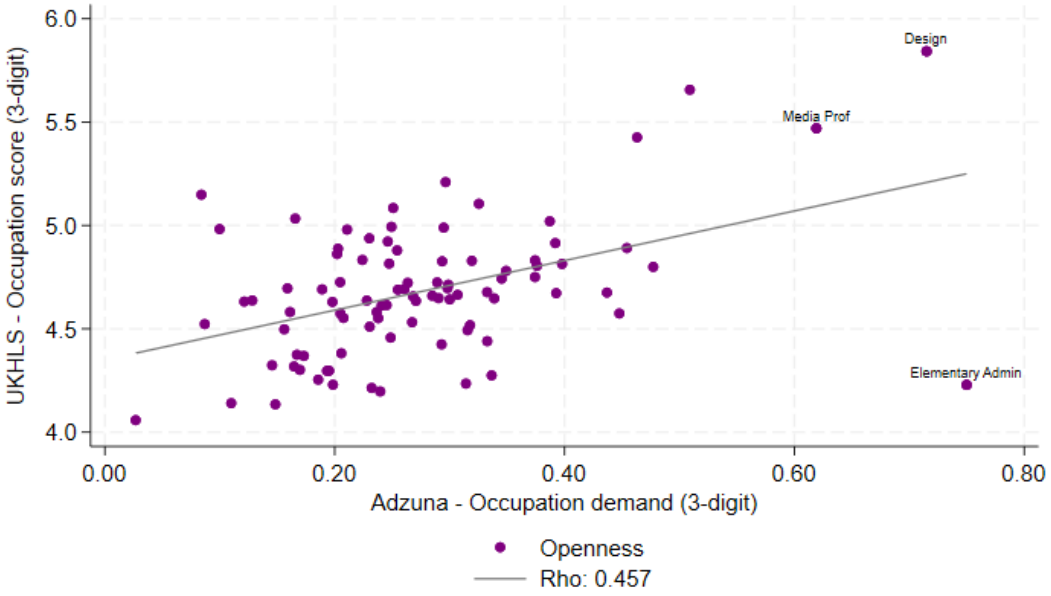
(a) All Personality Traits



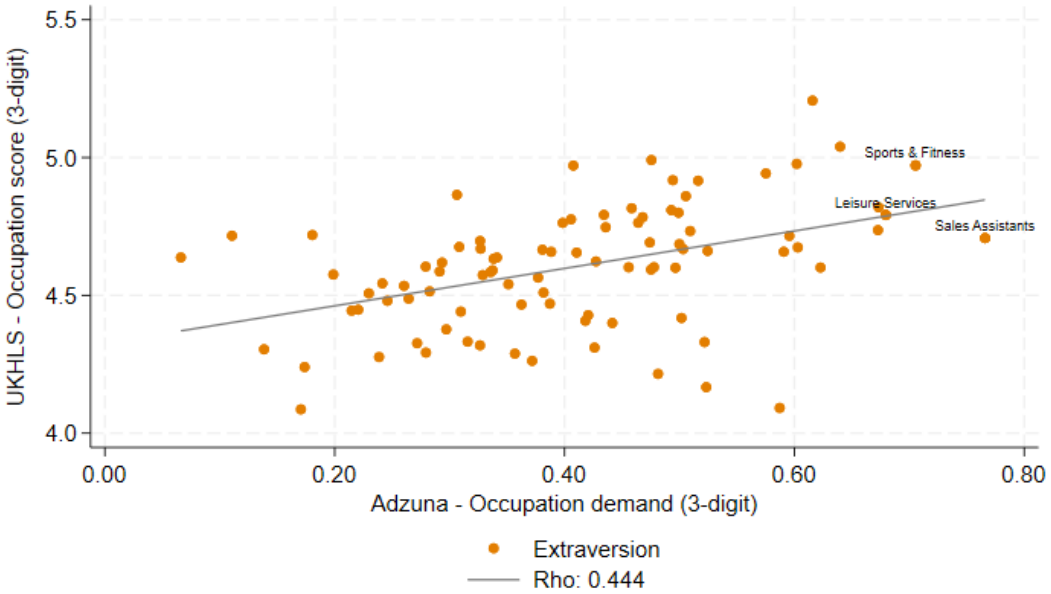
(b) Agreeableness



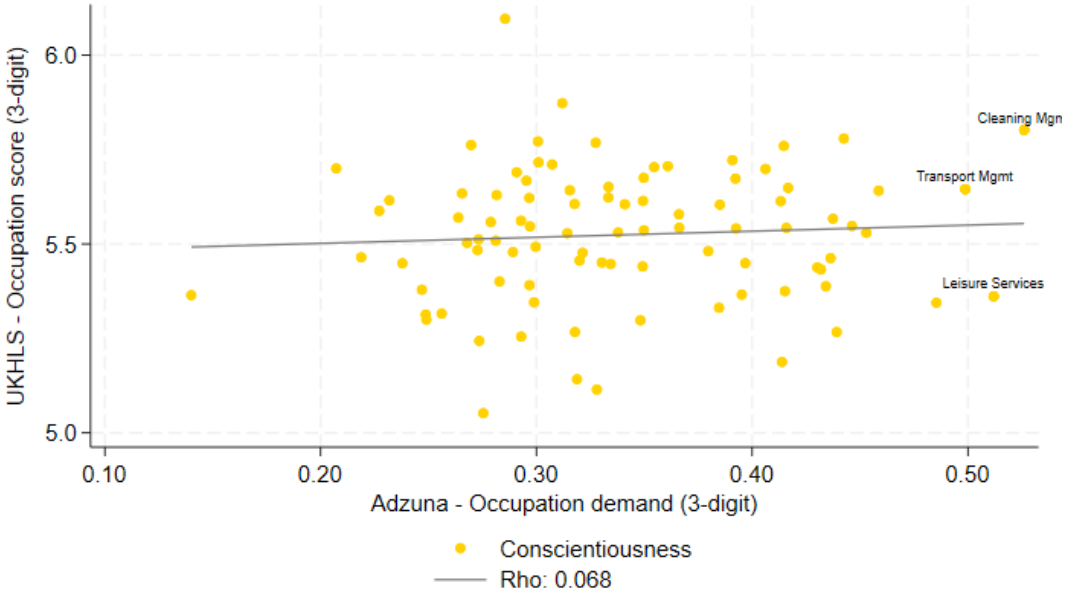
(c) Openness



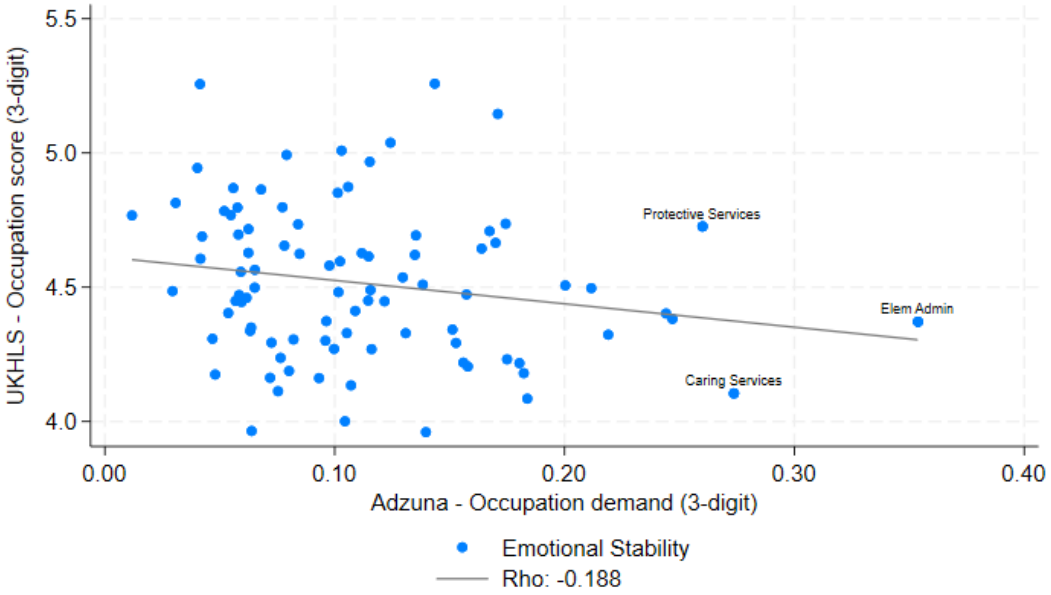
(d) Extraversion



(e) Conscientiousness



(f) Emotional Stability



Notes: Author's creation using UKHLS and Adzuna.

The relationship varies across traits and occupations. For Openness, occupations with the highest demands (exceeding 60 per cent of vacancies) also exhibit correspondingly high self-reported scores, reflecting the creative nature of these roles. This pattern reinforces the validity of using job postings to measure personality trait requirements.

Emotional Stability presents an exception to the assortative matching pattern. While some occupations frequently demand Emotional Stability (over 25 per cent of vacancies), the highest self-reported scores appear in different occupation groups entirely. This discrepancy likely stems from two measurement issues. First, job postings emphasise situational requirements (e.g., "ability to work under pressure"), while self-reported scores reflect broader personality tendencies. Second, the self-reported scores come from UKHLS questions about neuroticism that were inversely transformed to approximate Emotional Stability. This transformation may not capture the same construct that employers seek when requesting emotional stability, as employers may use different terminology or find it difficult to articulate these requirements in job postings.

To understand which occupation-level characteristics relate to each traits demand, I estimate a linear probability model following Equation 1:

$$P(Trait_i = 1) = \alpha + \beta_0 + \beta_1' Z_k + \beta_2 S_k + \beta_3 X_i + \delta_l + \theta_t + \phi_j + \epsilon_i \tag{1}$$

where Z_k is a vector of UKHLS occupation-level characteristics including gender composition, education levels, migrant share and mean worker age, S_k represents the standardised UKHLS trait score for occupation k, and X_i captures job posting characteristics such as advertisement length and other requirements. The model includes location (δ_l), time (θ_t), and firm (ϕ_j) fixed effects to account for geographic, temporal, and employer-specific factors that might affect trait demands.

Table 3 shows how occupation-level characteristics influence trait demands. Workforce education level shows a consistent negative relationship with trait demands: occupations with more university graduates are less likely to mention personality requirements (-0.8 to -1.9 percentage points across traits). Female-dominated occupations show higher demands for Extraversion (2.6 percentage points) and Emotional Stability (2.1 percentage points). Occupations with more migrant workers show mixed effects: positive for Emotional Stability (0.7 percentage points) and Openness (1.6 percentage points), but negative for Conscientiousness (-0.9 percentage points). Age composition negatively affects both overall personality traits (-1.1 percentage points) and Extraversion (-1.6 percentage points).

Table 3: Assortative matching: Standarised occupation characteristics effects on personality trait demand

	(1) P(All P.T.)	(2) P(Agree.)	(3) P(Consc.)	(4) P(Emot.)	(5) P(Extra.)	(6) P(Open.)
Women share	0.017***	−0.015	−0.006	0.021**	0.024***	−0.002
	(0.006)	(0.013)	(0.004)	(0.011)	(0.004)	(0.005)
Degree share	−0.014**	−0.019**	−0.012**	−0.012**	−0.008*	−0.010**
	(0.007)	(0.009)	(0.005)	(0.005)	(0.004)	(0.004)
Migrant share	−0.001	0.000	−0.009**	0.007***	−0.002	0.016***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Management share	−0.001	0.005	0.005	−0.002	−0.009*	−0.001
	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
Age mean	−0.009**	−0.007	−0.002	0.001	−0.016***	0.004
	(0.005)	(0.005)	(0.006)	(0.003)	(0.004)	(0.004)
Tenure mean	−0.009	−0.016	0.010	0.007	−0.008**	−0.003
	(0.008)	(0.012)	(0.007)	(0.005)	(0.003)	(0.006)
R^2	0.346	0.357	0.274	0.283	0.339	0.326

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N = 9,888,315. "Any P.T." refers to Any Personality Trait, "Consc." to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. All variables refer to the occupation-level 2011-2017 UKHLS standardized value. Regression controls for the standardised UKHLS trait score, traits and skills demand, remote work requirements, length of advertisement, contract type, work schedule, bonus offerings, education, and negotiable salary option. Location, year and firm fixed effects. SE clustered at the firm level. Author's creation using Adzuna and UKHLS.

6 Empirical design

I use two empirical approaches to study how AI adoption affects personality trait demands: a high-dimensional fixed effects linear probability model[12] and a staggered difference-in-differences (DID) design[13] following Callaway and Sant'Anna (2021). The first approach compares AI and non-AI vacancies within firms and occupations. The second tracks the same firms before and after their first AI hire, using firms that have not yet adopted AI as the comparison group. To make treated and control firms comparable on observable characteristics, I apply entropy balancing as a pre-processing step within the DID framework

Personality-related words in job advertisements may correlate with unobserved characteristics that could bias estimates. For instance, personality trait demands might relate to occupation-specific skill requirements, industry practices, or local labour market conditions. Firms that emphasise certain personality traits might also offer different compensation structures, working conditions, or advancement opportunities. Even within the same occupation and location, firm-specific HR practices could influence both job posting language and actual job requirements.

This model controls for these sources of unobserved heterogeneity across multiple dimensions, reducing potential bias. I also account for observable job characteristics

[12]Implemented using the STATA command *reghdfe* from (Correia, 2016).
[13]Implemented using the STATA *csdid* command from (Callaway & Sant'Anna, 2021).

that might correlate with personality demands, including contract type, work schedule, seniority level, required education, experience, and performance-based compensation. By isolating within-location, within-occupation, and within-firm variation while controlling for these observable characteristics, this method reduces bias in estimates of how personality requirements relate to other job characteristics, particularly AI demand.

To examine whether personality traits appear together in job postings and how they relate to other job requirements, I estimate:

$$P(Trait_i = 1) = \beta_0 + \beta_1 X_i + \delta_l + \theta_t + \phi_j + \gamma_k + \epsilon_i \tag{2}$$

where X_i is a vector of job posting characteristics including the demand for other personality traits, δ_l are location fixed effects, θ_t are month-year fixed effects, ϕ_j are firm fixed effects and γ_k are occupation fixed effects. To measure how these relationships change over time, I use a variation of Equation 2 where I replace month-year fixed effects with year dummies:

$$P(Trait_i = 1) = \beta_0 + \beta_1 X_i + \sum_{t=2018}^{2022} \alpha_t Year_t + \gamma_k + \phi_j + \epsilon_i \tag{3}$$

where $Year_t$ are dummy variables (with 2017 as the base year).

To study how technological change affects personality demands, I examine the relationship between AI adoption and personality requirements[14]. This analysis is relevant given AI's rapid adoption across industries and the interest on how AI might complement or substitute for human traits. While research has examined AI's impact on technical skill requirements (Acemoglu et al., 2022) and cognitive tasks (Babina et al., 2024), little is known about its effects on personality trait demands. Johnson and Acemoglu (2023) suggest firms adopting AI technologies might increasingly value workers with personality characteristics that complement AI capabilities. To test these relationships, I estimate:

$$P(Trait_i = 1) = \beta_0 + \beta_1 X_i + \beta_2 AI_i + \delta_l + \theta_t + \phi_j + \gamma_k + \epsilon_i \tag{4}$$

where X_i is a vector of job posting characteristics, δ_l are location fixed effects, θ_t are time fixed effects, ϕ_j are firm fixed effects and γ_k are occupation fixed effects. I show how AI coefficients change when incrementally introducing each fixed effect.

While the fixed effects analysis shows how AI adoption relates to personality demands, I complement this with a staggered difference-in-differences strategy that uses variation in the timing of firms' first AI adoption, following Callaway and Sant'Anna (2021)'s methodology for multiple time periods.

[14] All specifications use Linear Probability Models. While these have known limitations such as predicted probabilities potentially outside the [0,1] range, this approach offers easier interpretation and avoids the incidental parameter problem in non-linear models with multiple fixed effects (Greene, 2004).

This approach's key innovation is using never-treated and not-yet-treated units as controls, which helps avoid bias from treatment effect heterogeneity in traditional two-way fixed effects models. For each firm, I identify the day they first post a job advertisement requiring AI skills as their treatment timing, allowing me to construct quarter, half-year and annual treatment variables. This staggered adoption design compares changes in personality trait demands between firms adopting AI at different times while using both never-treated and not-yet-treated firms as controls.

A key requirement for this identification strategy is that AI adoption was minimal before 2017. Alekseeva et al. (2021) document that AI-related job postings accounted for just 0.02 per cent of all online postings in 2010, rising gradually to 0.16 per cent by 2019. In the Adzuna data, fewer than 0.3 per cent of vacancies mention AI skills in 2017, consistent with this evidence of early-stage adoption at the start of the sample period. Section 7.2 provides further verification of this pattern.

This near-absence of AI adoption before 2017 has two implications for identification. First, it means that the pre-treatment period captures a genuine "before" state, where firms had not yet begun to reorganise hiring around AI capabilities. Second, because personality trait demands were themselves relatively stable across firms before AI adoption became widespread, the entropy balancing approach, which matches treated and control firms on their pre-treatment personality demands, firm size, profit, and occupational composition, is more likely to achieve balance on the unobserved determinants of personality requirements that could confound the post-adoption comparisons. Together, the minimal pre-2017 AI adoption and the stable pre-treatment personality demand patterns across treated and control firms support the parallel trends assumption underlying the staggered difference-in-differences design.

Let $Y_{i,t}$ denote the proportion of vacancies requiring a personality trait at time t for firm i. Let $G_{i,g} = 1$ indicate that firm i is first treated in period g. Following Callaway and Sant'Anna (2021), the group-time average treatment effects are:

$$ATT(g,t) = E[Y_{i,t}(g) - Y_{i,t}(0)|G_i = g] \tag{5}$$

where $Y_{i,t}(g)$ represents the proportion of vacancies requiring a personality trait when first treated in period g, and $Y_{i,t}(0)$ represents the proportion had the firm not adopted AI. Under the parallel trends assumption, this parameter is identified by:

$$ATT(g,t) = E[Y_{i,t} - Y_{i,g-1}|G_i = g] - E[Y_{i,t} - Y_{i,g-1}|G_i \in G_{comp}] \tag{6}$$

where G_{comp} consists of not-yet-treated units. For pre-treatment periods, G_{comp} includes

all cohorts not yet treated at time t, rather than only those not treated at time g[15]. The estimation uses the doubly-robust inverse probability weighted estimator with standard errors clustered at the firm level.

The model is estimated using three specifications: a base specification without controls, a specification that includes pre-treatment covariates as control variables, and a specification that uses entropy balancing weights. Let X_i^{pre} denote the vector of pre-treatment covariates for firm i, which includes: (1) the average pre-treatment demand for personality traits and skills, (2) pre-treatment firm characteristics (firm size measured by number of employees, profit and whether the firm is publicly quoted), and (3) pre-treatment occupational composition at the 1-digit SOC level. The covariates specification includes X_i^{pre} directly in the estimation, functioning as a regression adjustment approach. The entropy balancing specification instead constructs weights w_i for control units such that the weighted distribution of X_i^{pre} in the control group matches the distribution in the treated group in both mean and variance. These weights are estimated using Hainmueller (2012)'s method, targeting first and second moments of the pre-treatment covariates[16].

This model addresses several identification and estimation threats. First, it avoids the "negative weights" problem in TWFE models by using clean comparison groups. Second, entropy balancing addresses concerns about selection into AI adoption based on observable characteristics by ensuring treated and control firms are comparable across pre-treatment covariates. Third, using both never-treated and not-yet-treated comparison units maximises statistical power while maintaining comparison group integrity.

While balancing weights ensure treated and control firms are comparable across covariates before treatment, the DID framework accounts for time-invariant unobserved differences between firms and time-varying unobservables affecting groups differently. This allows for parsimonious estimation while accounting for selection on observables through matching procedures.

I test the parallel trends assumption validity by examining pre-treatment effects, which should be statistically indistinguishable from zero if the assumption holds. I also use event study plots to visualise both pre-treatment trends and post-treatment effects over time.

The data are analysed at the firm-period level, with each observation representing a firm's average personality trait demands in a given period (year, half-year or quarter). Although the original data are weekly, aggregating to longer periods allows more precise

[15]This distinction in the comparison group construction for pre-treatment periods is implemented using the `asinr` option in the `csdid` command.

[16]Unlike propensity score methods that model the probability of treatment assignment, entropy balancing directly reweights control units to achieve exact balance on specified moments of pre-treatment covariates, without requiring an intermediate propensity score estimation step.

treatment timing identification while maintaining sufficient observations for each firm and minimising anticipation effects. The staggered adoption timing combined with matching on pre-treatment characteristics helps address endogeneity concerns about firms' decisions to adopt AI technologies.

Together, the fixed effects and staggered DiD models provide complementary evidence on how personality demands in job advertisements respond to firm AI adoption, capturing both immediate within-firm effects and dynamic changes in requirements following adoption.

7 Results

I present results in four steps. I begin with summary statistics on the prevalence and occupational distribution of personality trait demands (Section 7.1). I then examine how these demands have changed over time, progressively controlling for compositional shifts in occupations and firms (Section 7.1). Next, I turn to the vacancy-level relationship between AI skill requirements and personality trait demands, showing how estimates change as firm and occupation fixed effects are introduced (Section 7.2). Finally, I present the firm-level and staggered difference-in-differences analyses, which examine how firms adjust their overall personality demands after adopting AI for the first time (Sections 7.3 and 7.4).

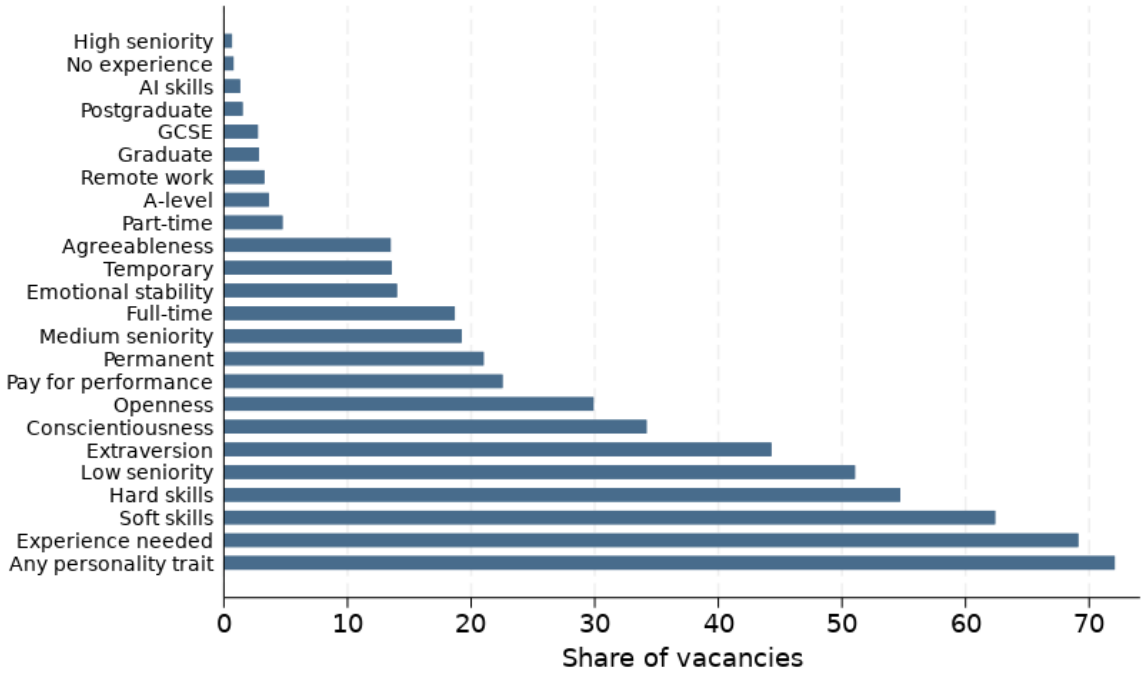
Analysis of 11.7 million UK job advertisements shows that 72 per cent contain at least one personality trait-related word (Table A.2 in Appendix A). Extraversion appears most frequently, in 44 per cent of advertisements, followed by Conscientiousness (34 per cent), Openness (30 per cent), Emotional Stability (14 per cent), and Agreeableness (13 per cent). Figures A.1 to A.5 in Appendix A show the terms within each trait dictionary. For Extraversion, "passion" and "friendly" are most common. For other traits, the most frequent terms are "innovative" and "flexible" (Openness), "efficient" and "ambitious" (Conscientiousness), "work under pressure" and "emotional" (Emotional Stability), and "warm" and "trustful" (Agreeableness).

Personality traits appear more often in job postings than traditional human capital measures (Figure 1). Soft skills and hard skills are mentioned in 62 per cent and 55 per cent of advertisements, while experience appears in 69 per cent and seniority in 70 per cent. Only 11 per cent explicitly mention education requirements - roughly the same percentage that mention being passionate. This low figure aligns with recent research by Bone et al. (2024), who found that only about 12 per cent of UK job postings explicitly state formal educational requirements[17]. While education requirements might be implied

[17]This low education demand is specific to the UK online labour market and contrasts with nearly 50 per cent in US online vacancies (Stahle, 2024). Several factors may contribute to this difference,

through job titles, occupational categories, or task descriptions, employers clearly choose to emphasise personality characteristics in their recruitment.

Figure 1: Frequency of Job Requirements



Notes: Author's creation using Adzuna.

Personality trait demands vary systematically across occupations[18]. Administrative and customer-facing roles show the highest trait demands: Elementary Administration (88.6 per cent), Sales Assistants and Retail Cashiers (87.3 per cent), and Customer Service (85.3 per cent). In contrast, manual and construction-related occupations show much lower demands: Elementary Construction (20.2 per cent), Metal Forming and Welding Trades (33.1 per cent), and Building Finishing Trades (33.9 per cent). This pattern aligns with theories about the importance of non-cognitive attributes in jobs requiring frequent interpersonal interaction (Deming, 2017).

Analysis at the major occupational group level (1-digit SOC code) reveals clearer patterns in both overall trait demands and specific trait preferences (Figure 2). Sales and Customer Service occupations have the highest overall trait demands (85.3 per cent), with particular emphasis on Extraversion (66.8 per cent) and Conscientiousness (41.1 per cent), reflecting needs for both social interaction and organisational skills. Caring

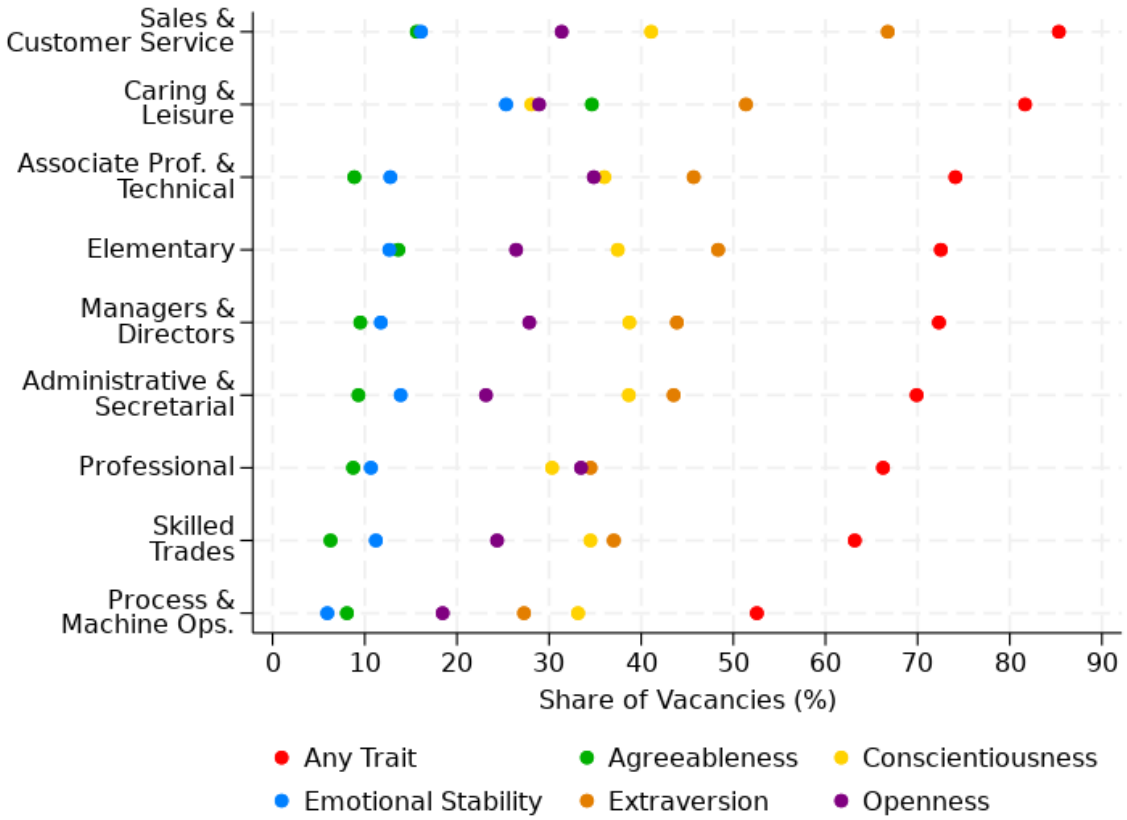
including different recruitment practices between countries and variations in how education requirements are communicated during hiring. While formal qualifications might be verified later in the recruitment process, there may also be an emerging shift toward skills-based hiring in certain sectors (Bone et al., 2024).

[18]Demand for each 3-digit SOC group is shown in Table A.3 in Appendix A.

and Leisure Services follow with 81.7 per cent overall trait demand and show the highest requirements for Agreeableness (34.7 per cent) and Emotional Stability (25.3 per cent) among all groups, consistent with the empathetic and emotionally demanding nature of care work. Associate Professional and Technical occupations show balanced demands across traits, with notably high requirements for Openness (34.9 per cent), suggesting the value placed on adaptability and creativity in these roles.

In contrast, Process and Machine Operatives show the lowest overall trait demands (52.6 per cent), with particularly low requirements for Emotional Stability (5.9 per cent) and Openness (18.5 per cent). Roles focused on routine technical tasks place less emphasis on personality characteristics, although Conscientiousness remains relatively important (33.1 per cent) even in these occupations, reflecting the general need for efficiency and attention to detail across job types.

Figure 2: Personality Trait demand by 1-digit SOC group



Notes: Author's creation using Adzuna.

These occupational variations in trait demands align with task-based job design and skill-biased technological change frameworks. Jobs involving complex social interactions or non-routine cognitive tasks show higher demands for personality traits, while those focused on routine or manual tasks show lower requirements. This supports the literature on the growing importance of social and personality characteristics in an increasingly

service-oriented economy (Borghans et al., 2014).

Further analysis shows these occupational patterns align with skill requirements. When ranked by hard skill requirements (Figure A.6a in Appendix A), occupations show relatively stable demands for Extraversion (average 41 per cent) and Conscientiousness (average 34 per cent), suggesting these traits are valued regardless of technical skill intensity. However, Openness demand increases notably in occupations with more hard skill requirements, rising from an average of 24 per cent in the bottom 20 occupations to over 34 per cent in the top 20 hard-skill intensive occupations, reflecting the need for adaptability and creativity in complex technical jobs. When occupations are ranked by soft skills requirements (Figure A.6b in Appendix A), there is a clear positive relationship for most traits. Extraversion demand shows the strongest increase, rising from 28 per cent in the bottom 20 to over 45 per cent in the top 20 soft skill-intensive activities. Openness and Conscientiousness also show positive trends, though less pronounced. While certain traits (like Extraversion and Conscientiousness) are broadly valued across the skill spectrum, others serve specific functions depending on occupational characteristics: Openness complements occupations with greater technical skills, while the full range of personality traits becomes increasingly important as jobs require more interpersonal interaction. These relationships are formally tested in Equation 2.

7.1 Personality Traits demand and their evolution

7.1.1 Complementarities in Trait Demands

Understanding how personality traits are demanded together provides insight into how employers conceptualise the multidimensional nature of personality in the workplace. Research has studied individual personality traits' effects on labour market outcome, such as Flinn et al. (2025) finding that conscientiousness and emotional stability predict job stability, and Gensowski (2018) showing openness and extraversion affect lifetime earnings. However, little is known about how these traits are demanded in combination. This analysis helps show whether employers see certain personality characteristics as complementary or substitutable, similar to the relationship between cognitive and social skills identified by Deming (2017).

Table 4 shows the relationships between different personality trait requirements. The results show significant complementarities in trait demands, even after controlling for occupation fixed effects and other job characteristics. Job postings mentioning one personality trait are consistently more likely to mention others, though relationship strengths vary across trait pairs.

Both Conscientiousness and Extraversion show positive relationships with other traits: advertisements mentioning Conscientiousness are 5 percentage points more likely to

demand Extraversion and 4.1 percentage points more likely to demand Openness. Those mentioning Extraversion are significantly associated with all other traits, with effects ranging from 4.5 to 8.4 percentage points. The strongest complementarities are between Extraversion and Emotional Stability (8.4 percentage points), Openness (7.8 percentage points), and Agreeableness (7.8 percentage points).

These complementarity patterns suggest employers tend to demand clusters of traits rather than isolated characteristics. These relationships persist even after controlling for occupation and firm fixed effects, indicating that trait clustering reflects broader requirements rather than just occupational differences or firm-specific HR strategies. This aligns with Rohrbach-Schmidt et al. (2023), who found that trait combinations are valued differently in specific job contexts.

Table 4: Coefficients between traits demands

	(1) P(Agree.)	(2) P(Consc.)	(3) P(Emot.)	(4) P(Extra.)	(5) P(Open.)
Agreeableness		0.015	0.048**	0.078***	0.020**
		(0.014)	(0.019)	(0.021)	(0.009)
Conscientiousness	0.007		0.027***	0.050***	0.041***
	(0.007)		(0.010)	(0.010)	(0.008)
Emotional Stability	0.043***	0.050***		0.084***	0.009
	(0.015)	(0.018)		(0.016)	(0.012)
Extraversion	0.037***	0.050***	0.045***		0.067***
	(0.008)	(0.010)	(0.011)		(0.006)
Openness	0.011**	0.048***	0.006	0.078***	
	(0.005)	(0.009)	(0.007)	(0.007)	
R^2	0.364	0.281	0.287	0.342	0.331

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N = 9,888,315. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Regressions control for traits and skills demand, remote work requirements, length of advertisement, contract type, work schedule, bonus offerings, education, and negotiable salary option. Location, year, occupation and firm fixed effects. SE clustered at the occupation and firm level. Author's creation using Adzuna.

7.1.2 Relationship with Other Job Requirements

Beyond understanding trait combinations, examining how personality traits relate to other job requirements reveals how employers package worker attributes. Deming and Kahn (2018) provides evidence about strong associations between cognitive and social skills in job postings. However, little is known about how specific job requirements relate to the Big Five in actual hiring. Understanding these patterns shows how firms combine personality demands with other requirements to create job profiles, as suggested by Brenčič and McGee (2023b).

Table 5 shows the relationships between personality trait demands and other job requirements, controlling for occupation, location, month-year and firm fixed effects. The

results reveal systematic patterns in how employers combine personality traits with other job characteristics.

Table 5: Correlation with other job requirements

	(1) P(All P.T.)	(2) P(Agree.)	(3) P(Consc.)	(4) P(Emot.)	(5) P(Extra.)	(6) P(Open.)
Soft skill	0.073***	0.017	0.044***	0.032**	0.026**	0.032***
	(0.007)	(0.013)	(0.007)	(0.015)	(0.012)	(0.007)
Hard skill	0.021**	−0.013*	0.019**	−0.002	0.024**	0.008
	(0.009)	(0.007)	(0.009)	(0.008)	(0.011)	(0.007)
Remote work	−0.013**	−0.008	−0.013	−0.005	−0.009	0.010
	(0.006)	(0.006)	(0.014)	(0.005)	(0.012)	(0.015)
Pay for performance	0.016**	0.006	−0.005	0.004	0.037***	−0.017**
	(0.006)	(0.010)	(0.008)	(0.007)	(0.009)	(0.007)
Advert's length	0.260***	0.063***	0.168***	0.095***	0.156***	0.176***
	(0.009)	(0.005)	(0.009)	(0.020)	(0.009)	(0.012)
Offers negotiable salary	−0.004	−0.008**	−0.001	−0.008**	−0.002	−0.002
	(0.005)	(0.003)	(0.005)	(0.003)	(0.003)	(0.004)
Temporary job	−0.018***	−0.000	−0.014**	0.002	0.005	−0.016**
	(0.006)	(0.006)	(0.007)	(0.005)	(0.008)	(0.008)
Permanent job	−0.000	−0.000	−0.005	−0.002	0.011*	0.004
	(0.004)	(0.004)	(0.003)	(0.004)	(0.005)	(0.006)
Full time job	−0.003	−0.003	−0.010**	0.005	−0.000	−0.007
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.008)
Part time job	0.013*	0.019	−0.017	0.007	0.001	0.010
	(0.007)	(0.016)	(0.020)	(0.016)	(0.017)	(0.008)
High seniority	0.054***	0.018	0.025	0.018	0.015	−0.028
	(0.016)	(0.018)	(0.018)	(0.018)	(0.019)	(0.022)
Medium seniority	0.025**	−0.005	0.023*	−0.009	0.007	−0.011*
	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.006)
Low seniority	0.019	−0.009	0.021*	−0.006	0.002	−0.015***
	(0.014)	(0.010)	(0.012)	(0.009)	(0.008)	(0.005)
GCSE	−0.006	−0.011	0.012	0.030***	0.003	−0.027**
	(0.011)	(0.011)	(0.013)	(0.008)	(0.015)	(0.011)
A-level	−0.014	0.007	−0.009	−0.004	−0.018	−0.007
	(0.011)	(0.010)	(0.010)	(0.009)	(0.014)	(0.008)
Bachelor's degree	0.023**	−0.014***	0.029**	−0.011	0.022**	0.024***
	(0.011)	(0.004)	(0.014)	(0.012)	(0.009)	(0.007)
Posgraduate degree	−0.012	−0.012*	0.009	−0.018	−0.022*	0.023
	(0.009)	(0.007)	(0.023)	(0.027)	(0.011)	(0.027)
Request experience	−0.008	−0.011	−0.002	−0.006	−0.010	−0.014**
	(0.006)	(0.008)	(0.007)	(0.010)	(0.009)	(0.006)
No experience needed	−0.028	0.006	−0.060*	−0.004	−0.008	−0.033
	(0.043)	(0.026)	(0.032)	(0.020)	(0.040)	(0.028)
R^2	0.349	0.364	0.281	0.287	0.342	0.331

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N = 9,888,315. "Any P.T." refers to Any Personality Trait, "Consc." to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Regression controls for traits and skills demand, remote work requirements, length of advertisement, contract type, work schedule, bonus offerings, education, and negotiable salary option. Location, year, occupation and firm fixed effects. SE clustered at the occupation and firm level. Author's creation using Adzuna.

The most important finding is that soft skills requirements strongly predict personality trait demands, with advertisements mentioning soft skills being 7.3 percentage points more likely to demand personality traits overall. This relationship is consistent across all Big Five traits, with effects ranging from 2.6 percentage points for Extraversion to 4.4 percentage points for Conscientiousness. Hard skills requirements show weaker but still positive associations (2.1 percentage points overall). This pattern confirms that

employers view personality traits as complementary to interpersonal capabilities rather than technical competencies, supporting Woessmann (2024)' argument about modern workplace skill combinations.

Advertisement length is the strongest predictor of trait demands. The variable is measured as the natural logarithm of the number of characters, with a mean of 2,297 characters in the sample. The estimated coefficient of 0.260 implies that doubling the length of a job advertisement raises the probability of mentioning any personality trait by approximately 18 percentage points ($0.260 \times \ln 2 \approx 0.18$), indicating that detailed job descriptions systematically include more personality requirements.

Pay-for-performance schemes reveal strategic trait emphasis, increasing Extraversion demands by 3.7 percentage points while reducing Openness demands by 1.7 percentage points. This suggests firms align personality requirements with their incentive structures, with performance-based compensation favouring assertive traits over creative ones.

Organisational hierarchy systematically shapes trait demands. High seniority positions are 5.4 percentage points more likely to mention personality traits overall, while entry-level positions emphasise Conscientiousness (2.1-2.3 percentage points increase for low and medium seniority). Remote work reduces overall trait demands by 1.3 percentage points, and temporary positions show negative associations with Conscientiousness (-1.4 percentage points) and Openness (-1.6 percentage points), indicating that stable, office-based positions drive personality requirements.

Educational requirements reveal distinct trait patterns that reflect job complexity. Bachelor's degree requirements increase demands for Conscientiousness (2.9 percentage points), Extraversion (2.2 percentage points), and Openness (2.4 percentage points) while reducing Agreeableness demands (-1.4 percentage points). In contrast, GCSE-level positions emphasise Emotional Stability (3.0 percentage points) but reduce Openness demands (-2.7 percentage points). The strongest education effect occurs for no-experience positions, which are 6 percentage points less likely to demand Conscientiousness, suggesting that personality requirements substitute for work experience in signalling candidate quality.

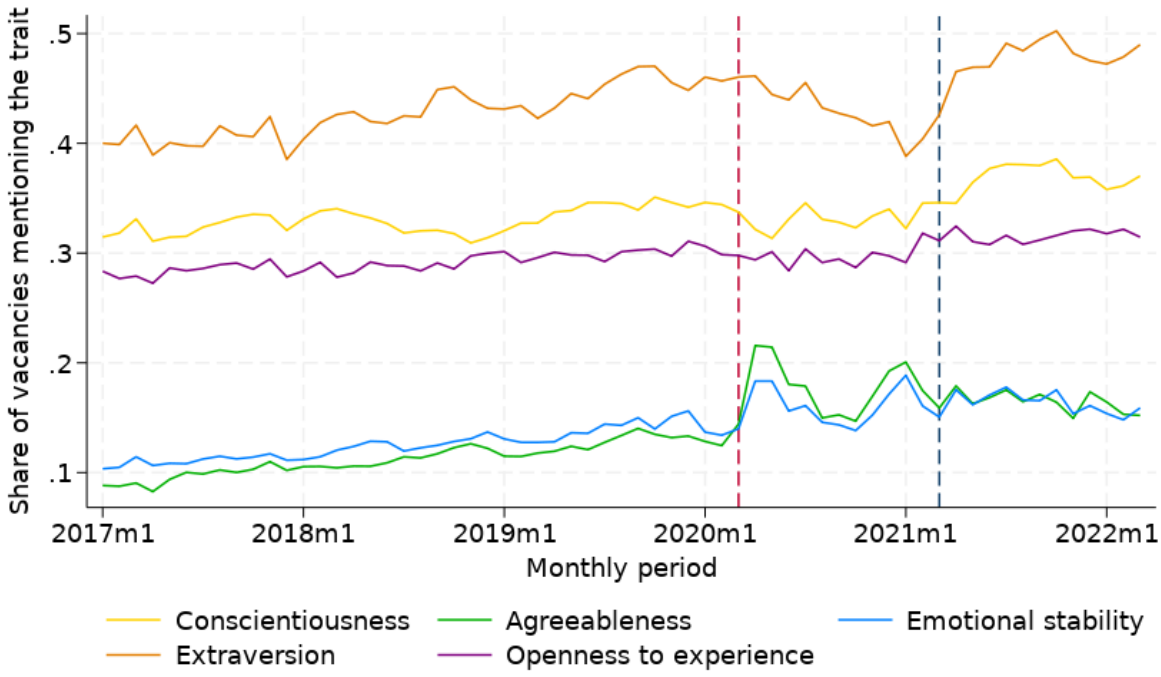
These findings reveal that personality trait demands reflect strategic employer choices rather than random job posting variation. Soft skills requirements and detailed job descriptions strongly predict trait mentions, while organisational factors (seniority, incentive structures, employment arrangements) systematically shape which traits employers emphasise. Educational requirements create distinct trait profiles, with higher education levels associated with greater emphasis on analytical and interpersonal traits. Together, these patterns demonstrate that personality demands are embedded within broader job design decisions and organisational practices.

7.1.3 Changes over time

Examining how personality trait demands change over time helps us understand evolving work and skill requirements in modern labour markets. These changes may reflect both economic structural shifts and changing employer preferences, especially during periods of technological change and economic disruption. This analysis covers 2017 to 2022, a period marked by the COVID-19 pandemic and accelerated workplace digitalisation.

With substantial increases in AI adoption during this period (Schmidt et al., 2024), analysing these trends provides context for understanding how technological change might affect personality requirements. The raw data shows increasing demand for personality traits (Figure 3). The share of vacancies mentioning at least one personality trait rose from 68 per cent in 2017 to 76 per cent in 2022. However, this overall change hides considerable variation across traits and time.

Figure 3: Change over time of the demand for personality traits.



Notes: Dashed lines show the beginning (red) and end (blue) of Covid-19 restrictions.
Author's creation using Adzuna.

To analyse how personality trait demands have evolved, I use the Linear Probability Model from Equation 3 to estimate time effects while progressively adding controls for job characteristics, location, industry, occupation, and firm fixed effects. Figures A.7 to A.12 in Appendix A show the monthly evolution of coefficients for each trait as these controls are added, revealing how compositional factors shape the observed patterns.

Table 6 shows how the estimated effect of 2022 (compared to 2017) changes with

different controls and fixed effects. Each row reports results from a separate regression, with the dependent variable being the personality trait indicator named in that row. The basic specification shows significant differences across most traits, but as fixed effects are added, many become statistically insignificant. In the most comprehensive model, only Conscientiousness shows a significant decline (-2.1 percentage points), while Extraversion maintains a positive trend (1.7 percentage points) that approaches significance.

Table 6: Effect of 2022 vs 2017 on Personality Traits: Progression across model specifications

	(1) Basic	(2) +Location	(3) +Industry	(4) +Occupation	(5) +Firm
Personality Traits	-0.002***	-0.001	-0.004	-0.005	-0.014
	(0.001)	(0.001)	(0.009)	(0.009)	(0.009)
Agreeableness	0.020***	0.018***	0.014	0.013	0.003
	(0.000)	(0.000)	(0.012)	(0.010)	(0.009)
Conscientiousness	-0.005***	-0.006***	-0.011	-0.016	-0.021**
	(0.001)	(0.001)	(0.012)	(0.010)	(0.011)
Emotional Stability	0.001***	0.000	-0.002	-0.001	0.006
	(0.000)	(0.000)	(0.008)	(0.008)	(0.008)
Extraversion	0.009***	0.013***	0.013	0.013	0.017
	(0.001)	(0.001)	(0.012)	(0.009)	(0.012)
Openness	-0.019***	-0.020***	-0.018	-0.012	-0.003
	(0.001)	(0.001)	(0.012)	(0.010)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes
Location FE	No	Yes	Yes	Yes	Yes
Industry FE	No	No	Yes	Yes	No
Occupation FE	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes
R-squared range	[0.082, 0.181]	[0.083, 0.181]	[0.103, 0.203]	[0.118, 0.223]	[0.280, 0.364]
Observations	11,662,124	9,923,701	9,923,697	9,923,697	9,888,315

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each row reports the coefficient on the 2022 year dummy from a separate linear probability model, where the dependent variable is an indicator for the personality trait named in that row. Each column adds a further set of fixed effects to the specification from the previous column. Baseline year is 2017. All regressions include controls for traits and skills demand, remote work requirements, length of advertisement, contract type, work schedule, bonus offerings, education, and negotiable salary option.

Controlling for compositional changes is particularly important during the COVID-19 period. The data shows a shift in occupational composition: Caring Personal Services increased sharply from about 10 per cent of vacancies in early 2020 to over 30 per cent by mid-2020, before stabilising at around 15 per cent in 2021-22. Other major occupational groups like Customer Service, Business Professionals, and Sales remained relatively stable, fluctuating between 5-10 per cent throughout the period (see Figure A.13 in Appendix A). This rise in care-related vacancies helps explain some observed changes in trait demands. For example, the apparent increase in Agreeableness (2.0 percentage points in the basic specification) disappears completely once the full set of fixed effects is included (0.3

percentage points, non-significant), suggesting it was driven by the increased proportion of caring roles rather than changing requirements within occupations.

The point estimates change modestly as fixed effects are added incrementally. The most consequential step is the inclusion of firm fixed effects in column (5): the standard errors fall substantially, rendering Conscientiousness statistically significant at the 5 per cent level. This reflects the large increase in explained variation—from an R^2 of around 0.22 to above 0.28—when within-firm variation is isolated. Occupational composition, while important for understanding trait demand levels, does not substantially alter the estimated trend coefficients once firm-level heterogeneity is also controlled for. However, the negative effect for Conscientiousness and the positive trend for Extraversion persist across specifications, suggesting these reflect genuine shifts in trait demands within occupations and firms.

After accounting for compositional changes in location, occupation, and firm, Table 7 shows the estimated year effects relative to 2017 for the full period. Conscientiousness shows a gradual decline, becoming significantly negative by 2022. Extraversion shows a consistent positive trend, with significant increases during 2019-2021 (ranging from 1.7 to 2.8 percentage points). Emotional Stability shows temporary positive effects during 2021 (1.1 percentage points), while Openness and Agreeableness show no significant time trends after controlling for composition.

Table 7: Year effect on each Personality Trait. Baseline: 2017

	(1) P(All P.T.)	(2) P(Agree.)	(3) P(Consc.)	(4) P(Emot.)	(5) P(Extra.)	(6) P(Open.)
2018	−0.003	0.002	−0.010	0.004	0.009	0.002
	(0.004)	(0.004)	(0.008)	(0.004)	(0.006)	(0.005)
2019	−0.003	0.003	−0.012	0.011	0.017∗	0.006
	(0.006)	(0.007)	(0.010)	(0.007)	(0.009)	(0.007)
2020	0.001	0.012	−0.014	0.009	0.028∗∗	−0.000
	(0.008)	(0.013)	(0.011)	(0.007)	(0.012)	(0.007)
2021	−0.005	0.018	−0.016	0.011∗	0.022∗	−0.001
	(0.008)	(0.016)	(0.010)	(0.007)	(0.012)	(0.008)
2022	−0.014	0.003	−0.021∗∗	0.006	0.017	−0.003
	(0.009)	(0.009)	(0.011)	(0.008)	(0.012)	(0.009)
R^2	0.349	0.364	0.280	0.286	0.341	0.331

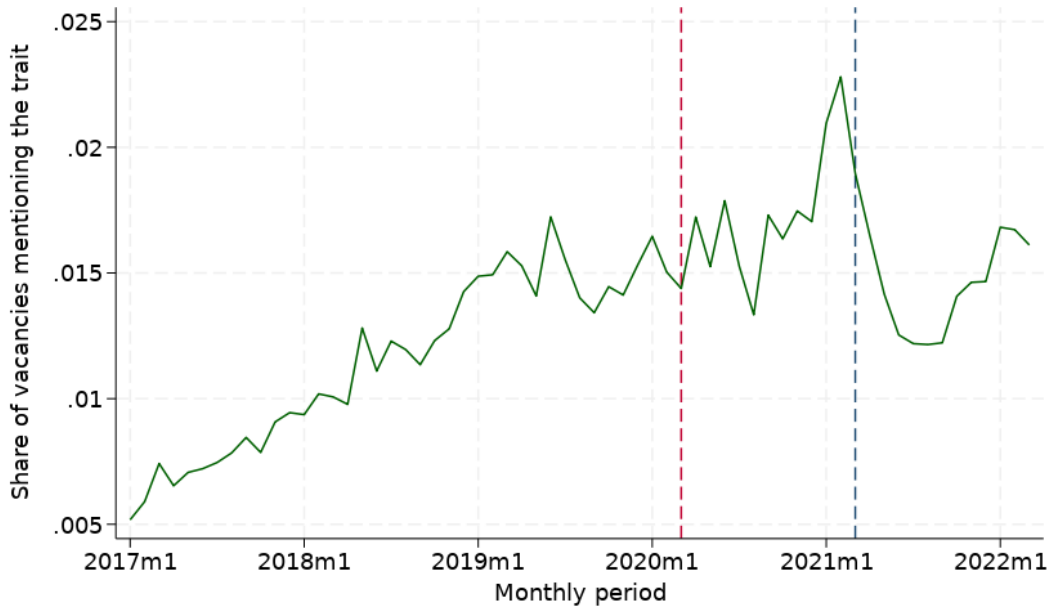
Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N = 9,888,315. "Any P.T." refers to Any Personality Trait, "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Regression controls for traits and skills demand, remote work requirements, length of advertisement, contract type, work schedule, bonus offerings, education, and negotiable salary option. Location, occupation and firm fixed effects. SE clustered at the occupation and firm level. Author's creation using Adzuna.

7.2 AI demand and Personality Traits

This section examines how artificial intelligence adoption relates to firms' demands for personality traits. While research has documented AI's impact on technical skill requirements (Acemoglu et al., 2022) and cognitive tasks (Babina et al., 2024), little is known about how AI adoption affects demands for specific personality characteristics. As Johnson and Acemoglu (2023) observe, firms adopting AI technologies might increasingly value workers with personality traits that complement AI capabilities, particularly those involving social interaction, adaptation and communication.

In the Adzuna dataset, 1.3 per cent of advertisements mention AI-related skills, with an increasing trend over time (Figure 4). This demand started at 0.75 per cent in 2017, peaked at 2.3 per cent in early 2021—coinciding with the end of Covid-19 restrictions—before stabilising around 1.5 per cent. This modest but growing trend matches findings from Schmidt et al. (2024), who report that AI-hiring intensity in the UK peaked at the end of 2020 before returning to pre-COVID levels.

Figure 4: Change over time of AI-skills demand



Notes: Dashed lines show the beginning (red) and end (blue) of Covid-19 restrictions.
Author's creation using Adzuna.

AI-demanding vacancies strongly require technical capabilities, with 91.3 per cent requiring hard skills compared to 54.2 per cent in non-AI positions. Software skills are the most frequently co-demanded proficiency (59.2 per cent), followed by engineering skills (45.3 per cent) and data analysis capabilities (35.1 per cent). While technical skills dominate, these positions also need substantial interpersonal capabilities, with

communication skills (29.3 per cent), teamwork (16.4 per cent) and problem-solving abilities (13.8 per cent) frequently appearing alongside AI requirements (Figures A.14 and A.15 in Appendix A). This aligns with research by Squicciarini and Nachtigall (2021), who found that skills related to communication, problem-solving, creativity, and teamwork increasingly appear alongside software-related and AI-specific competencies in job postings. However, AI requirements remain specialised - among positions requiring software skills, only 8.5 per cent demand AI capabilities, while for engineering and data analysis positions these figures are 7.3 per cent and 5.5 per cent respectively. This suggests that while AI skills complement other technical capabilities, they represent a distinct technical requirement.

AI-demanding and non-AI-demanding positions show strong differences in personality requirements. These are compared in Table 8. Openness demand is substantially higher in AI-demanding positions (52.1 per cent versus 29.6 per cent), indicating that employers value adaptability and creative thinking alongside AI skills. This aligns with Alekseeva et al. (2021), who identified increased demand for creative and adaptive skills in AI-intensive jobs. Conversely, AI-demanding positions show lower demands for Agreeableness (6.2 per cent versus 13.6 per cent), Conscientiousness (29.9 per cent versus 34.2 per cent), Emotional Stability (7.8 per cent versus 14.1 per cent), and Extraversion (35.2 per cent versus 44.4 per cent). Interestingly, the overall demand for soft skills and personality traits remains relatively stable between AI and non-AI positions (62.8 per cent versus 62.4 per cent for soft skills, and 73.1 per cent versus 72.1 per cent for personality traits), suggesting that AI adoption could reshape the specific personality traits demanded rather than reducing the overall importance of non-technical attributes.

Table 8: Demand for Personality Traits and Skills in vacancies by their AI demand, in percentages

Demands AI	Hard	Soft	All P.T.	Agree.	Consc.	Emot.	Extra.	Open.
Yes	91.3	62.8	73.1	6.2	29.9	7.8	35.2	52.1
No	54.2	62.4	72.1	13.6	34.2	14.1	44.4	29.6

Notes: "Hard" refers to Hard Skills, "Soft" to Soft Skills,, "Any P.T." to Any Personality Trait, "Consc." to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Author's creation using Adzuna.

While the overall demand for AI skills in job advertisements is modest, this figure masks substantial variation across occupations. AI requirements are most prevalent in Natural and Social Science Professionals jobs (10.1 per cent of vacancies), followed by IT and Telecommunications Professionals (9.2 per cent) and Engineering Professionals (5.9 per cent). These figures suggest that AI skills have become significant in hiring requirements for knowledge-intensive sectors, particularly those involving research, data analysis, and technical development, similar to patterns identified by Babina et al. (2024). The relatively high demand in Natural and Social Science Professionals (10.1 per cent) is

particularly noteworthy, indicating AI's growing importance beyond purely programming domains, possibly reflecting increased use of AI tools in research methodologies and data analysis across scientific disciplines.

AI-demanding vacancies are concentrated in specific occupations. While AI skills are demanded across the 90 3-digit SOC occupational categories, 78 per cent of all AI-demanding vacancies appear in just five occupations, with IT and Telecommunications Professionals alone accounting for 49.1 per cent. This is followed by Business, Research and Administrative Professionals (10.3 per cent), Engineering Professionals (7.2 per cent), Legal Associate Professionals (6.2 per cent) and Electrical and Electronic Trades (5.1 per cent). This concentration aligns with Schmidt et al. (2024), who identified Information and Communication, Finance and Insurance, and Professional, Scientific and Technical Activities as the most AI job-intensive sectors across all countries studied. This concentration suggests that skills required to implement artificial intelligence remain primarily within specialised roles. Without accounting for occupation fixed effects, the analysis of AI's relationship with personality traits would be heavily influenced by the characteristics and requirements of these dominant sectors.

To examine how AI adoption relates to personality trait demands while accounting for potential confounding factors, I estimate:

$$P(Trait_i = 1) = \beta_0 + \beta_1 AI_i + \beta_2 X_i + \delta_l + \theta_t + \epsilon_i \tag{7}$$

where AI_i is a dummy equal to 1 when the advertisement mentions any AI-related skill, X_i is a vector of job posting characteristics, δ_l represents location fixed effects, and θ_t captures time fixed effects. The vector X_i includes job posting characteristics: soft and hard skills requirements, education levels, seniority, experience, contract type, work schedule, bonus availability and advertisement length.

This comparison with detailed controls addresses potential confounding factors. Some firms might systematically demand more personality traits, locations might have different conditions affecting trait demands, and requirements might vary over time or across occupations. Also, the relationship between AI and personality traits might stem from correlations with other job requirements rather than a direct relationship. The extensive controls and fixed effects help isolate the specific relationship between AI adoption and personality trait demands[19].

[19]For example, in the most comprehensive model, consider two software developer positions advertised by Google in London during September 2018: one requiring Conscientiousness and another not mentioning it. Even in this narrow comparison, the positions might differ in other requirements—one might require more experience or offer different contract terms. By controlling for these job-specific characteristics while including fixed effects, I can isolate how AI requirements relate to the probability of demanding Conscientiousness, keeping constant occupation, location, time, and firm factors, as well as other job posting attributes.

Table 9 shows the relationship between AI adoption and personality trait demands across different specifications. The baseline specification (column 1) shows substantial heterogeneity in how AI relates to different traits. While the overall effect is modestly positive (1.0 percentage point), this masks strong opposing forces: a large positive association with Openness (22.5 percentage points) counterbalanced by negative relationships with other traits, particularly Extraversion (-9.2 percentage points) and Agreeableness (-7.4 percentage points).

Adding job requirement controls (column 2) substantially alters these relationships, while including time and location fixed effects (columns 3–4) produces modest changes. Column 5 adds industry fixed effects: coefficients remain broadly stable relative to column 4, with the exception of Conscientiousness, which becomes notably more negative (-6.7 percentage points), and Extraversion, whose negative association strengthens to -9.8 percentage points. The positive effect on Openness rises to 16.2 percentage points. This pattern suggests that industry sorting explains relatively little of the AI–personality relationship, and that the associations persist within narrowly defined sectors.

After controlling for hard and soft skills, experience, seniority, and contract types, the overall effect becomes negative (-1.2 percentage points), suggesting the initial positive relationship was driven by AI positions requiring more comprehensive skill sets. At the trait level, this shift reflects both a reduction in the positive effect on Openness (from 22.5 to 14.1 percentage points) and in the negative effect of Extraversion (from -9.2 to -8.2 percentage points). On the contrary, adding job controls generates a stronger negative association with Conscientiousness (from -4.3 to -6.5 percentage points). These changes likely reflect the high correlation between personality traits and other job requirements, in particular with hard skills, which are substantially more prevalent in AI-demanding positions.

After including firm and occupation fixed effects, the effect of AI adoption on personality trait demands remains consistently positive: whether controlling for firm (1.1 percentage points, column 6), occupation (2.1 percentage points, column 7), or both (1.6 percentage points, column 8). Trait-level analysis reveals substantial heterogeneity. Even in the most demanding specification with all fixed effects (column 8), AI adoption shows a strong positive relationship with Openness (6.7 percentage points), but negative associations with other traits, particularly Extraversion (-2.5 percentage points), Emotional Stability (-2.2 percentage points), and Conscientiousness (-1.7 percentage points). The positive effect on Openness is more than twice the size of any negative effect, suggesting that AI adoption leads to a reallocation of emphasis across personality traits rather than an overall reduction in personality requirements. These patterns align with theories about AI complementing creativity and adaptability while potentially substituting for interpersonal and routine-oriented traits. This view is supported by Acemoglu et al. (2022), who

found that AI exposure is associated with a shift toward non-routine analytical skills, and by Alekseeva et al. (2021), who documented that AI-intensive jobs increasingly require creative problem-solving abilities alongside technical competencies.

Including firm fixed effects (column 6) substantially reduces coefficient magnitudes compared to specifications without firm controls (columns 1–5), suggesting that firm-specific practices explain a large portion of the relationship between AI and personality demands. The effect is particularly strong for Openness, where the coefficient drops from 14.3 to 7.3 percentage points, and for Extraversion, decreasing from -8.3 to -3.4 percentage points. This reduction indicates that much of the raw correlation between AI adoption and personality demands reflects systematic differences across firms rather than AI adoption itself.

When controlling for occupation fixed effects instead (column 7), the coefficients show a different pattern. While still smaller than the raw correlations, the occupation-only specification maintains stronger effects than the firm fixed effects model, particularly for Openness (11.1 percentage points) and Conscientiousness (-3.8 percentage points). This suggests that occupational requirements explain less of the variation in AI-personality trait relationships than firm characteristics. The way firms implement AI appears to have a stronger influence on personality demands than occupation-specific patterns of AI adoption.

These relationships persist in the most comprehensive specification with both firm and occupation fixed effects (column 8). The coefficients generally decrease further, with Openness showing a 6.7 percentage point effect and other traits showing modest negative associations ranging from -1.2 to -2.5 percentage points. This leads to a positive overall effect (1.6 percentage points), as the strong complementarity with Openness outweighs the more modest negative associations with other traits. These results suggest two key insights: first, both firm practices and occupational requirements play important roles in mediating AI-personality relationships; second, there exists a complementarity between AI adoption and demand for Openness that persists even after controlling for both firm and occupation characteristics. The fact that firm fixed effects produce a larger reduction in coefficients than occupation fixed effects suggests that firm-level decisions about AI adoption and implementation may be more crucial in shaping personality requirements than occupation-specific technological constraints.

The analysis reveals complex relationships between AI adoption and personality trait demands across multiple empirical specifications. While AI-demanding positions initially appear to require more personality traits overall, once accounting for job characteristics and fixed effects, I find evidence of significant reallocation in personality demands. As Johnson and Acemoglu (2023) argue, AI tends to automate routine cognitive tasks while

increasing demand for human attributes like adaptability and creativity, which aligns with the finding of increased demand for Openness, alongside modest reductions in demands for traits related to interpersonal interaction and routine task execution. This pattern is consistent with Autor (2014)'s observations on how technological change tends to "amplify the comparative advantage of workers in supplying problem-solving, adaptability and creativity" (p. 2), while substituting for more routine activities. The magnitude of these effects varies substantially with the empirical specification, with firm-specific practices explaining a larger portion of the variation than occupational requirements, echoing Deming and Kahn (2018)'s work on the importance of firm heterogeneity in skill demands.

Table 9: Impact of AI on Personality Traits

	(1) Baseline	(2) +Controls	(3) +Time	(4) +Location	(5) +Industry	(6) +Firm	(7) +Occ	(8) All FE
Panel A: Overall Effect								
All Personality Traits	0.010***	-0.012***	-0.012***	-0.011***	-0.006***	0.011	0.021***	0.016**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.007)	(0.005)	(0.007)
Panel B: Effects by Trait								
Agreeableness	-0.074***	-0.028***	-0.031***	-0.033***	-0.026***	-0.015***	-0.011***	-0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.004)
Conscientiousness	-0.043***	-0.065***	-0.064***	-0.064***	-0.067***	-0.023***	-0.038***	-0.017**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.007)	(0.009)	(0.009)
Emotional Stability	-0.063***	-0.044***	-0.046***	-0.047***	-0.045***	-0.027***	-0.026***	-0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.010)	(0.007)
Extraversion	-0.092***	-0.082***	-0.082***	-0.083***	-0.098***	-0.034***	-0.038***	-0.025***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.008)	(0.009)	(0.007)
Openness	0.225***	0.141***	0.143***	0.143***	0.162***	0.073***	0.111***	0.067***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.008)	(0.014)	(0.013)
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	No	No	No	Yes	Yes	Yes	Yes	Yes
Industry FE	No	No	No	No	Yes	No	No	No
Firm FE	No	No	No	No	No	Yes	No	Yes
Occupation FE	No	No	No	No	No	No	Yes	Yes
R-squared range	[0.001, 0.012]	[0.080, 0.121]	[0.082, 0.122]	[0.084, 0.123]	[0.103, 0.203]	[0.273, 0.351]	[0.105, 0.169]	[0.281, 0.364]
N observations	11,662,124	11,662,124	11,662,124	9,923,701	9,923,697	9,888,315	9,923,701	9,888,315
N firms	128,559	128,559	128,559	117,229	117,229	81,843	117,229	81,843

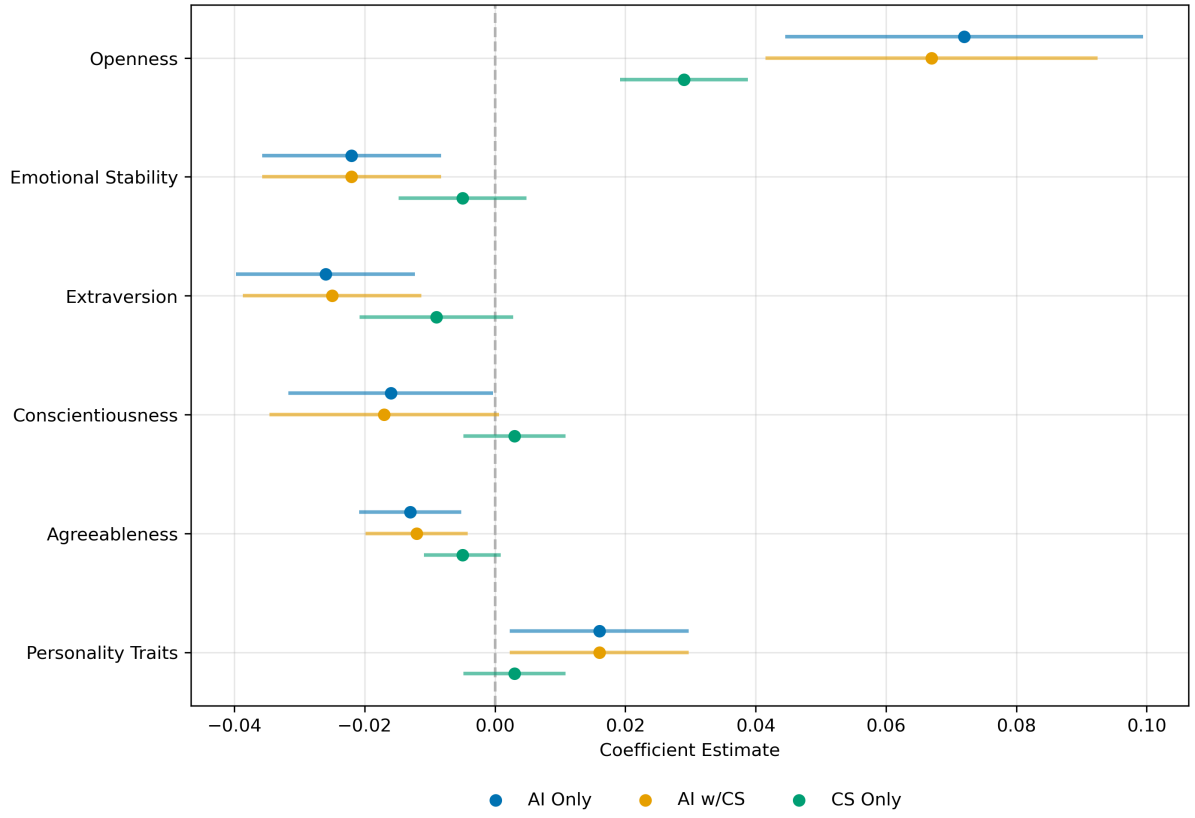
Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column and coefficient represents a different specification with progressively added controls and fixed effects. The R-squared range shows the minimum and maximum values across the six trait regressions for each specification. All regressions include computer skills controls and standard controls (traits and skills demand, remote work requirements, length of advertisement, contract type, work schedule, bonus offerings, education, and negotiable salary option). Standard errors are clustered at the occupation and firm level in specifications (6)–(8).

7.2.1 Robustness: Distinguishing AI from General Computer Skills

A potential concern is that the identified effects of AI skills might simply reflect broader technological requirements rather than AI-specific demands. To address this, I examine whether general computer skills requirements drive the results. I define computer skills as job requirements related to software development and data analysis capabilities within the broader hard skills category. This includes 1,911,086 vacancies, 12 times more than those demanding AI skills. Table A.5 presents estimates excluding computer skills controls, while Table A.6 shows the effects of computer skills—instead of AI—on personality traits.

The comparison shows significant differences between AI and computer skills effects. Without controlling for computer skills, the AI coefficient on overall personality traits is 1.6 percentage points ($p < 0.05$), similar to the main specification. However, examining trait-specific effects shows that computer skills alone generate substantially weaker relationships. As shown in Figure 5, while both AI and computer skills show positive associations with Openness, the AI effect (7.2 percentage points) is more than twice the magnitude of the computer skills effect (2.9 percentage points). This difference is statistically significant, with $p < 0.01$ in the most demanding specification with all fixed effects (Table A.7).

Figure 5: Effects on Personality Trait demand: AI vs Computer Skills



Notes: Author's creation using Adzuna.

The other traits show no significant effect in the computer skill specification. Formal

tests comparing AI and computer skills coefficients reject coefficient equality for all traits except Agreeableness in the full specification. This is because AI and computer skills have very small impacts on Agreeableness (-1.13 vs -0.5 percentage points). Additionally, when testing the difference in AI coefficients when removing computer skills controls (Table A.8 in Appendix A), I cannot reject equality of means in the occupation and firm fixed effects estimations. This placebo exercise strengthens confidence in the main results by showing that AI adoption influences personality requirements through channels distinct from general technological skill demands. The substantially larger effects of AI, particularly for Openness, suggest that the identified relationships reflect specific features of AI technology rather than simply correlating with broader computer skill requirements.

7.3 From Vacancy-Level to Firm-Level Effects

The vacancy-level analysis shows that firm-specific practices play a crucial role in mediating the relationship between AI adoption and personality trait demands, with firm fixed effects producing larger coefficient reductions than occupational controls. This suggests that firm-level decisions about AI adoption and implementation may be more important in shaping personality requirements than occupation-specific technological constraints. To further investigate these firm-level dynamics, I conduct two complementary analyses that differ conceptually from the previous fixed effects estimation.

While the vacancy-level analysis examines how personality requirements differ between AI and non-AI positions within firms, the firm-level approach focuses on how firms' overall hiring patterns change when they demand AI technologies. First, I examine changes in the average personality requirements across all vacancies posted by firms in months when they also recruit for AI positions. Second, I analyse changes in the volume of vacancies requiring different personality traits, providing insight into potential substitution patterns between AI and non-AI roles.

For the firm-level analysis, the unit of observation is the firm-month. Let $\bar{Y}_{j,t}$ denote the average share of vacancies posted by firm j in month t that mention a given personality trait. I estimate:

$$\bar{Y}_{j,t} = \alpha + \beta\, AI_{j,t} + \boldsymbol{\gamma}' X_{j,t} + \delta_l + \theta_t + \phi_j + \epsilon_{j,t} \tag{8}$$

where $AI_{j,t}$ is a dummy equal to one if firm j posts at least one AI-skill vacancy in month t, $X_{j,t}$ is a vector of average job characteristics for firm j in month t (the same controls as in Equation 4), δ_l are location fixed effects, θ_t are month-year fixed effects, and ϕ_j are firm fixed effects included in the most demanding specification. Standard errors are clustered at the firm level. The coefficient β captures the within-firm difference in personality trait demands between months when a firm recruits for AI positions and months when it does not.

Table 10 presents the firm-level analysis of personality requirements. The results align with the vacancy-level findings. After controlling for average job characteristics, location, industry, and occupational composition (Column 5), firms posting AI vacancies show a significant reduction in overall personality requirements (-3.0 percentage points, $p < 0.01$). There is a substantial increase in demands for Openness (3.4 percentage points, $p < 0.01$) alongside reductions in requirements for other traits, particularly Extraversion (-4.3 percentage points, $p < 0.01$) and Conscientiousness (-3.6 percentage points, $p < 0.01$).

When including firm fixed effects (Column 6), many of these effects become smaller but remain consistent in their sign. For instance, the effect on Openness remains positive and statistically significant (1.8 percentage points, $p < 0.01$), while the negative effects on Extraversion and Emotional Stability persist (-1.1 percentage points for both, $p < 0.01$). This within-firm analysis suggests that as firms adopt AI, they systematically adjust their personality requirements across all vacancies. The persistence of these patterns at the firm level suggests that AI adoption relates to broader changes in firms' desired worker characteristics rather than just requirements specific to AI positions.

Table 11 examines how AI posting relates to the volume of vacancies requiring different personality traits. The results confirm the changes in firms' hiring patterns. After controlling for average job characteristics, industry, location, and occupational composition (Column 5), firms posting AI vacancies show a significant reduction in positions requiring personality traits overall (-5.6 vacancies, $p < 0.01$), despite a large increase in total vacancies (80.3 vacancies, $p < 0.01$). This pattern is consistent across all traits except Openness: firms significantly reduce vacancies requiring Agreeableness (-13.5 vacancies), Emotional Stability (-5.6 vacancies), Extraversion (-3.6 vacancies), and Conscientiousness (-1.7 vacancies). In contrast, they increase positions requiring Openness (7.4 vacancies, $p < 0.01$).

The within-firm analysis (Column 6) shows smaller but still substantial effects, with firms significantly increasing total vacancies (30.5 vacancies, $p < 0.01$) when posting AI positions. The model also confirms that firms continue to increase positions requiring Openness (3.6 vacancies, $p < 0.01$) while reducing Agreeableness-demanding positions (-5.8 vacancies, $p < 0.01$).

Interestingly, firms posting AI vacancies also show a substantial reduction in non-AI demanding positions (-4.5 vacancies in Column 5 and -2.8 vacancies in Column 6, both $p < 0.01$), suggesting significant substitution between AI and non-AI roles in their hiring patterns. This result aligns with Acemoglu et al. (2022) and Choi and Leigh (2024).

Table 10: Impact of AI on Personality Traits: Firm-Month Level Analysis

	(1) Baseline	(2) +Controls	(3) +Location	(4) +Industry	(5) +Occupation	(6) +Firm
Panel A: Overall Effect						
All Personality Traits	0.006***	-0.049***	-0.051***	-0.044***	-0.030***	-0.001
	(0.002)	(0.001)	(0.001)	(0.005)	(0.004)	(0.002)
Panel B: Effects by Trait						
Agreeableness	-0.031***	-0.025***	-0.025***	-0.018***	-0.010***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)
Conscientiousness	-0.015***	-0.057***	-0.056***	-0.051***	-0.036***	-0.006**
	(0.002)	(0.002)	(0.002)	(0.006)	(0.005)	(0.003)
Emotional Stability	-0.029***	-0.041***	-0.040***	-0.035***	-0.021***	-0.011***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.002)
Extraversion	-0.053***	-0.068***	-0.071***	-0.068***	-0.043***	-0.011***
	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.003)
Openness	0.153***	0.066***	0.058***	0.053***	0.034***	0.018***
	(0.002)	(0.002)	(0.002)	(0.008)	(0.006)	(0.003)
Controls	No	Yes	Yes	Yes	Yes	Yes
Location FE	No	No	Yes	Yes	Yes	Yes
Industry FE	No	No	No	Yes	Yes	No
Occupation FE	No	No	No	No	Yes	Yes
Firm FE	No	No	No	No	No	Yes
R-squared range	[0.000, 0.005]	[0.067, 0.237]	[0.068, 0.238]	[0.082, 0.248]	[0.091, 0.264]	[0.476, 0.574]
N observations	728161	728161	701816	701814	701814	651808

Notes: Analysis conducted at firm-month level. Standard errors in parentheses, clustered at firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a different specification with progressively added controls and fixed effects. Column (6) includes firm fixed effects, capturing within-firm variation in personality trait demands associated with AI adoption. All regressions include computer skills controls. Regressions with controls include traits and skills demand, remote work requirements, length of advertisement, contract type, work schedule, bonus offerings, education, and negotiable salary option.

The comparison between proportions and counts of personality-demanding vacancies shows consistent patterns in firms' hiring strategies during months of AI adoption. Both specifications show negative effects across all traits except Openness. For instance, firms reduce both the share of vacancies requiring Extraversion (-4.3 percentage points) and their absolute numbers (-3.6 vacancies). The effect on Openness stands out for its consistency in the opposite direction: firms increase both the share of vacancies requiring this trait (3.4 percentage points) and the absolute number of such positions (7.4 vacancies). The firm fixed effects analysis confirms these patterns, showing that they persist even when examining within-firm changes over time. This robust pattern reinforces the complementarity between AI adoption and demand for adaptive and creative characteristics, while suggesting substitution away from other personality traits and non-AI vacancies.

Table 11: Impact of AI Adoption on Vacancy Counts: Firm-Month Level Analysis

	(1) Baseline	(2) +Controls	(3) +Location	(4) +Industry	(5) +Occupation	(6) +Firm
Panel A: Effect on Personality Trait Demanding Vacancies						
All Personality Traits	-5.999***	-6.254***	-6.353***	-6.073***	-5.646***	-0.836
	(0.534)	(0.558)	(0.572)	(1.941)	(1.935)	(0.830)
Panel B: Effects by Trait						
Agreeableness	-13.316***	-13.486***	-13.776***	-13.579***	-13.521***	-5.811***
	(1.239)	(1.301)	(1.335)	(4.897)	(4.925)	(2.059)
Conscientiousness	-2.179***	-2.159***	-2.145***	-2.101	-1.742	0.495
	(0.608)	(0.632)	(0.649)	(1.602)	(1.653)	(0.916)
Emotional Stability	-5.914***	-5.894***	-5.999***	-5.822*	-5.616*	-2.415
	(1.046)	(1.098)	(1.127)	(3.009)	(3.071)	(1.778)
Extraversion	-4.549***	-4.309***	-4.325***	-4.269**	-3.624*	-0.681
	(0.836)	(0.871)	(0.894)	(2.104)	(2.182)	(1.165)
Openness	7.504***	7.348***	7.499***	7.593***	7.421***	3.563***
	(0.655)	(0.677)	(0.694)	(2.163)	(2.168)	(0.786)
Panel C: Effects on Non-AI Vacancies						
Non-AI Vacancies	-4.574***	-4.525***	-4.583***	-4.602***	-4.463***	-2.768***
	(0.058)	(0.056)	(0.057)	(0.317)	(0.279)	(0.092)
Controls	No	Yes	Yes	Yes	Yes	Yes
Location FE	No	No	Yes	Yes	Yes	Yes
Industry FE	No	No	No	Yes	Yes	No
Occupation FE	No	No	No	No	Yes	Yes
Firm FE	No	No	No	No	No	Yes
R-squared range	[0.348, 0.906]	[0.350, 0.907]	[0.350, 0.907]	[0.352, 0.907]	[0.352, 0.907]	[0.573, 0.950]
N observations	728161	728161	701816	701814	701814	651808

Notes: Analysis conducted at firm-month level. The dependent variables are counts of vacancies. Panels A and B show effects on the number of vacancies demanding personality traits. Panel C shows effects on total vacancies and non-AI vacancies. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a different specification with progressively added controls and fixed effects. Column (6) includes firm fixed effects, capturing within-firm variation in vacancy patterns associated with AI adoption. All regressions control for total number of vacancies.

7.4 Staggered Difference-in-Differences Analysis

These firm-level findings motivate a deeper investigation of the causal effects of AI adoption on personality demands. While the fixed effects analyses show strong correlations between AI adoption and personality requirements, they cannot fully address potential endogeneity in firms' AI adoption decisions. Firms might adopt AI precisely when planning to reorganise their workforce or shift skill requirements. To provide causal evidence on how AI adoption affects firms' personality trait demands, I use a staggered difference-in-differences strategy that exploits variation in the timing of firms' first AI adoption. This approach captures organisational changes that might occur after firms begin integrating AI skills, even in positions not directly requiring them.

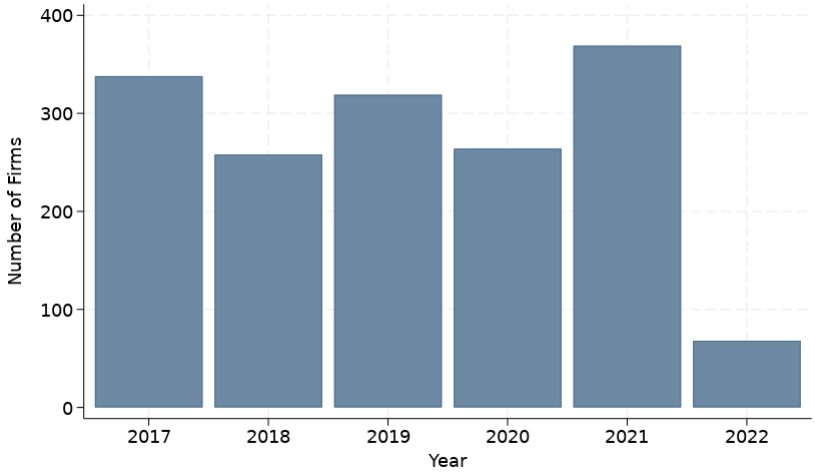
Following Callaway and Sant'Anna (2021)'s methodology, I identify for each firm the first time they post a job advertisement requiring AI skills. This creates a staggered treatment design where firms adopt AI at different times. To ensure comparability between treated and control firms, I implement three specifications: a base specification without pre-treatment controls, a specification with pre-treatment covariates, and entropy balancing. The covariates specification includes pre-treatment characteristics as control variables in the estimation, functioning similarly to a regression adjustment approach where observable differences between treated and control firms are accounted for through linear controls. The entropy balancing specification instead reweights the control group to match the treated group on both the mean and variance of pre-treatment characteristics, creating a balanced pseudo-population before estimation. Both approaches match on pre-treatment personality trait demands (specifically, the average demand for the trait being analysed in the pre-treatment period), pre-treatment firm characteristics (firm size and profit), and pre-treatment occupational composition at the 1-digit SOC level. Entropy balancing is the preferred specification as it ensures better balance across the full distribution of covariates rather than just controlling for their linear effects.

This approach offers advantages over the standard two-way fixed effects (TWFE) difference-in-differences model. First, preliminary analysis showed substantial treatment effect heterogeneity across adoption timing, with later adopters (2021-2022) having different effects compared to early adopters (2018-2019), which TWFE cannot properly account for. Second, while TWFE relies primarily on never-treated firms as controls, the staggered design uses the more comparable group of not-yet-treated firms, better accounting for selection into treatment timing. The staggered design captures the within-firm evolution of personality trait demands in the months and years around the adoption event while avoiding the contamination that arises in standard two-way fixed effects models when treatment effects differ across adoption cohorts.

Of the 128,559 firms in the database, only 1,610 posted vacancies demanding AI skills.

These AI-adopting firms collectively posted 61,431 AI-related job advertisements, indicating that firms demanding AI skills tend to post multiple relevant positions rather than single vacancies. This staggered adoption pattern creates variation in treatment timing essential for the difference-in-differences approach. Firms adopting AI in 2017 (21 per cent of adopters) are considered "always treated" in the yearly analysis, while those adopting in subsequent years (2018-2022) form distinct treatment cohorts (Figure 6). Since the data extends only to March 2022, adoption figures for this year are censored, representing just 4.2 per cent of total adopters. This temporal distribution allows identification of treatment effects while accounting for potentially heterogeneous impacts across different adoption periods, which would not be possible with a standard two-way fixed effects estimation.

Figure 6: Number of firms adopting AI by year



Notes: Author's creation using Adzuna.

The entropy balancing results at the yearly aggregation show modest but significant effects. As shown in Table 12, firms increase their overall personality trait demands by 2.7 percentage points ($p < 0.05$) following AI adoption. Looking at individual traits, the effect is strongest for Openness, with a significant increase of 3.6 percentage points ($p < 0.10$). Emotional Stability shows a significant decline of 2.6 percentage points ($p < 0.05$), while effects on other traits remain small and statistically insignificant. The covariates specification shows similar patterns, with a 2.0 percentage point increase in overall personality traits ($p < 0.05$) and a larger effect on Openness (4.9 percentage points, $p < 0.01$). The base specification without controls produces smaller and generally insignificant effects, highlighting the importance of accounting for pre-treatment differences between treated and control firms.

Table 12: ATT Estimates: Year Aggregation

	Entropy Balance	Base	Controls
Panel A: Overall Effect			
All Personality Traits	0.027**	0.004	0.020**
	(0.014)	(0.006)	(0.009)
Panel B: Effects by Trait			
Openness	0.036*	0.021***	0.049***
	(0.021)	(0.007)	(0.014)
Agreeableness	-0.015	-0.002	-0.005
	(0.016)	(0.005)	(0.009)
Conscientiousness	-0.001	0.007	0.009
	(0.019)	(0.007)	(0.012)
Emotional Stability	-0.026**	-0.005	-0.011
	(0.012)	(0.005)	(0.009)
Extraversion	0.001	0.003	-0.007
	(0.020)	(0.007)	(0.013)
Pre-trend Test P-values			
All PT	0.930	0.563	0.097
Openness	0.196	0.081	0.161
Agreeableness	0.086	0.325	0.131
Conscientiousness	0.022	0.826	0.446
Emotional Stability	0.331	0.446	0.358
Extraversion	0.716	0.584	0.341
Number of Observations			
Observations	22,009	174,250	22,009

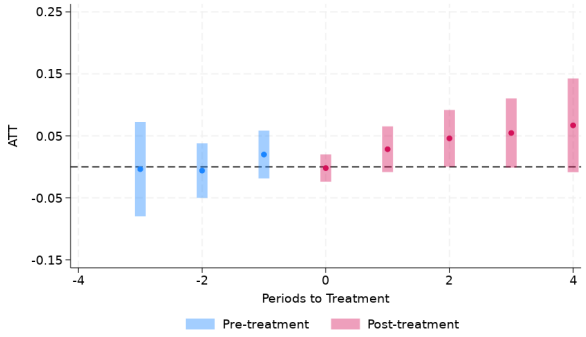
Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table compares ATT estimates using three different estimators at year level. Standard errors in parentheses, clustered at the firm level. Pre-trend test p-values from chi-square test of joint significance of pre-treatment coefficients.

The event study plots in Figure 7 show the dynamic effects of AI adoption on trait demands using the entropy balancing specification. For overall personality traits (panel a), pre-treatment trends remain flat and close to zero, followed by a gradual increase in the treatment effect over time, becoming more pronounced three to four years after adoption. The pattern for Openness (panel d) shows the most consistent positive trajectory, with the effect building steadily in the post-treatment period and reaching approximately 7 percentage points by the fourth year after adoption. For Emotional Stability (panel f), the pattern shows a decline in the post-treatment period. The patterns for Agreeableness (panel b), Conscientiousness (panel c), and Extraversion (panel e) show relatively stable trends with confidence intervals including zero throughout most of the post-treatment period, consistent with the insignificant average treatment effects for these traits.

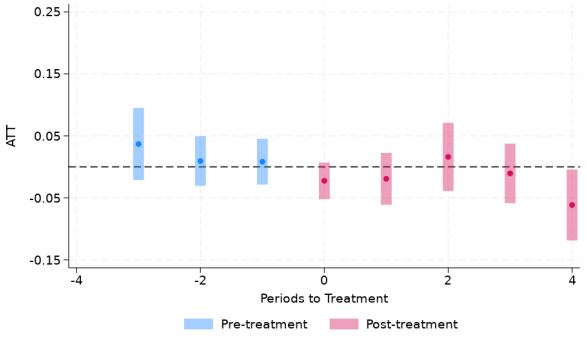
The entropy balancing specification lends support to the parallel trends assumption for most traits, as shown by the pre-trend test p-values in Table 12. I cannot reject the

Figure 7: Staggered DID ATT. Year-level aggregation, Entropy Balance

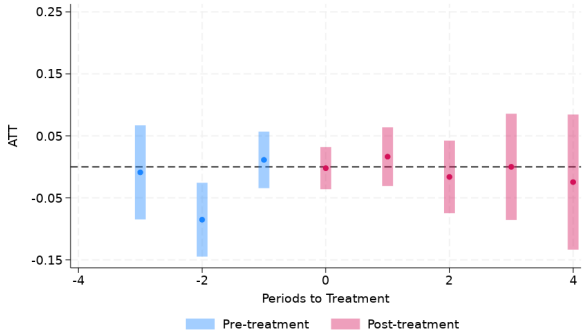
(a) All Personality Traits



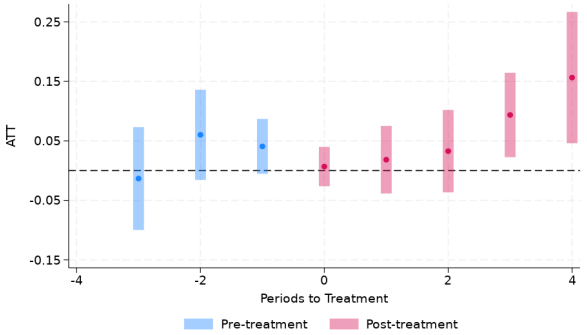
(b) Agreeableness



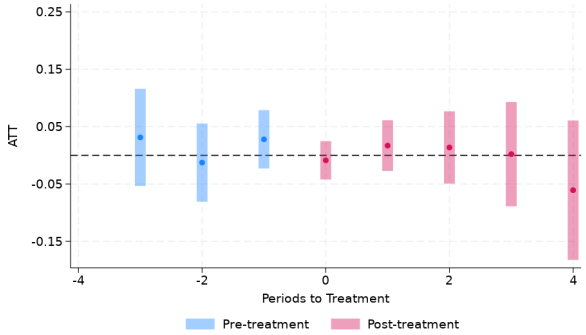
(c) Conscientiousness



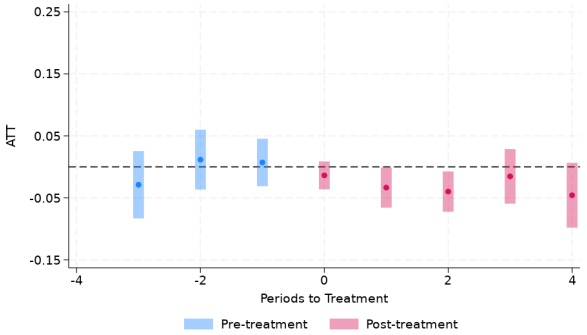
(d) Openness



(e) Extraversion



(f) Emotional Stability



Notes: Author's creation using Adzuna.

null hypothesis of parallel pre-trends for overall personality traits ($p = 0.930$), Openness ($p = 0.196$), Emotional Stability ($p = 0.331$), and Extraversion ($p = 0.716$). However, the test suggests potential pre-trend differences for Conscientiousness ($p = 0.022$) and Agreeableness ($p = 0.086$).

Robustness checks using different temporal aggregations (half-year and quarter) are presented in Tables A.9 and A.10 and Figures A.16 and A.17 in Appendix A. The half-year entropy balancing analysis confirms the main findings, showing a significant positive effect on overall personality traits (4.1 percentage points, $p < 0.05$) with particularly strong effects for Openness (6.4 percentage points, $p < 0.01$). The quarterly analysis also shows positive effects using entropy balancing, with overall personality traits increasing by 3.1 percentage points ($p < 0.10$) and Openness by 4.7 percentage points ($p < 0.05$). However, the precision of estimates decreases at finer time aggregations, likely reflecting both the challenges of identifying effects when AI skills appear in less than 2 per cent of vacancies and the increased noise in measuring firm-level changes over shorter periods, when firms may anticipate changes in their worforce structure.

An important consideration is whether the observed effects reflect changes in AI-demanding positions specifically or broader changes in firms' overall hiring patterns. To address this, I conduct a robustness check where I exclude all AI-demanding vacancies from the analysis while maintaining the treatment timing based on when firms first post AI positions. This approach measures personality trait demands only in non-AI positions before and after firms adopt AI technologies. If the effects were driven solely by the composition of AI versus non-AI positions, it would be expected that the effects would disappear when excluding AI positions. Instead, Table A.11 in Appendix A shows that the results remain largely unchanged. The entropy balancing specification shows a positive effect on overall personality traits (2.5 percentage points, $p < 0.10$), marginally smaller than the 2.7 percentage points found in the full sample. For Openness, the effect is 3.0 percentage points ($p > 0.15$), compared to 3.6 percentage points in the full sample. The decline in Emotional Stability persists at 2.5 percentage points ($p < 0.05$), nearly identical to the full sample result. The similarity of results when excluding AI positions implies that AI adoption leads to changes in personality requirements across firms' entire workforce, not just in AI-intensive roles. This suggests broader organisational adaptation to AI technologies rather than simple compositional shifts.

The staggered DID analysis provides complementary evidence to the fixed effects results, with both approaches capturing different aspects of the AI-personality relationship. The consistency between the results, especially regarding Openness, improves our understanding of how AI affects labour markets through both immediate and dynamic channels. The positive effects on personality trait demands appear both in specific AI-intensive positions and more broadly across firms' hiring patterns following AI adoption. This

suggests that the relationship between AI and personality demands reflects both targeted job-specific requirements and broader organisational adaptation to AI technologies. The stronger effects at more aggregated time periods likely reflect the gradual nature of these organisational changes, as firms adjust their hiring practices to complement their AI adoption. The significant negative effect on Emotional Stability suggests that AI adoption may reduce demand for traits associated with managing workplace stress and pressure.

8 Conclusions

This study provides comprehensive evidence on the demand for personality traits in the UK labour market and how this demand interacts with the increasing adoption of AI technologies. By analysing over 11.7 million job advertisements between 2017-2022, I identify several key patterns that improve our understanding of how personality traits are valued in online labour markets.

Main Findings

Addressing the first research question on the nature and extent of personality trait demands, I find that personality traits appear in 72 percent of job advertisements, more than traditional human capital measures like education (11 percent) and soft skills (62 percent). Extraversion is the most frequently demanded trait (44 percent), followed by Conscientiousness (34 percent), Openness (30 percent), Emotional Stability (14 percent), and Agreeableness (13 percent). The demand varies systematically across occupations, with customer-facing and administrative roles showing the highest requirements: Elementary Administration (88.6 per cent), Sales Assistants and Retail Cashiers (87.3 per cent), and Customer Service (85.3 per cent). In contrast, manual and construction-related occupations show much lower demands: Elementary Construction (20.2 per cent), Metal Forming and Welding Trades (33.1 per cent), and Building Finishing Trades (33.9 per cent). This pattern supports theoretical predictions about the importance of non-cognitive attributes in jobs requiring frequent interpersonal interaction (Deming, 2017).

There is strong evidence of assortative matching between personality trait demands and worker characteristics across occupations, particularly for Agreeableness ($\rho = 0.814$). This alignment between job requirements and worker self-reported traits supports theories of comparative advantage in job matching and validates our text-based approach to measuring personality demands. However, the exception of Emotional Stability highlights measurement challenges when comparing employer-stated requirements with worker self-assessments, particularly for traits that may be expressed differently in job advertisements versus personality surveys.

Although raw trait requirements have increased, most changes reflect shifts in occupa-

tional composition rather than within-occupation changes. After controlling for occupation and firm fixed effects, only two significant trends emerge during 2017-2022: a steady decline in Conscientiousness (-2.1 percentage points) and a modest increase in Extraversion (1.7 percentage points). This suggests that changes in personality requirements primarily reflect structural changes in the labour market rather than broad shifts in how firms value personality traits within occupations.

For the third research question, I use two complementary empirical strategies to examine how AI adoption affects personality trait demands. Jobs requiring AI skills are significantly more likely to demand Openness (7.2 percentage points) but show reduced demands for Extraversion (-2.7 percentage points), Emotional Stability (-2.2 percentage points), and Conscientiousness (-1.7 percentage points). This aligns with theoretical predictions that AI may complement creative and adaptive capabilities but potentially replace routine-oriented and interpersonal tasks.

The staggered difference-in-differences analysis provides causal evidence of how firms' personality demands change after their first AI adoption. Using entropy balancing to ensure comparable treated and control firms, I find significant effects on overall personality trait demands (2.7 percentage points, $p < 0.05$), with particularly strong effects for Openness (3.6 percentage points, $p < 0.10$) and a decline in Emotional Stability (2.6 percentage points, $p < 0.05$). The consistency of the Openness effect across both empirical approaches strengthens confidence in AI's role in increasing demand for creative and adaptive traits. The robustness of these findings to alternative time aggregations and to the exclusion of AI positions from the analysis further supports the causal interpretation of these effects, entailing that AI adoption influences personality requirements across firms' entire workforce rather than only in AI-intensive positions.

Mechanisms and Interpretation

This alignment between job-level effects and firm-wide adoption impacts suggests that AI adoption influences personality requirements through multiple channels. Although the strongest effects appear in AI-intensive positions, the positive impact on Openness extends more broadly across firms' hiring patterns following AI adoption. This indicates that as organisations integrate AI technologies, they increasingly value workers' capacity for innovation and flexible thinking, even in positions not directly involving AI skills.

These findings suggest that models of human capital and skill-biased technological change need to better incorporate personality traits as distinct factors in labour market outcomes. The consistent evidence that AI adoption increases demand for Openness, using both contemporaneous and dynamic analyses, indicates that technological change may affect personality requirements as systematically as it affects skill demands. This

highlights the need for models that explicitly consider how automation complements or substitutes for different personality traits.

Policy Implications

For policy makers and educators, these results emphasise the importance of including personality development, particularly traits related to creativity and adaptability, in workforce preparation strategies. The finding that personality traits appear in 72 per cent of job advertisements, far exceeding formal education requirements, suggests that the traditional focus on academic qualifications may be insufficient. Furthermore, the strong complementarity between AI and Openness indicates that educational institutions should particularly emphasise the development of creative thinking and adaptability as AI adoption continues to expand across sectors.

For firms, the findings suggest the value of explicitly considering personality traits in hiring strategies, especially when adopting AI technologies. The systematic evidence of assortative matching between job requirements and worker characteristics indicates that clearly stating personality requirements can help achieve better worker-job fit. Moreover, the causal evidence that AI adoption leads to increased demand for Openness suggests that firms should proactively develop skills related to these traits in their workforce as they plan AI implementation.

Limitations and Future Research

Several important questions emerge for future research. First, although this study establishes causal evidence of AI adoption on personality demands, understanding the mechanisms behind these effects requires further investigation. The strong complementarity between AI and Openness suggests that AI may be changing how tasks are performed, but directly measuring these changes in work organisation would provide valuable insights. Second, the robust evidence of assortative matching between personality requirements and worker characteristics raises questions about efficiency and equity. Does better personality-job matching improve productivity and worker satisfaction, and how does this affect wage inequality? Finally, as AI technology continues to evolve, longer-term analysis will be crucial for understanding whether the identified patterns persist or evolve as firms gain more experience with AI implementation.

Contribution

In conclusion, this study shows that personality traits play a crucial role in the modern labour market, appearing in 72 per cent of job advertisements and showing systematic patterns across occupations. The evidence that AI adoption causally increases demand

for Openness but potentially reduces requirements for other traits provides important insights into how technological change reshapes personality demands. Through both fixed effects and staggered difference-in-differences analyses, I find consistent evidence that AI complements creative and adaptive capabilities. These findings suggest that as AI adoption continues to expand, the ability to think creatively and adapt to change may become increasingly valuable in the labour market.

References

Abraham, L., Hallermeier, J., & Stein, A. (2024). Words matter: Experimental evidence from job applications. *Journal of Economic Behavior & Organization*, *225*, 348–391.

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022, April). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, *40*(S1), 293–340.

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, *33*(2), 3–30.

Alekseeva, L., Azar, J., Gine, M., Samila, S., & Taska, B. (2021). The demand for ai skills in the labor market. *Labour Economics*, *71*, 102002.

Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality psychology and economics. In *Handbook of the Economics of Education* (Vol. 4, pp. 1–181). Elsevier.

Autor, D. (2014). *Polanyi's paradox and the shape of employment growth* (Working paper No. 20485). National Bureau of Economic Research.

Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, *151*, 103745.

Banfi, S., & Villena-Roldan, B. (2019). Do high-wage jobs attract more applicants? directed search evidence from the online labor market. *Journal of Labor Economics*, *37*(3), 715–746.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, *44*(1), 1–26.

Batra, H., Michaud, A. M., & Mongey, S. (2023). *Online job posts contain very little wage information* (Working paper No. 31984). National Bureau of Economic Research.

Bellatin, A., & Galassi, G. (2022). *What covid-19 may leave behind: Technology-related job postings in canada* (Discussion Paper No. 15209). Institute of Labor Economics (IZA).

Bone, M., Ehlinger, E., & Stephany, F. (2024). *Skills or degree? the rise of skill-based hiring for ai and green jobs* (Tech. Rep. No. 20/2023). Bruegel.

Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, *43*(4), 972–1059.

Borghans, L., Ter Weel, B., & Weinberg, B. A. (2014). People skills and the labor-market outcomes of underrepresented groups. *ILR Review*, *67*(2), 287–334.

Brenčič, V. (2012). Wage posting: evidence from job ads. *Canadian Journal of Economics/Revue canadienne d'économique*, *45*(4), 1529–1559.

Brenčič, V., & McGee, A. (2023a). *Demand for personality traits, tasks, and sorting* (Discussion Paper No. 16576). Institute of Labor Economics (IZA).

Brenčič, V., & McGee, A. (2023b). *Employers' demand for personality traits* (Discussion

Paper No. 16083). Institute of Labor Economics (IZA).

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology*, *38*(4), 476–506.

Chaturvedi, S., Mahajan, K., & Siddique, Z. (2021). *Words matter: Gender, jobs and applicant behavior* (Discussion Paper No. 14497). Institute of Labor Economics (IZA).

Choi, T., & Leigh, N. G. (2024). Artificial intelligence's creation and displacement of labor demand. *Technological Forecasting and Social Change*, *209*, 123824.

Cobb-Clark, D. A., & Tan, M. (2011). Noncognitive skills, occupational attainment, and relative wages. *Labour Economics*, *18*(1), 1–13.

Collischon, M. (2020). The returns to personality traits across the wage distribution. *Labour*, *34*(1), 48–79.

Correia, S. (2016). *A feasible estimator for linear models with multi-way fixed effects.* (Preprint at http://scorreia. com/research/hdfe.pdf)

Costa Jr, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, *6*(4), 343–359.

Deming. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593–1640.

Deming, & Kahn. (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, *36*(S1), S337–S369.

Dickerson, A., & Morris, D. (2019). *The Changing Demand for Skills in the UK* (Discussion Paper No. 20). Centre for Vocational Education Research.

Draca, M., Duchini, E., Rathelot, R., Turrell, A., & Vattuone, G. (2022). *Revolution in progress? the rise of remote work in the uk* (Working Paper No. 616). Centre for Competitive Advantage in the Global Economy (CAGE).

Edin, P.-A., Fredriksson, P., Nybom, M., & Öckert, B. (2022). The rising return to noncognitive skill. *American Economic Journal: Applied Economics*, *14*(2), 78–100.

Fletcher, J. M. (2013). The effects of personality traits on adult labor market outcomes: Evidence from siblings. *Journal of Economic Behavior & Organization*, *89*, 122–135.

Flinn, C. J., Todd, P. E., & Zhang, W. (2025). Labor market returns to personality: A job search approach to understanding gender gaps. *Journal of Political Economy*, *133*(4), 1169–1234.

Gensowski, M. (2018). Personality, IQ, and lifetime earnings. *Labour Economics*, *51*, 170–183.

Goldberg, L. R. (1990). An alternative "description of personality": the big-five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216-–1229.

Goldberg, L. R. (1992). The development of markers for the big-five factor structure.

Psychological Assessment, *4*(1), 26—42.

Greene, W. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *The Econometrics Journal*, *7*(1), 98–119.

Grugulis, I., & Vincent, S. (2009, December). Whose skill is it anyway?: "soft" skills and polarization. *Work, Employment and Society*, *23*(4), 597–615.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, *20*(1), 25–46.

Heckman, J. J. (2011). *Integrating personality psychology into economics* (Working paper No. 17378). National Bureau of Economic Research.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, *19*(4), 451–464.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the ged testing program. *American Economic Review*, *91*(2), 145–149.

Heineck, G., & Anger, S. (2010). The returns to cognitive abilities and personality traits in germany. *Labour Economics*, *17*(3), 535–546.

Hershbein, B., & Kahn, L. B. (2018). Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review*, *108*(7), 1737–1772.

John. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In *Handbook of personality: Theory and research* (p. 66—100). Guilford Press.

John, O. P., Srivastava, S., et al. (1999). The big-five trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research* (pp. 102–138). Guilford Press.

Johnson, S., & Acemoglu, D. (2023). *Power and progress: Our thousand-year struggle over technology and prosperity*. Hachette UK.

Judge, T. A., & Cable, D. M. (1997). Applicant personality, organizational culture, and organization attraction. *Personnel Psychology*, *50*(2), 359–394.

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, *52*(3), 621–652.

Khaouja, I., Kassou, I., & Ghogho, M. (2021). A Survey on Skill Identification From Online Job Ads. *IEEE Access*, *9*, 118134–118153.

King, E. B., George, J. M., & Hebl, M. R. (2005). Linking personality to helping behaviors at work: An interactional perspective. *Journal of Personality*, *73*(3), 585–608.

Kruyen, P. M., Keulemans, S., Borst, R. T., & Helderman, J.-K. (2020). Searching for the renaissance bureaucrat: A longitudinal computer-assisted study of personality

traits in government vacancies. *International Journal of Public Sector Management*, *33*(1), 22–44.

Kuhn, P., & Shen, K. (2023). What happens when employers can no longer discriminate in job ads? *American Economic Review*, *113*(4), 1013–1048.

Kuhn, P., Shen, K., & Zhang, S. (2020). Gender-targeted job ads in the recruitment process: Facts from a chinese job board. *Journal of Development Economics*, *147*, 102531.

Lindqvist, E., & Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics*, *3*(1), 101–128.

Lovaglio, P. G., Cesarini, M., Mercorio, F., & Mezzanzanica, M. (2018). Skills in demand for ict and statistical occupations: Evidence from web-based job vacancies. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *11*(2), 78–91.

Marinescu, I., & Wolthoff, R. (2020). Opening the black box of the matching function: The power of words. *Journal of Labor Economics*, *38*(2), 535–568.

Matteson, M. L., Anderson, L., & Boyden, C. (2016). "soft skills": A phrase in search of meaning. *Libraries and the Academy*, *16*(1), 71–88.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175–215.

McGee, A., & McGee, P. (2020). *Whoever you want me to be: personality and incentives* (Discussion Paper No. 13809). Institute of Labor Economics (IZA).

Mueller, G., & Plug, E. (2006). Estimating the effect of personality on male and female earnings. *ILR Review*, *60*(1), 3–22.

Nandi, A., & Nicoletti, C. (2014). Explaining personality pay gaps in the uk. *Applied Economics*, *46*(26), 3131–3150.

Nyhus, E. K., & Pons, E. (2005). The effects of personality on earnings. *Journal of Economic Psychology*, *26*(3), 363–384.

REC. (2022). *REC Annual Report and Accounts 2021* (Tech. Rep.). UK Recruitment and Employment Confederation.

Rohrbach-Schmidt, D., Wehner, C., Krueger, S., & Ebner, C. (2023). Wage returns to job tasks and personality traits in germany. *International Journal of Manpower*, *44*(9), 55–71.

Saucier, G. (1994). Mini-markers: A brief version of goldberg's unipolar big-five markers. *Journal of Personality Assessment*, *63*(3), 506–516.

Schmidt, J., Pilgrim, G., & Mourougane, A. (2024). *Measuring the demand for ai skills in the United Kingdom* (OECD Publishing No. 25). Organisation for Economic Co-operation and Development (OECD).

Seibert, S. E., & Kraimer, M. L. (2001). The five-factor model of personality and career success. *Journal of Vocational Behaviour*, *58*(1), 1–21.

Squicciarini, M., & Nachtigall, H. (2021). *Demand for ai skills in jobs: Evidence from online job postings* (Working Paper No. 2021/03). Organisation for Economic Co-operation and Development (OECD).

Stahle, C. (2024). *Educational requirements are gradually disappearing from job postings.*

Stanek, K. C., & Ones, D. S. (2023). Meta-analytic relations between personality and cognitive ability. *Proceedings of the National Academy of Sciences*, *120*(23).

Todd, P. E., & Zhang, W. (2020). A dynamic model of personality, schooling, and occupational choice. *Quantitative Economics*, *11*(1), 231–275.

Turrell, A., Speigner, B. J., Djumalieva, J., Copple, D., & Thurgood, J. (2019). *Transforming naturally occurring text data into economic statistics: The case of online job vacancy postings* (Working paper No. 25837). National Bureau of Economic Research.

Viinikainen, J., & Kokko, K. (2012). Personality traits and unemployment: evidence from longitudinal data. *Journal of Economic Psychology*, *33*(6), 1204–1222.

Webb, M. (2019). *The impact of artificial intelligence on the labor market.* (Webb, Michael, The Impact of Artificial Intelligence on the Labor Market (November 6, 2019). Available at SSRN: https://ssrn.com/abstract=3482150)

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, *96*(5), 849–861.

Woessmann, L. (2024). Skills and earnings: A multidimensional perspective on human capital. *Annual Review of Economics*, *17*.

9 Appendices

Appendix A - Figures and Tables

Table A.1: Occupation Characteristics in 2021 vs 2021 Census data

SOC	Name	Salary	Vacancies	Share	vs Census
111	Chief Executives and Senior Officials	32710	2172	0.1%	-67.0%
112	Production Managers and Directors	39471	30125	1.0%	-29.7%
113	Functional Managers and Directors	53967	31597	1.1%	-43.7%
115	Financial Institution Managers and Directors	42494	1572	0.1%	-86.9%
116	Managers and Directors in Transport and Logistics	32847	28909	1.0%	94.5%
117	Senior Officers in Protective Services	33109	429	0.0%	-90.8%
118	Health and Social Services Managers and Directors	36402	7820	0.3%	-10.8%
119	Managers and Directors in Retail and Wholesale	25263	5599	0.2%	-87.2%
121	Managers and Proprietors in Agriculture Related S.	33062	290	0.0%	-91.7%

Continued on next page

Table A.1 – continued from previous page

SOC	Name	Salary	Vacancies	Share	vs Census
122	Managers and Proprietors in Hospitality and Leisure S.	26514	25038	0.9%	-23.9%
124	Managers and Proprietors in Health and Care S.	36509	16373	0.6%	108.6%
125	Managers and Proprietors in Other Services	37251	20207	0.7%	-68.5%
211	Natural and Social Science Professionals	40216	6002	0.2%	-57.6%
212	Engineering Professionals	46820	38630	1.3%	7.8%
213	I.T. and Telecommunications Professionals	53345	174096	6.0%	187.2%
214	Conservation and Environment Professionals	37023	3692	0.1%	15.3%
215	Research and Development Managers	47057	3516	0.1%	9.0%
221	Health Professionals	49102	13124	0.5%	-64.9%
222	Therapy Professionals	36280	9741	0.3%	-2.7%
223	Nursing and Midwifery Professionals	37639	101933	3.5%	78.3%
231	Teaching and Educational Professionals	34845	38653	1.3%	-68.9%
241	Legal Professionals	55087	12525	0.4%	-29.5%
242	Business, Research and Administrative Professionals	50150	118196	4.1%	93.1%
243	Architects, Town Planners and Surveyors	48297	22169	0.8%	-3.2%
244	Welfare Professionals	34791	10862	0.4%	-22.5%
245	Librarians and Related Professionals	28842	254	0.0%	-91.4%
246	Quality and Regulatory Professionals	42531	22690	0.8%	206.7%
247	Media Professionals	36107	8310	0.3%	-41.1%
311	Science, Engineering and Production Technicians	30750	28458	1.0%	33.0%
312	Draughtspersons and Related Architectural Technicians	38480	3686	0.1%	-22.5%
313	Information Technology Technicians	31164	42571	1.5%	144.7%
321	Health Associate Professionals	33549	2808	0.1%	-74.2%
323	Welfare and Housing Associate Professionals	24577	46856	1.6%	80.1%
331	Protective Service Occupations	34952	3029	0.1%	-92.6%
341	Artistic, Literary and Media Occupations	34375	6381	0.2%	-80.1%
342	Design Occupations	45596	10375	0.4%	-25.1%
344	Sports and Fitness Occupations	29774	8470	0.3%	-31.0%
351	Transport Associate Professionals	35421	235	0.0%	-91.6%
352	Legal Associate Professionals	37626	6443	0.2%	11.7%
353	Business, Finance and Related Associate Prof.	41443	85773	3.0%	62.1%
354	Sales, Marketing and Related Associate Prof.	35470	150925	5.2%	88.6%
355	Conservation and Environmental Associate Prof.	25624	1127	0.0%	48.9%
356	Public Services and Other Associate Prof.	29428	98623	3.4%	132.6%
411	Admin. Occ.: Government and Related Organisations	31395	1812	0.1%	-96.3%
412	Admin. Occ.: Finance	28818	52018	1.8%	-24.7%
413	Admin. Occ.: Records	23728	27051	0.9%	-38.3%
415	Other Administrative Occupations	22249	44454	1.5%	-27.4%

Continued on next page

Table A.1 – continued from previous page

SOC	Name	Salary	Vacancies	Share	vs Census
416	Administrative Occupations: Office Managers and Sup.	28716	10931	0.4%	-42.5%
421	Secretarial and Related Occupations	28587	23336	0.8%	-72.6%
511	Agricultural and Related Trades	23400	3393	0.1%	-89.7%
521	Metal Forming, Welding and Related Trades	28009	4934	0.2%	-53.5%
522	Metal Machining, Fitting and Instrument Making Trades	32900	25082	0.9%	-18.3%
523	Vehicle Trades	30638	15090	0.5%	-48.9%
524	Electrical and Electronic Trades	41256	51441	1.8%	28.6%
525	Skilled Metal, Electrical and Electronic Trades Sup.	41915	1406	0.0%	-75.3%
531	Construction and Building Trades	40592	30457	1.1%	-64.9%
532	Building Finishing Trades	29613	3749	0.1%	-84.4%
533	Construction and Building Trades Sup.	37667	2878	0.1%	-37.1%
541	Textiles and Garments Trades	24473	393	0.0%	-91.7%
542	Printing Trades	24185	634	0.0%	-90.9%
543	Food Preparation and Hospitality Trades	24170	85146	3.0%	81.1%
544	Other Skilled Trades	28751	1090	0.0%	-89.1%
612	Childcare and Related Personal Services	22397	44325	1.5%	-42.6%
613	Animal Care and Control Services	23719	2133	0.1%	-71.2%
614	Caring Personal Services	20965	444155	15.4%	277.2%
621	Leisure and Travel Services	22782	8134	0.3%	-54.6%
622	Hairdressers and Related Services	23358	2221	0.1%	-92.4%
623	Housekeeping and Related Services	20995	7838	0.3%	-43.9%
624	Cleaning and Housekeeping Managers and Sup.	29088	25796	0.9%	256.8%
711	Sales Assistants and Retail Cashiers	20952	49762	1.7%	-71.4%
712	Sales Related Occupations	27191	12116	0.4%	-23.7%
713	Sales Supervisors	22404	5631	0.2%	-61.5%
721	Customer Service Occupations	22099	110675	3.8%	215.8%
722	Customer Service Managers and Sup.	28938	50708	1.8%	533.1%
811	Process Operatives	24503	1115	0.0%	-96.3%
812	Plant and Machine Operatives	23756	2563	0.1%	-88.9%
813	Assemblers and Routine Operatives	32306	6170	0.2%	-73.0%
814	Construction Operatives	27234	15322	0.5%	-24.7%
821	Road Transport Drivers	30771	123764	4.3%	38.7%
822	Mobile Machine Drivers and Operatives	24399	13550	0.5%	-3.6%
823	Other Drivers and Transport Operatives	28500	2512	0.1%	-64.7%
911	Elementary Agricultural Occupations	22736	1312	0.0%	-84.5%
912	Elementary Construction Occupations	24769	10550	0.4%	-40.8%
913	Elementary Process Plant Occupations	21575	7294	0.3%	-75.4%
921	Elementary Administration Occupations	21333	11370	0.4%	-53.5%
923	Elementary Cleaning Occupations	16478	76428	2.7%	-1.3%
924	Elementary Security Occupations	24552	44234	1.5%	33.4%

Continued on next page

Table A.1 – continued from previous page

SOC	Name	Salary	Vacancies	Share	vs Census
925	Elementary Sales Occupations	20926	564	0.0%	-95.4%
926	Elementary Storage Occupations	21762	62555	2.2%	33.9%
927	Other Elementary Services Occupations	20590	99993	3.5%	6.9%

Table A.2: Summary statistics

	Category	Mean	SD
Basic Statistics	Number of Ads	11662124	
	Mean length characters	2,297	1236
	Number of firms	128,559	
Baseline	No trait requirements	0.28	0.45
	No traits nor skills requirements	0.09	0.29
	Obs. with no requirements	89,620	
Traits	Any Personality Trait	0.72	0.45
	Conscientiousness	0.34	0.47
	Openness	0.30	0.46
	Emotional Stability	0.14	0.35
	Extraversion	0.44	0.50
	Agreeableness	0.13	0.34
Skills	Soft skills	0.62	0.48
	Hard skills	0.55	0.50
	AI skills	0.01	0.11
Seniority	High Seniority	0.01	0.08
	Middle Seniority	0.19	0.39
	Low Seniority	0.51	0.50
Education	No education specified	0.89	0.31
	GCSE	0.03	0.16
	A-level	0.04	0.19
	Bachelor's degree	0.03	0.17
	Master or Doctoral degree	0.02	0.12
Job schedule	Full time job	0.30	0.46
	Part time job	0.08	0.28
Job duration	Temporary job	0.13	0.34
	Permanent job	0.35	0.48
Experience	Experience requested	0.69	0.46
	No experience requested	0.01	0.09
Other job info	Remote Work	0.03	0.18
	Average Salary	31940	18690
	Negotiable salary	0.49	0.50
	Bonus offered	0.23	0.42

Notes: No trait requirements refers to the share of vacancies that do not mention words from Traits dictionaries. *No traits nor skills* refers to the share of vacancies that do not mention traits or Soft or Hard skills. *Obs. with no requirements* refers to the observations that did not match any of the job requirements in the text, hence the baseline sample for regression.

Table A.3: Personality Traits by 3-digit Occupation group

SOC	Name	P.T.	Agree.	Consc.	Emot.	Extra.	Open.
111	Chief Executives and Senior Officials	61.1%	11.0%	29.8%	10.6%	30.8%	24.4%
112	Production Managers and Directors	59.4%	4.2%	35.0%	7.8%	27.9%	23.0%
113	Functional Managers and Directors	68.7%	7.3%	33.6%	10.4%	34.6%	37.9%
115	Financial Institution Managers and Directors	63.7%	7.5%	35.5%	11.6%	33.0%	21.0%
116	Managers and Directors in Transport and Logistics	79.3%	5.9%	49.9%	10.2%	51.6%	26.1%
117	Senior Officers in Protective Services	68.2%	7.9%	30.2%	11.7%	40.4%	24.8%
118	Health and Social Services Managers and Directors	82.1%	19.2%	41.5%	16.8%	47.5%	35.0%
119	Managers and Directors in Retail and Wholesale	77.3%	6.0%	39.5%	15.7%	59.0%	33.0%
121	Managers and Proprietors in Agriculture Related S.	74.1%	10.3%	44.6%	8.3%	42.1%	29.1%
122	Managers and Proprietors in Hospitality and Leisure S.	81.5%	13.9%	40.8%	13.4%	64.0%	20.3%
124	Managers and Proprietors in Health and Care S.	75.2%	21.6%	36.3%	17.5%	40.6%	25.6%
125	Managers and Proprietors in Other Services	69.9%	9.8%	36.6%	12.7%	45.7%	24.7%
211	Natural and Social Science Professionals	71.5%	5.3%	26.7%	5.5%	37.6%	45.9%
212	Engineering Professionals	55.6%	3.2%	25.9%	6.3%	24.1%	29.4%
213	I.T. and Telecommunications Professionals	63.2%	4.0%	25.3%	5.7%	30.2%	39.4%
214	Conservation and Environment Professionals	74.7%	11.2%	35.6%	10.5%	49.8%	37.7%
215	Research and Development Managers	70.0%	8.9%	29.0%	5.9%	36.5%	48.1%
221	Health Professionals	69.7%	13.6%	23.3%	10.1%	42.7%	28.8%
222	Therapy Professionals	70.2%	10.2%	26.8%	9.6%	46.7%	21.1%
223	Nursing and Midwifery Professionals	75.1%	23.1%	39.3%	24.7%	41.1%	23.8%
231	Teaching and Educational Professionals	77.6%	11.1%	31.8%	13.2%	52.2%	32.5%
241	Legal Professionals	59.7%	10.4%	29.7%	6.4%	32.9%	18.9%
242	Business, Research and Administrative Professionals	64.5%	6.4%	32.9%	9.8%	31.8%	35.3%
243	Architects, Town Planners and Surveyors	51.3%	4.4%	28.7%	5.5%	24.4%	20.2%
244	Welfare Professionals	71.1%	15.1%	29.1%	17.2%	38.7%	37.7%
245	Librarians and Related Professionals	79.2%	12.9%	33.9%	14.0%	48.2%	46.2%
246	Quality and Regulatory Professionals	62.3%	5.1%	33.6%	8.5%	29.5%	29.5%
247	Media Professionals	82.9%	8.4%	31.9%	11.6%	50.2%	61.9%
311	Science, Engineering and Production Technicians	67.1%	4.9%	39.1%	10.6%	35.4%	26.5%
312	Draughtspersons and Related Architectural Technicians	57.7%	3.7%	27.3%	8.0%	27.2%	29.7%
313	Information Technology Technicians	69.5%	8.5%	34.1%	13.7%	41.4%	30.1%

Continued on next page

Table A.3 – continued from previous page

SOC	Name	P.T.	Agree.	Consc.	Emot.	Extra.	Open.
321	Health Associate Professionals	69.2%	9.4%	30.7%	7.2%	43.5%	19.6%
323	Welfare and Housing Associate Professionals	79.4%	18.0%	31.5%	24.4%	49.4%	39.3%
331	Protective Service Occupations	76.1%	11.2%	33.2%	26.4%	38.3%	31.8%
341	Artistic, Literary and Media Occupations	77.4%	6.2%	27.9%	10.2%	45.8%	51.8%
342	Design Occupations	84.7%	4.4%	25.0%	7.6%	48.0%	72.0%
344	Sports and Fitness Occupations	79.4%	16.2%	29.2%	5.5%	70.4%	24.4%
351	Transport Associate Professionals	50.9%	1.5%	31.6%	14.5%	26.4%	15.0%
352	Legal Associate Professionals	61.8%	8.2%	32.8%	10.7%	33.6%	20.6%
353	Business, Finance and Related Associate Prof.	65.3%	7.6%	35.2%	10.2%	34.2%	29.1%
354	Sales, Marketing and Related Associate Prof.	77.6%	7.8%	36.6%	11.5%	49.9%	39.7%
355	Conservation and Environmental Associate Prof.	79.8%	15.9%	32.3%	5.9%	59.1%	24.4%
356	Public Services and Other Associate Prof.	77.9%	9.8%	41.4%	15.8%	50.4%	32.0%
411	Admin. Occ.: Government and Related Organisations	61.5%	10.8%	24.8%	8.2%	38.9%	26.8%
412	Admin. Occ.: Finance	62.4%	6.2%	32.0%	10.7%	37.8%	19.9%
413	Admin. Occ.: Records	72.9%	7.2%	43.1%	15.2%	42.7%	23.7%
415	Other Administrative Occupations	76.8%	10.6%	44.7%	15.9%	47.8%	26.8%
416	Administrative Occupations: Office Managers and Sup.	73.8%	11.1%	43.7%	15.6%	50.0%	24.2%
421	Secretarial and Related Occupations	75.5%	17.5%	41.8%	18.3%	51.0%	24.9%
511	Agricultural and Related Trades	76.1%	9.1%	46.2%	4.2%	52.3%	27.3%
521	Metal Forming, Welding and Related Trades	33.2%	0.8%	20.6%	4.1%	11.0%	8.4%
522	Metal Machining, Fitting and Instrument Making Trades	49.7%	2.8%	29.2%	6.8%	22.0%	16.1%
523	Vehicle Trades	66.2%	5.8%	43.3%	12.6%	35.2%	16.0%
524	Electrical and Electronic Trades	53.4%	3.6%	27.5%	5.2%	23.1%	25.7%
525	Skilled Metal, Electrical and Electronic Trades Sup.	41.3%	3.0%	28.6%	5.9%	17.1%	10.0%
531	Construction and Building Trades	46.3%	7.6%	29.1%	3.1%	19.9%	12.9%
532	Building Finishing Trades	33.9%	3.6%	21.8%	3.0%	13.8%	8.7%
533	Construction and Building Trades Sup.	42.3%	2.6%	28.3%	4.2%	17.4%	12.1%
541	Textiles and Garments Trades	67.8%	8.1%	31.3%	8.5%	47.6%	25.3%
542	Printing Trades	59.4%	2.8%	28.0%	12.2%	27.9%	29.6%
543	Food Preparation and Hospitality Trades	81.2%	9.4%	41.6%	19.7%	58.7%	33.1%
544	Other Skilled Trades	65.8%	2.2%	41.4%	13.7%	33.9%	16.8%
612	Childcare and Related Personal Services	78.8%	17.5%	22.7%	18.4%	59.7%	34.4%
613	Animal Care and Control Services	81.2%	9.3%	28.6%	6.6%	61.6%	45.8%
614	Caring Personal Services	82.6%	39.1%	26.6%	27.4%	50.2%	29.3%

Continued on next page

Table A.3 – continued from previous page

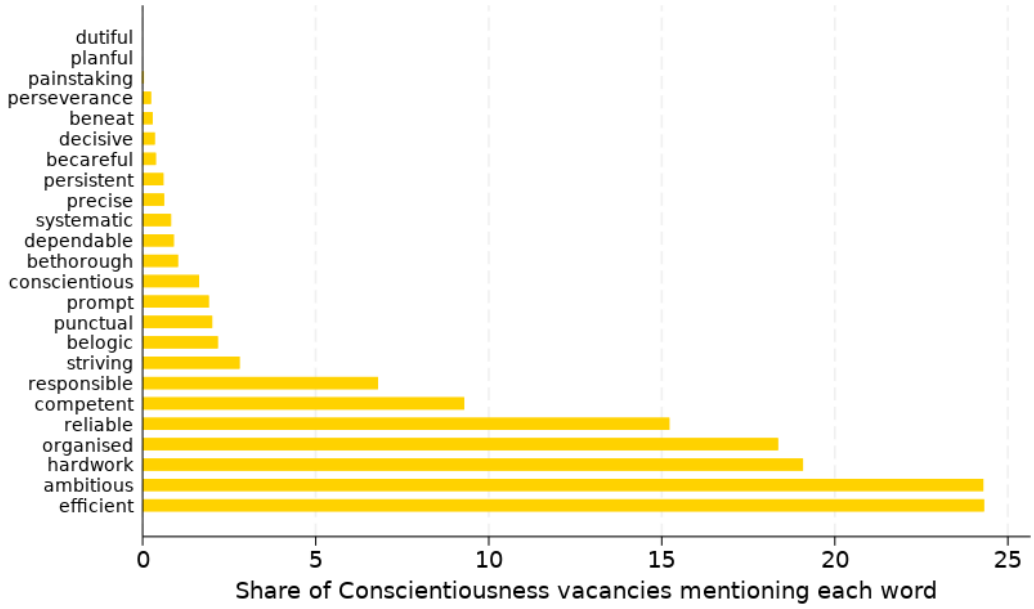
SOC	Name	P.T.	Agree.	Consc.	Emot.	Extra.	Open.
621	Leisure and Travel Services	83.6%	9.8%	51.7%	22.3%	67.6%	20.6%
622	Hairdressers and Related Services	78.8%	7.6%	33.1%	7.4%	61.8%	22.0%
623	Housekeeping and Related Services	72.2%	28.1%	33.8%	11.0%	37.9%	16.4%
624	Cleaning and Housekeeping Managers and Sup.	72.1%	20.5%	52.0%	15.4%	39.2%	15.8%
711	Sales Assistants and Retail Cashiers	87.3%	16.0%	42.8%	9.5%	76.4%	31.9%
712	Sales Related Occupations	82.1%	7.5%	35.2%	11.7%	49.3%	39.5%
713	Sales Supervisors	79.0%	13.0%	41.7%	4.7%	60.3%	43.8%
721	Customer Service Occupations	85.4%	17.8%	39.6%	17.7%	67.8%	30.3%
722	Customer Service Managers and Sup.	84.9%	13.0%	43.8%	20.9%	60.5%	30.1%
811	Process Operatives	49.5%	2.0%	29.6%	6.2%	21.4%	16.7%
812	Plant and Machine Operatives	49.2%	1.7%	29.8%	5.7%	24.5%	17.1%
813	Assemblers and Routine Operatives	52.0%	2.6%	29.3%	6.5%	26.5%	18.6%
814	Construction Operatives	62.6%	9.5%	39.7%	6.5%	31.6%	20.6%
821	Road Transport Drivers	52.6%	8.8%	33.5%	5.8%	27.9%	19.0%
822	Mobile Machine Drivers and Operatives	40.4%	5.2%	23.7%	4.0%	18.0%	10.9%
823	Other Drivers and Transport Operatives	61.9%	4.8%	45.3%	17.1%	30.8%	23.0%
911	Elementary Agricultural Occupations	66.9%	5.6%	38.3%	7.9%	32.4%	32.0%
912	Elementary Construction Occupations	20.2%	3.2%	13.9%	1.1%	6.6%	2.6%
913	Elementary Process Plant Occupations	55.6%	1.8%	29.8%	6.3%	29.2%	14.6%
921	Elementary Administration Occupations	88.6%	34.4%	48.5%	35.6%	42.4%	75.5%
923	Elementary Cleaning Occupations	63.9%	7.6%	39.4%	6.7%	33.5%	23.4%
924	Elementary Security Occupations	71.5%	9.4%	27.1%	11.0%	44.7%	34.0%
925	Elementary Sales Occupations	69.3%	10.8%	27.8%	4.6%	51.5%	17.1%
926	Elementary Storage Occupations	70.5%	6.3%	44.1%	6.2%	45.5%	19.5%
927	Other Elementary Services Occupations	82.6%	21.1%	38.5%	17.9%	67.0%	22.6%

Table A.4: Assortative matching: Personalities traits demand using UKHLS scores at the occupation level, no controls

	(1) P(Any P.T.)	(2) P(Agree.)	(3) P(Consc.)	(4) P(Emot.)	(5) P(Extra.)	(6) P(Open.)
Personalities score	0.023***					
	(0.000)					
Agreeableness score		0.089***				
		(0.000)				
Conscientiousness score			0.001***			
			(0.000)			
Emotional Stability score				−0.044***		
				(0.000)		
Extraversion score					0.089***	
					(0.000)	
Openness score						0.045***
						(0.000)
R^2	0.003	0.067	0.000	0.016	0.032	0.010

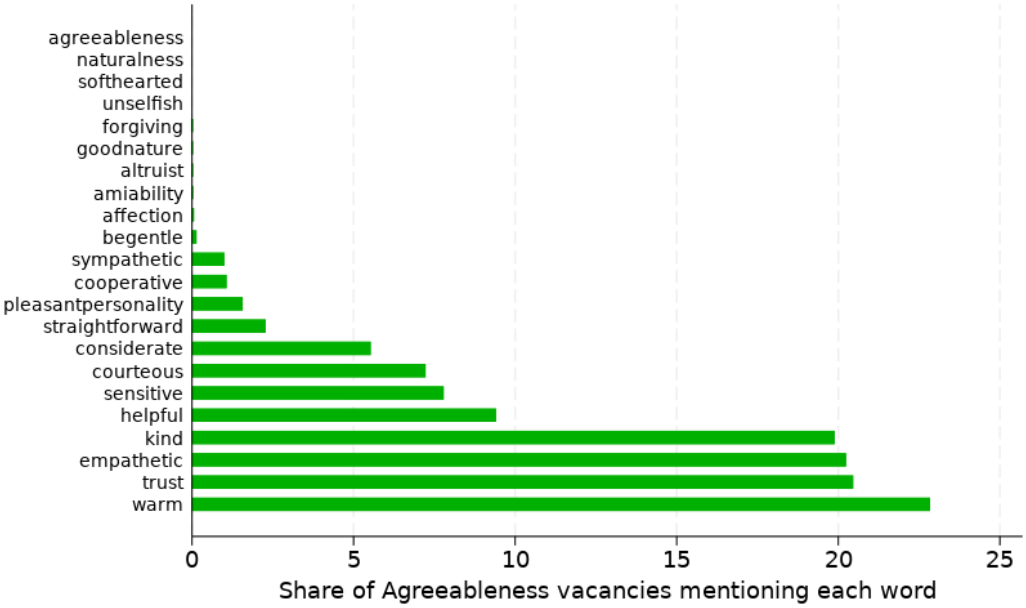
Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ $N = 9888315$. "Any P.T." refers to Any Personality Trait, "Consc." to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. All variables refer to the occupation-level 2011-2017 UKHLS standardized value. Author's creation using Adzuna and UKHLS.

Figure A.1: Frequency of words in Conscientiousness



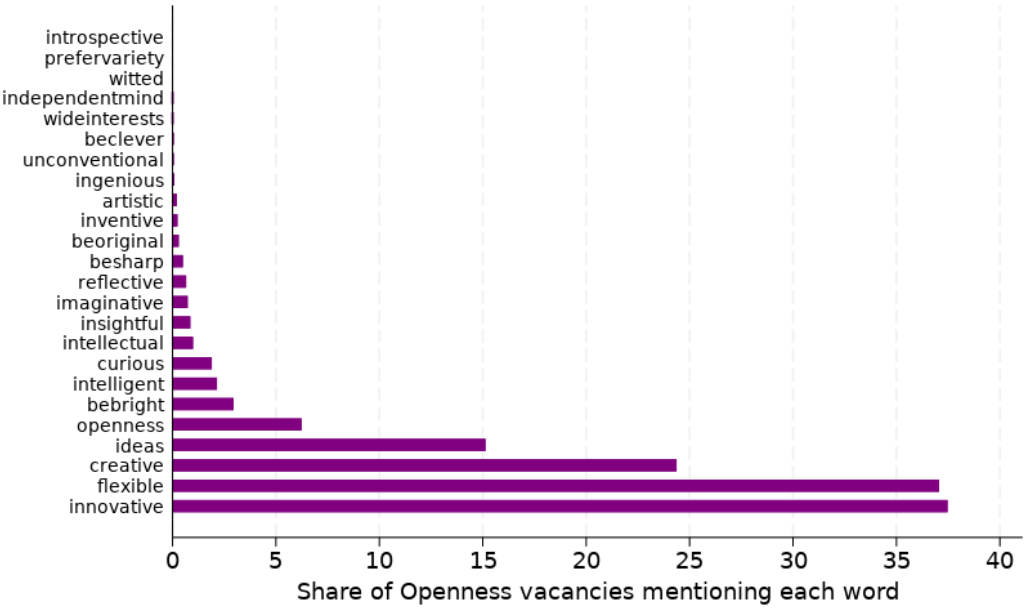
Notes: Author's creation using Adzuna.

Figure A.2: Frequency of words in Agreeableness



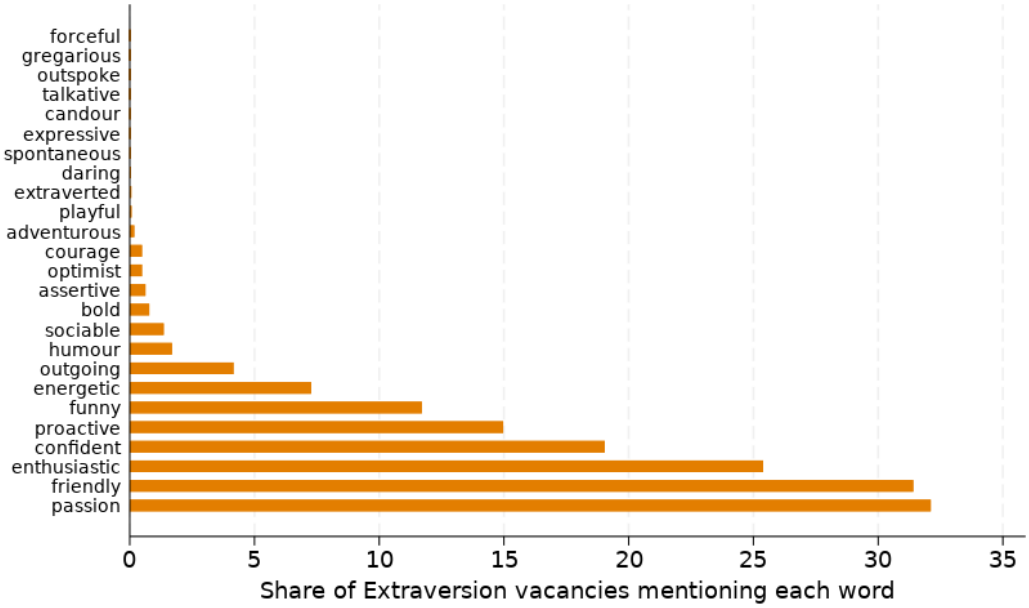
Notes: Author's creation using Adzuna.

Figure A.3: Frequency of words in Openness



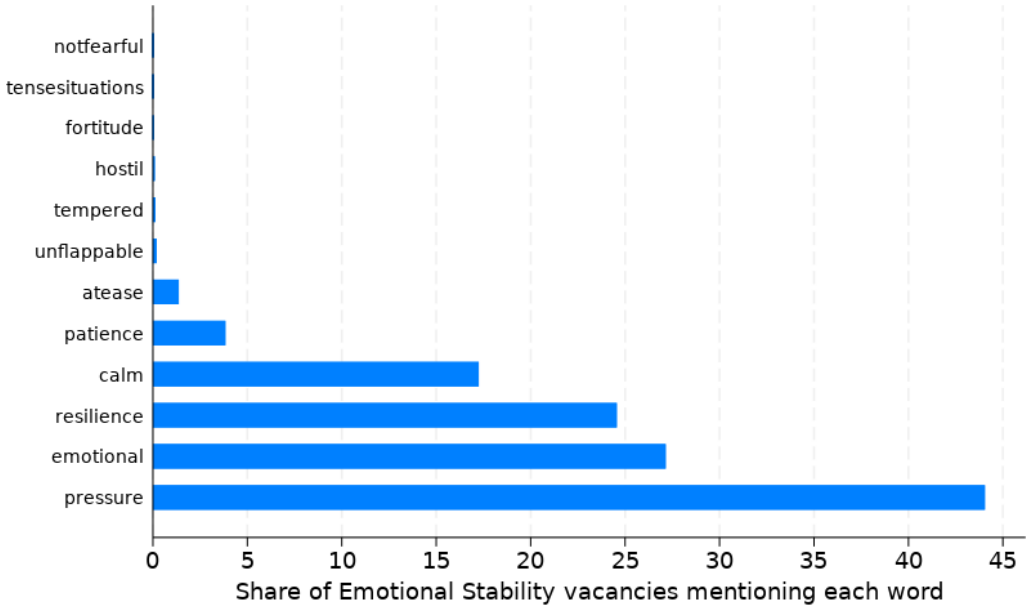
Notes: Author's creation using Adzuna.

Figure A.4: Frequency of words in Extraversion



Notes: Author's creation using Adzuna.

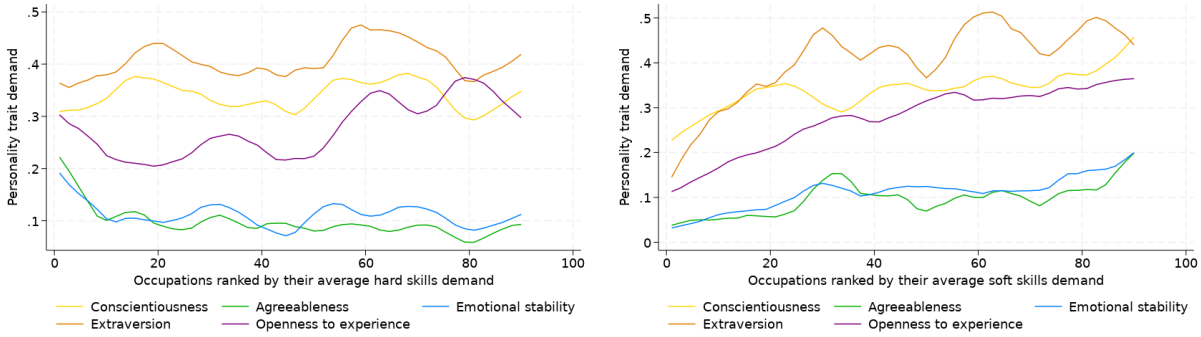
Figure A.5: Frequency of words in Emotional Stability



Notes: Author's creation using Adzuna.

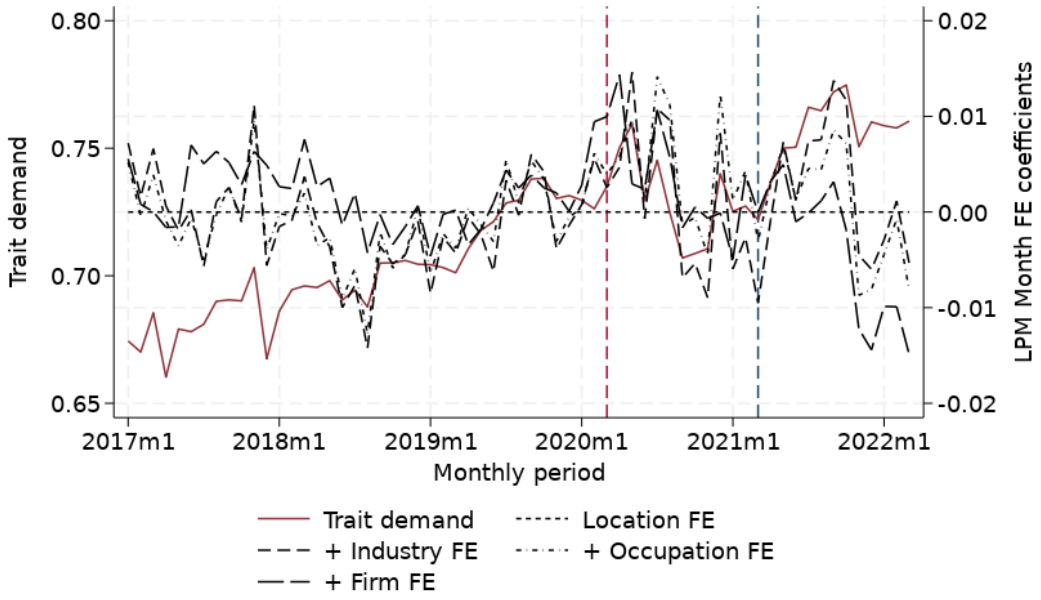
Figure A.6: Trait demand by hard and soft skill rank in Adzuna, at 3-digit SOC

(a) Demand by hard-skill ranking (b) Demand by soft-skill ranking



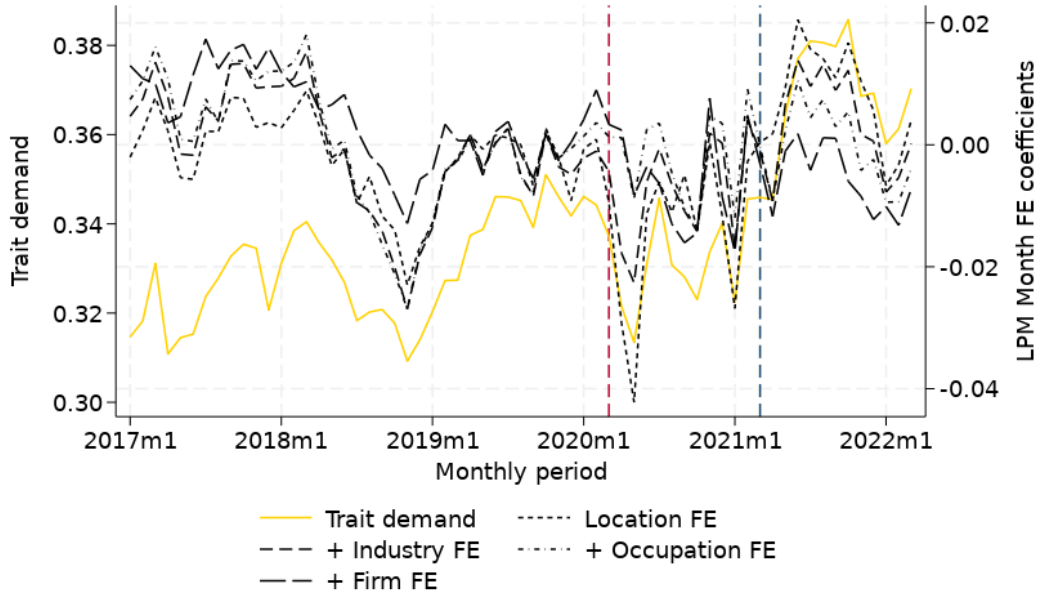
Notes: The figure plots a smoothed polynomial regression of the traits demand in each 3-digit SOC occupation against its rank in Adzuna demand of hard and soft skills by ranked occupations for 2017-2022. Author's creation using Adzuna.

Figure A.7: Month FE coefficients in all Personality Traits



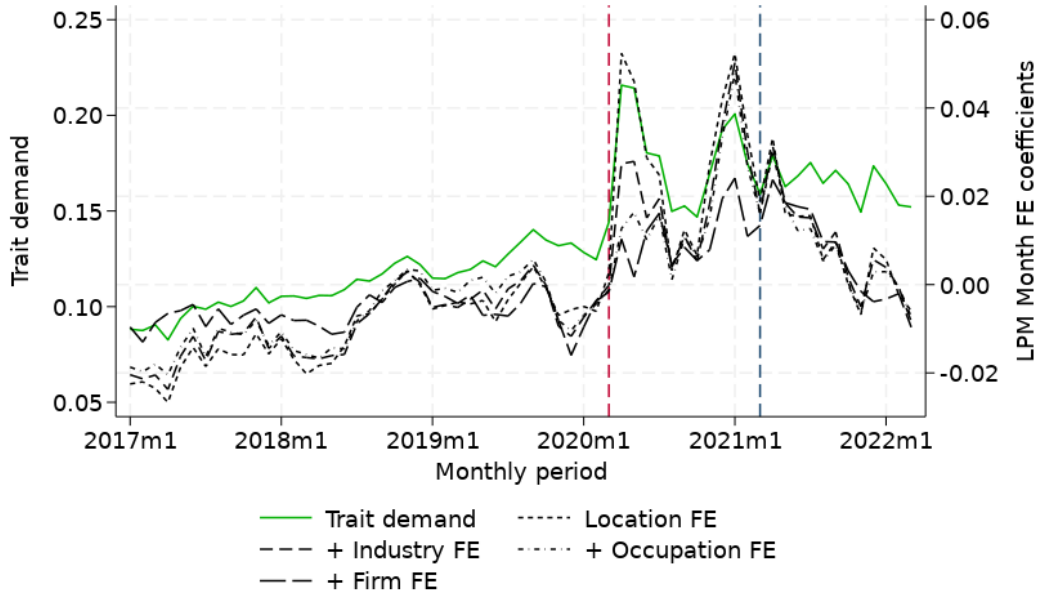
Notes: Dashed lines show the beginning (red) and end (blue) of Covid-19 restrictions. Author's creation using Adzuna.

Figure A.8: Month FE coefficients in Conscientiousness



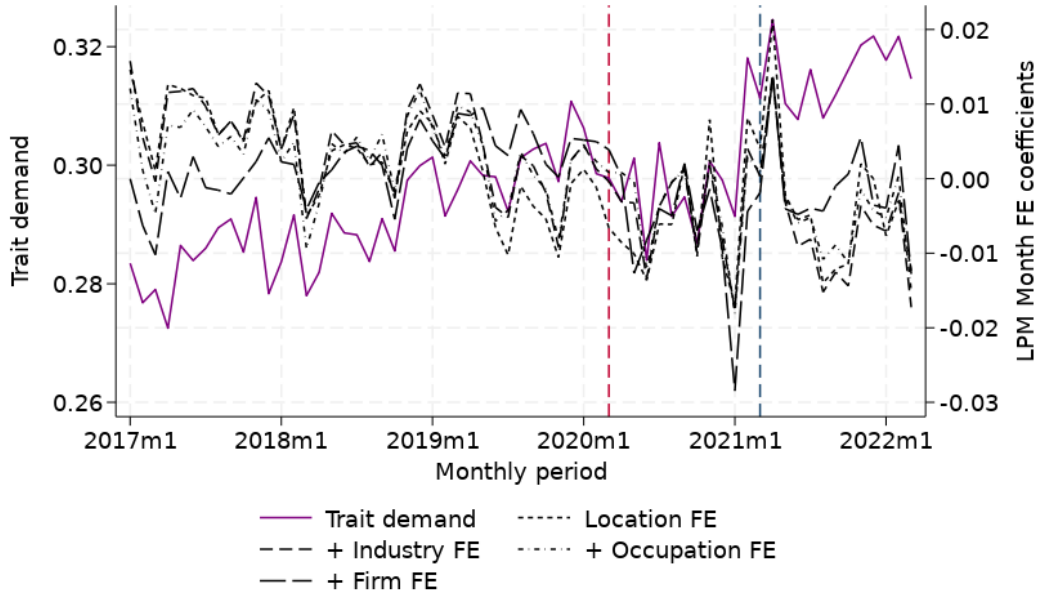
Notes: Dashed lines show the beginning (red) and end (blue) of Covid-19 restrictions. Author's creation using Adzuna.

Figure A.9: Month FE coefficients and trait demand for Agreeableness



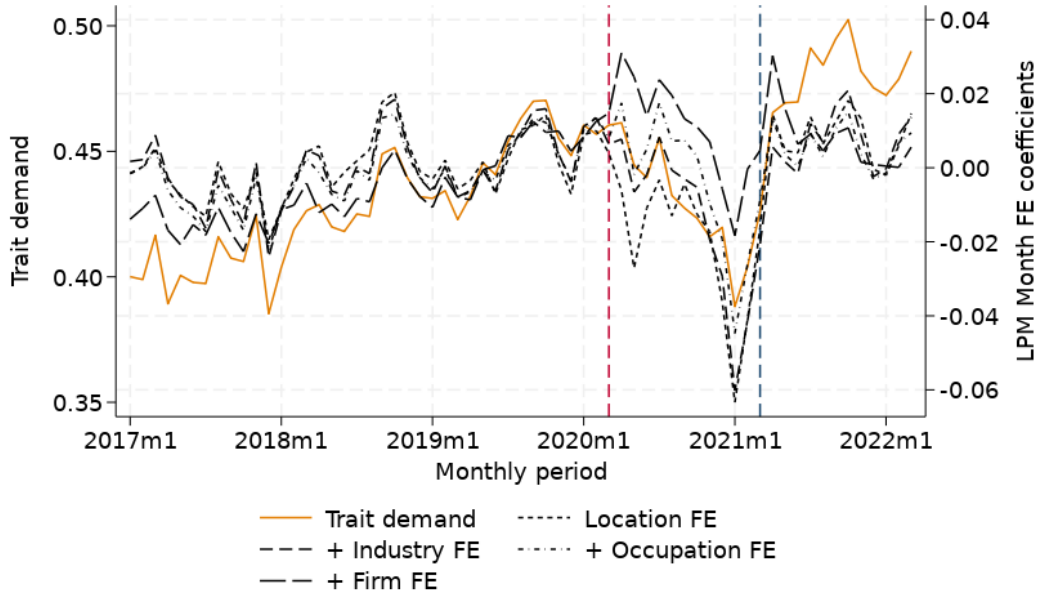
Notes: Dashed lines show the beginning (red) and end (blue) of Covid-19 restrictions. Author's creation using Adzuna.

Figure A.10: Month FE coefficients and trait demand for Openness



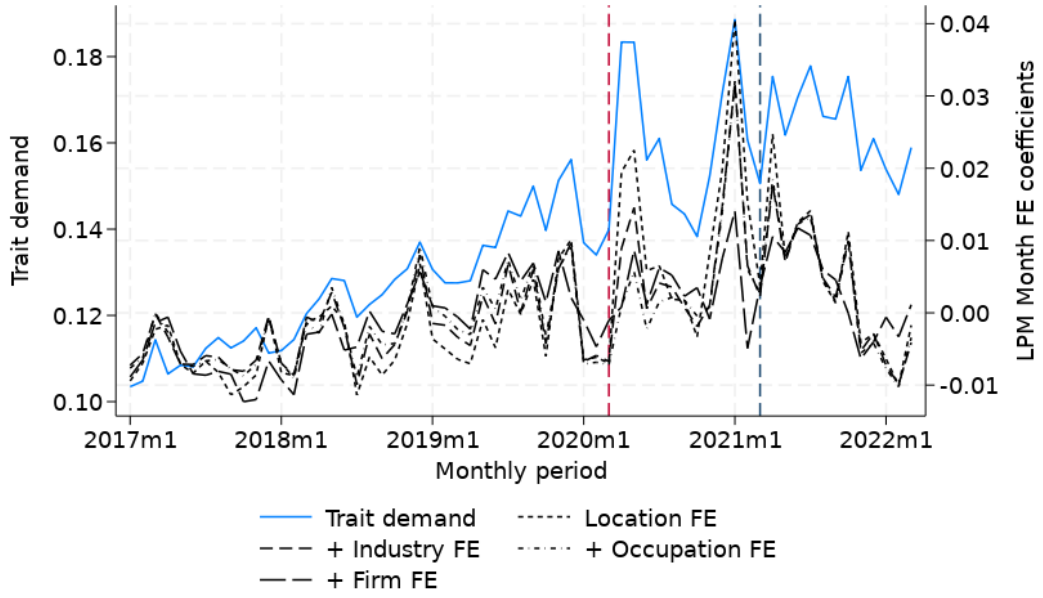
Notes: Dashed lines show the beginning (red) and end (blue) of Covid-19 restrictions. Author's creation using Adzuna.

Figure A.11: Month FE coefficients and trait demand for Extraversion



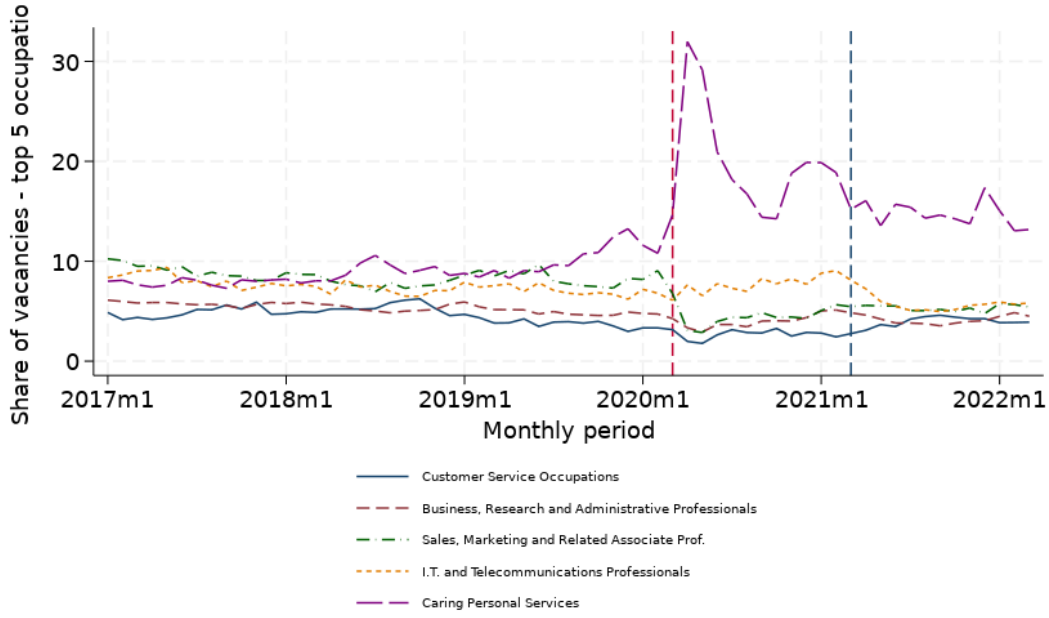
Notes: Dashed lines show the beginning (red) and end (blue) of Covid-19 restrictions. Author's creation using Adzuna.

Figure A.12: Month FE coefficients and trait demand for Emotional Stability



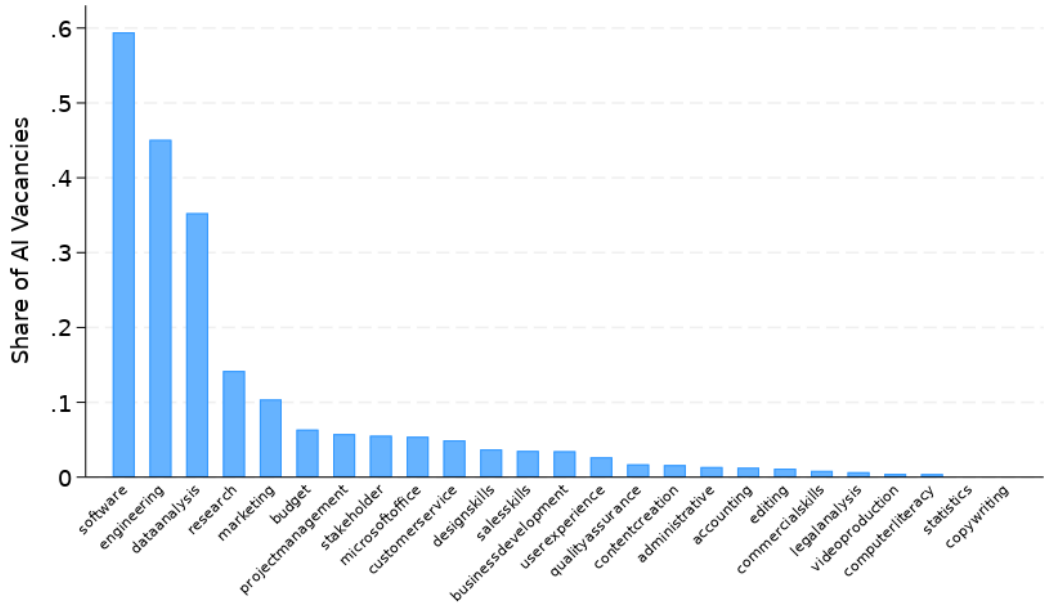
Notes: Dashed lines show the beginning (red) and end (blue) of Covid-19 restrictions. Author's creation using Adzuna.

Figure A.13: Change over time of top 5 occupations in Adzuna



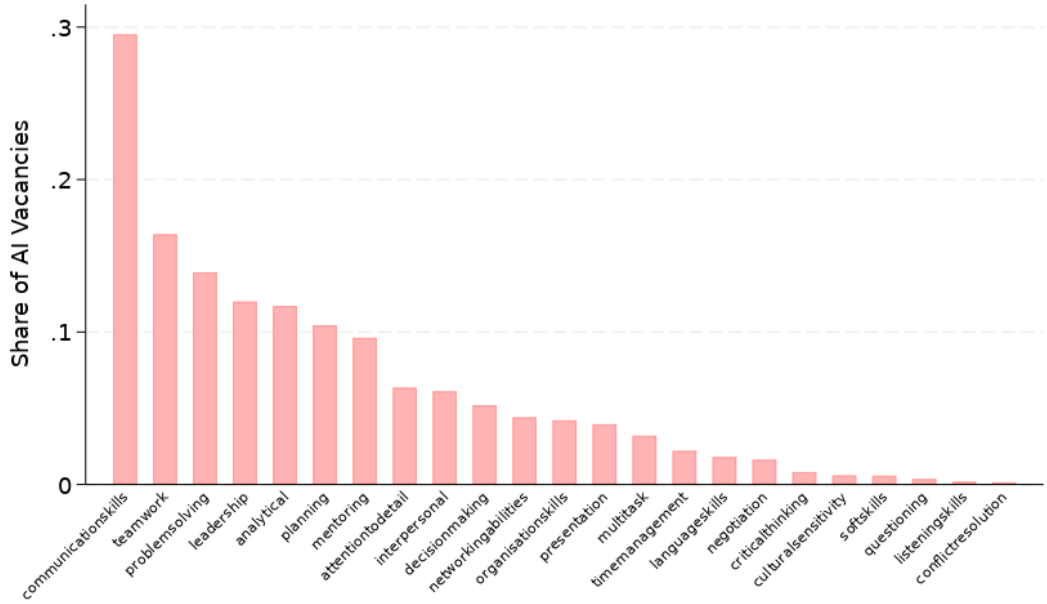
Notes: Author's creation using Adzuna.

Figure A.14: Hard Skills Demanded in AI Vacancies



Notes: Author's creation using Adzuna.

Figure A.15: Soft Skills Demanded in AI Vacancies



Notes: Author's creation using Adzuna.

Table A.5: Impact of AI on Personality Traits: Adding Controls Sequentially

	(1) Baseline	(2) +Controls	(3) +Time	(4) +Location	(5) +Firm	(6) +Occ	(7) All FE
Panel A: Overall Effect							
All Personality Traits	0.010***	-0.032***	-0.032***	-0.029***	0.012*	0.021***	0.016**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.007)	(0.005)	(0.007)
Panel B: Effects by Trait							
Agreeableness	-0.074***	-0.047***	-0.050***	-0.052***	-0.017***	-0.012***	-0.013***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.004)
Conscientiousness	-0.043***	-0.073***	-0.072***	-0.071***	-0.023***	-0.037***	-0.016**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.007)	(0.008)	(0.008)
Emotional Stability	-0.063***	-0.056***	-0.057***	-0.058***	-0.029***	-0.026***	-0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)	(0.009)	(0.007)
Extraversion	-0.092***	-0.114***	-0.113***	-0.112***	-0.038***	-0.041***	-0.026***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.008)	(0.009)	(0.007)
Openness	0.225***	0.171***	0.173***	0.172***	0.081***	0.120***	0.072***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.008)	(0.016)	(0.014)
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	Yes	Yes	Yes	Yes	Yes
Location FE	No	No	No	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	Yes	No	Yes
Occupation FE	No	No	No	No	No	Yes	Yes
R-squared range	[0.000, 0.003]	[0.078, 0.119]	[0.080, 0.120]	[0.084, 0.121]	[0.273, 0.351]	[0.105, 0.169]	[0.281, 0.364]
N observations	11662124	11662124	11662124	9923701	9888315	9923701	9888315
N firms	128559	128559	128559	117229	81843	117229	81843

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a different specification with progressively added controls and fixed effects. The R-squared range shows the minimum and maximum values across the six trait regressions for each specification. All regressions with controls include traits and skills demand, remote work requirements, length of advertisement, contract type, work schedule, bonus offerings, education, and negotiable salary option. Standard errors are clustered at the occupation and firm level in specifications (5)-(7).

Table A.6: Impact of Computer Skills on Personality Traits

	(1) Baseline	(2) +Controls	(3) +Time	(4) +Location	(5) +Firm	(6) +Occ	(7) All FE
Panel A: Overall Effect							
All Personality Traits	0.022***	-0.041***	-0.041***	-0.038***	0.003	0.001	0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.005)	(0.004)
Panel B: Effects by Trait							
Agreeableness	-0.053***	-0.041***	-0.041***	-0.042***	-0.009**	-0.005	-0.005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.005)	(0.003)
Conscientiousness	0.028***	-0.020***	-0.019***	-0.018***	0.000	0.005	0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.006)	(0.004)
Emotional Stability	-0.016***	-0.026***	-0.026***	-0.024***	-0.009**	-0.001	-0.005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.005)	(0.005)
Extraversion	-0.029***	-0.070***	-0.070***	-0.065***	-0.014***	-0.015*	-0.009
	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.009)	(0.006)
Openness	0.124***	0.070***	0.070***	0.068***	0.034***	0.037***	0.029***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.006)	(0.005)
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	Yes	Yes	Yes	Yes	Yes
Location FE	No	No	No	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	Yes	No	Yes
Occupation FE	No	No	No	No	No	Yes	Yes
R-squared range	[0.000, 0.010]	[0.080, 0.121]	[0.082, 0.122]	[0.084, 0.123]	[0.273, 0.351]	[0.105, 0.169]	[0.281, 0.364]
N observations	11662124	11662124	11662124	9923701	9888315	9923701	9888315
N firms	128559	128559	128559	117229	81843	117229	90

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a different specification with progressively added controls and fixed effects. The R-squared range shows the minimum and maximum values across the six trait regressions for each specification. All regressions with controls include traits and skills demand, remote work requirements, length of advertisement, contract type, work schedule, bonus offerings, education, and negotiable salary option. Standard errors are clustered at the occupation and firm level in specifications (5)-(7).

Table A.7: Comparing AI and Computer Skills Effects on Personality Traits

	(1) Baseline	(2) +Controls	(3) +Time	(4) +Location	(5) +Firm	(6) +Occ	(7) All FE
Panel A: Coefficient Differences (AI - Computer Skills)							
All Personality Traits	-0.012	0.009	0.009	0.009	0.009	0.020	0.013
Agree	-0.021	-0.006	-0.009	-0.010	-0.008	-0.007	-0.008
Consc	-0.071	-0.053	-0.053	-0.053	-0.023	-0.041	-0.020
Emot	-0.046	-0.030	-0.031	-0.033	-0.020	-0.025	-0.018
Extra	-0.063	-0.044	-0.043	-0.046	-0.023	-0.026	-0.017
Open	0.101	0.101	0.103	0.104	0.046	0.083	0.042
Panel B: T-Statistics							
All Personality Traits	-9.89	8.10	8.01	7.14	1.12	2.71	1.66
Agree	-31.15	-8.53	-10.01	-9.51	-1.48	-1.16	-1.57
Consc	-57.66	-45.01	-42.61	-37.01	-2.68	-4.06	-2.18
Emot	-62.00	-39.96	-33.88	-31.48	-3.21	-2.40	-2.06
Extra	-48.90	-34.98	-34.07	-31.77	-2.43	-2.04	-1.87
Open	75.53	80.89	87.37	77.17	5.12	4.71	2.86
Panel C: P-values							
All Personality Traits	0.000***	0.000***	0.000***	0.000***	0.264	0.007***	0.097*
Agree	0.000***	0.000***	0.000***	0.000***	0.138	0.247	0.116
Consc	0.000***	0.000***	0.000***	0.000***	0.007***	0.000***	0.029**
Emot	0.000***	0.000***	0.000***	0.000***	0.001***	0.016**	0.039**
Extra	0.000***	0.000***	0.000***	0.000***	0.015**	0.041**	0.061*
Open	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.004***
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	Yes	Yes	Yes	Yes	Yes
Location FE	No	No	No	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	Yes	No	Yes
Occupation FE	No	No	No	No	No	Yes	Yes

Notes: Panel A shows the difference between the effect of AI and Computer Skills when estimated separately. Panel B presents the z-statistics for these differences, calculated as $(\beta_{AI} - \beta_{CS})/\sqrt{SE(\beta_{AI})^2 + SE(\beta_{CS})^2}$. Panel C shows corresponding p-values, with stars indicating significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All comparisons use coefficients from separate regressions of each trait on AI and Computer Skills respectively.

Table A.8: Changes in AI Effects When Controlling for Computer Skills

	(1) Baseline	(2) +Controls	(3) +Time	(4) +Location	(5) +Firm	(6) +Occ	(7) All FE
Panel A: Coefficient Differences (AI only - AI with CS)							
All Personality Traits	0.013	-0.020	-0.020	-0.018	0.001	0.000	0.000
Agree	-0.029	-0.019	-0.019	-0.019	-0.002	-0.001	-0.001
Consc	0.018	-0.008	-0.008	-0.007	0.000	0.001	0.001
Emot	-0.008	-0.011	-0.011	-0.011	-0.002	-0.000	-0.001
Extra	-0.014	-0.031	-0.031	-0.029	-0.003	-0.003	-0.002
Open	0.066	0.030	0.030	0.029	0.008	0.009	0.005
Panel B: T-Statistics							
All Personality Traits	7.81	-13.38	-13.07	-10.54	0.06	0.00	0.05
Agree	-32.40	-21.04	-15.64	-13.52	-0.38	-0.28	-0.14
Consc	10.72	-4.87	-4.60	-3.70	0.02	0.12	0.06
Emot	-7.75	-11.66	-9.33	-7.37	-0.29	-0.00	-0.08
Extra	-7.95	-18.65	-18.16	-14.66	-0.27	-0.27	-0.16
Open	36.02	17.81	18.88	15.89	0.71	0.39	0.27
Panel C: P-values							
All Personality Traits	0.000***	0.000***	0.000***	0.000***	0.955	0.999	0.960
Agree	0.000***	0.000***	0.000***	0.000***	0.705	0.780	0.887
Consc	0.000***	0.000***	0.000***	0.000***	0.983	0.903	0.953
Emot	0.000***	0.000***	0.000***	0.000***	0.772	0.998	0.934
Extra	0.000***	0.000***	0.000***	0.000***	0.785	0.791	0.872
Open	0.000***	0.000***	0.000***	0.000***	0.481	0.693	0.786
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	Yes	Yes	Yes	Yes	Yes
Location FE	No	No	No	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	Yes	No	Yes
Occupation FE	No	No	No	No	No	Yes	Yes

Notes: Panel A shows the difference between AI coefficients estimated without and with computer skills controls. Panel B presents the z-statistics for these differences, calculated as $(\beta_{AI} - \beta_{AI|CS})/\sqrt{SE(\beta_{AI})^2 + SE(\beta_{AI|CS})^2}$. Panel C shows corresponding p-values, with stars indicating significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

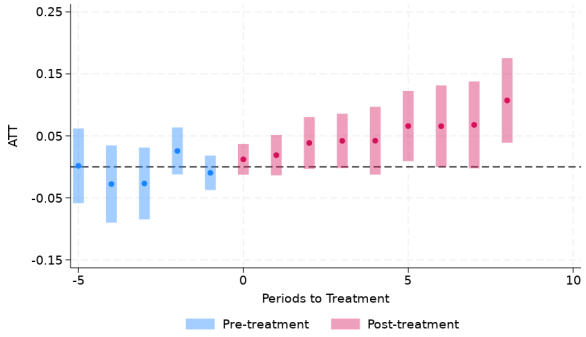
Table A.9: ATT Estimates: Half-year Aggregation

	Entropy Balance	Base	Controls
Panel A: Overall Effect			
All Personality Traits	0.041**	0.005	0.011
	(0.017)	(0.005)	(0.010)
Panel B: Effects by Trait			
Openness	0.064***	0.020***	0.045***
	(0.020)	(0.006)	(0.014)
Agreeableness	-0.007	-0.001	0.002
	(0.012)	(0.004)	(0.008)
Conscientiousness	0.027	0.006	0.001
	(0.021)	(0.006)	(0.012)
Emotional Stability	-0.010	-0.003	-0.017*
	(0.012)	(0.004)	(0.009)
Extraversion	0.025	-0.001	-0.004
	(0.021)	(0.006)	(0.014)
Pre-trend Test P-values			
All PT	0.060	0.067	0.090
Openness	0.000	0.846	0.001
Agreeableness	0.000	0.013	0.232
Conscientiousness	0.017	0.104	0.056
Emotional Stability	0.130	0.837	0.247
Extraversion	0.013	0.241	0.152
Number of Observations			
Observations	34,827	255,377	34,827

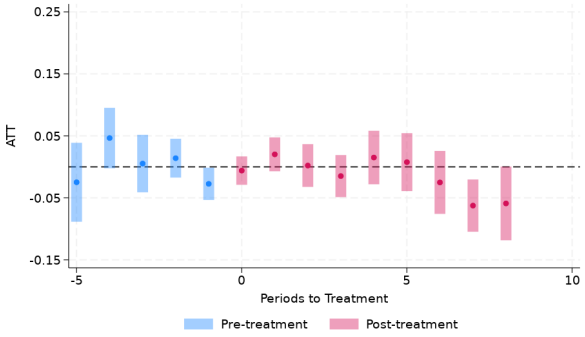
Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table compares ATT estimates using three different estimators at half year level. Standard errors in parentheses, clustered at firm level. Pre-trend test p-values from chi-square test of joint significance of pre-treatment coefficients.

Figure A.16: Staggered DID ATT. half-year aggregation

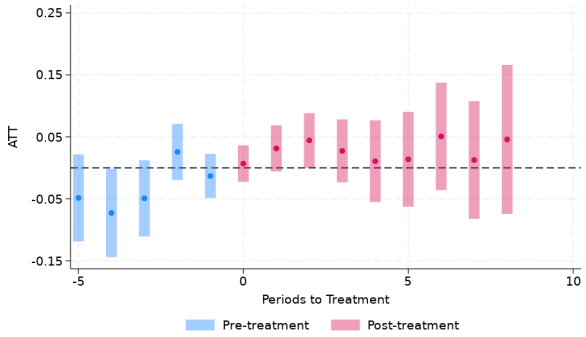
(a) All Personality Traits



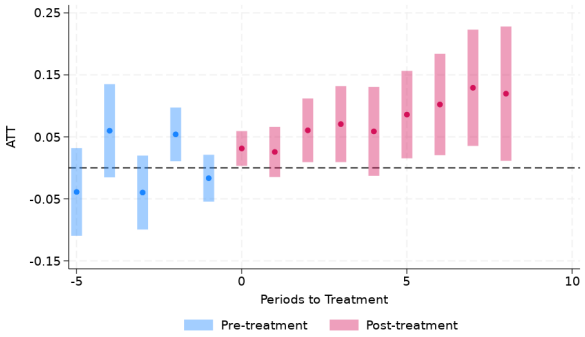
(b) Agreeableness



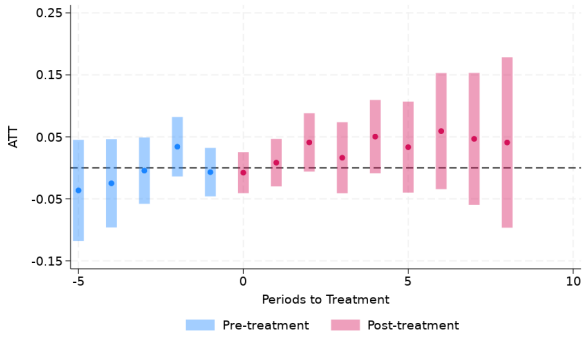
(c) Conscientiousness



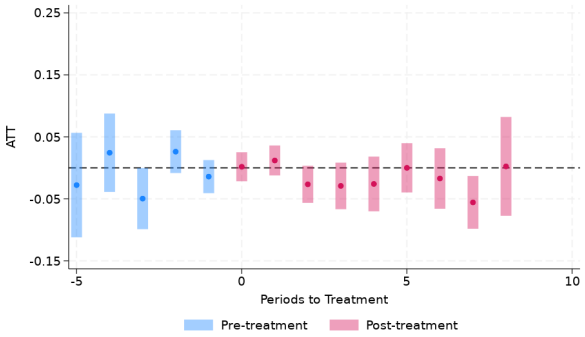
(d) Openness



(e) Extraversion



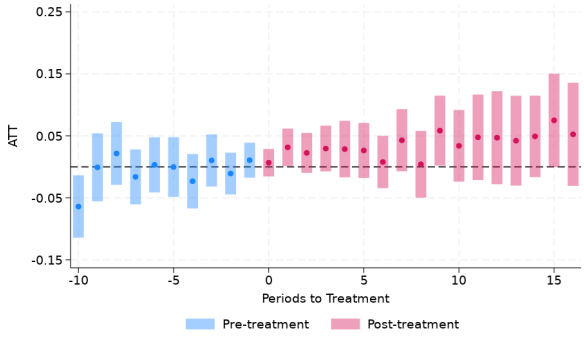
(f) Emotional Stability



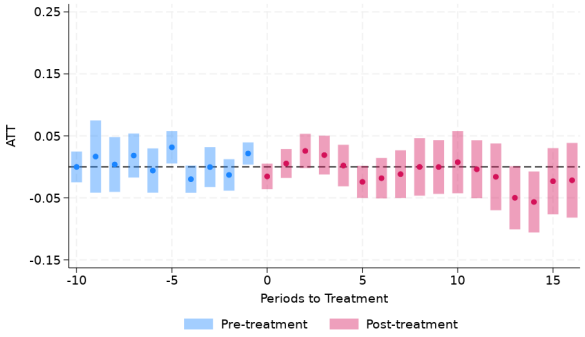
Notes: Author's creation using Adzuna.

Figure A.17: Staggered DID ATT. Quarter-level aggregation

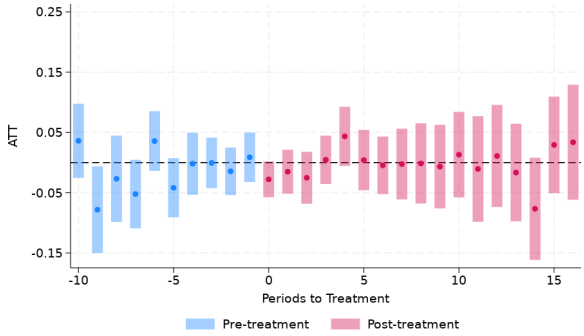
(a) All Personality Traits



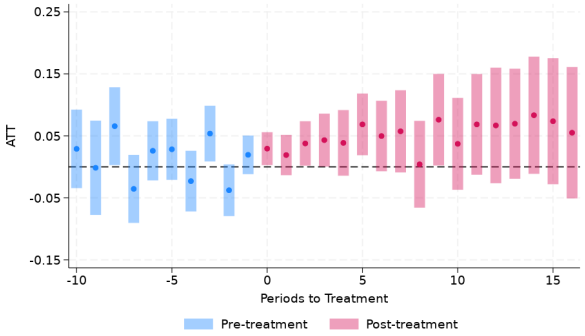
(b) Agreeableness



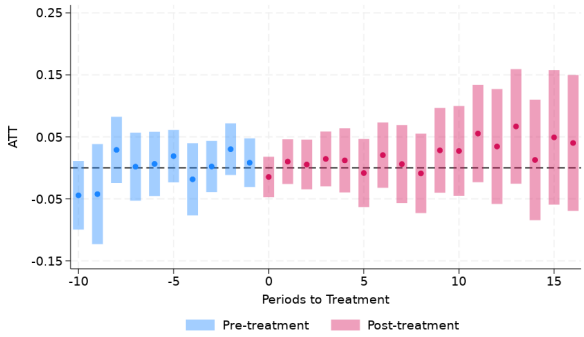
(c) Conscientiousness



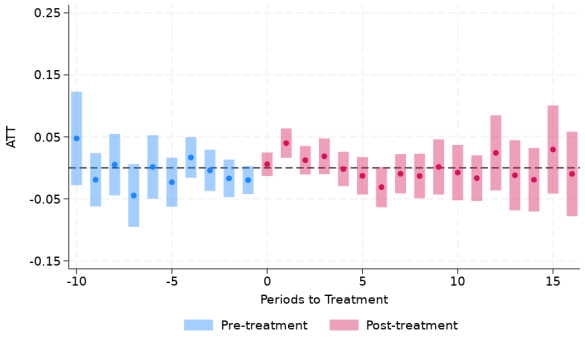
(d) Openness



(e) Extraversion



(f) Emotional Stability



Notes: Author's creation using Adzuna.

Table A.10: ATT Estimates: Quarter Aggregation

	Entropy Balance	Base	Controls
Panel A: Overall Effect			
All Personality Traits	0.031*	0.006	-0.007
	(0.016)	(0.005)	(0.010)
Panel B: Effects by Trait			
Openness	0.047**	0.019***	0.020
	(0.019)	(0.006)	(0.014)
Agreeableness	-0.007	-0.006	-0.008
	(0.012)	(0.004)	(0.009)
Conscientiousness	-0.003	0.001	-0.020
	(0.018)	(0.006)	(0.013)
Emotional Stability	0.004	0.004	-0.006
	(0.010)	(0.004)	(0.009)
Extraversion	0.013	0.005	-0.006
	(0.020)	(0.006)	(0.014)
Pre-trend Test P-values			
All PT	0.000	0.000	0.000
Openness	0.000	0.072	0.000
Agreeableness	0.000	0.005	0.000
Conscientiousness	0.000	0.000	0.000
Emotional Stability	0.000	0.022	0.000
Extraversion	0.000	0.003	0.000
Number of Observations			
Observations	52,626	361,993	52,626

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table compares ATT estimates using three different estimators at the quarterly-year level. Standard errors in parentheses, clustered at firm level. Pre-trend test p-values from chi-square test of joint significance of pre-treatment coefficients.

Table A.11: ATT Estimates for non-AI Positions: Year Aggregation

	Entropy Balance	Base	Controls
Panel A: Overall Effect			
All Personality Traits	0.025*	0.001	0.017*
	(0.014)	(0.006)	(0.009)
Panel B: Effects by Trait			
Openness	0.030	0.015**	0.043***
	(0.021)	(0.007)	(0.014)
Agreeableness	-0.014	-0.003	-0.003
	(0.017)	(0.005)	(0.009)
Conscientiousness	-0.002	0.001	0.007
	(0.019)	(0.007)	(0.012)
Emotional Stability	-0.025**	-0.004	-0.009
	(0.012)	(0.005)	(0.009)
Extraversion	0.003	0.002	-0.002
	(0.020)	(0.007)	(0.013)
Pre-trend Test P-values			
All PT	0.941	0.563	0.158
Openness	0.214	0.081	0.174
Agreeableness	0.081	0.325	0.125
Conscientiousness	0.024	0.826	0.451
Emotional Stability	0.365	0.446	0.395
Extraversion	0.775	0.584	0.349
Number of Observations			
Observations	22,009	174,250	22,009

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. ATT estimates for Non-AI positions only. Treatment (AI adoption) based on first AI vacancy posted by firm. Standard errors in parentheses, clustered at firm level. Pre-trend test p-values from chi-square test of joint significance of pre-treatment coefficients.

Appendix B - Notes on the cleaning process

Other datasets used

Companies House:

Companies House is an executive agency of the UK government that provides information about registered and dissolved companies in the UK. Companies House includes 5,166,124 companies registered in the UK since 1885. It provides information for most companies' addresses, postcode of registered office, and a 4 to 6 digit Standard Industrial Code (SIC). After removing companies with the same name (hence impossible to differentiate), 4,832,015 unique companies remained. Since SIC codes are only provided for private companies, public institutions like the NHS, Universities, or Charities were not considered. Institutions and dormant companies whose SIC codes are generic (99999) were also not considered. The final Companies House dataset consists of 4,581,872 private companies.

FAME:

A financial database covering 3.4 million companies in the UK and the Republic of Ireland. The dataset contains information on company accounts, activities, ownership and management. I retrieved information on whether companies are publicly listed.

UK Household Labour Survey:

Also known as Understanding Society, the survey is a large-scale longitudinal study that aims to provide valuable insights into the social and economic aspects of individuals and households over time. Surveys a wide range of information from participants, including demographic data, employment status, income, health, and subjective well-being. Data were collected with the same methodology from 2009 until 2022.

In Wave C, performed during 2011 and 2012, the survey requested individuals to complete a short questionnaire on self-perceived attitudes, where the Big Five traits were inferred. The module presented three affirmations for each personality, and the individuals had to choose a number that best described how they saw themselves *as persons*, using a scale from 1 to 7, where 1 meant *does not apply to me at all* and 7 meant *applies to me perfectly*[20]. Then, an average score of the three responses for each trait was calculated.

- **Agreeableness:**
  - I see myself as someone who is sometimes rude to others.
  - I see myself as someone who has a forgiving nature.

[20]UKHLS creates a score for the *Neuroticism* trait. For comparison purposes with the demand side study, the values had been inverted (1/x) to infer *Emotional Stability*.

  - I see myself as someone considerate and kind to almost everyone.
- **Conscientiousness:**
  - I see myself as someone who does a thorough job.
  - I see myself as someone who tends to be lazy.
  - I see myself as someone who does things efficiently.
- **Extraversion:**
  - I see myself as someone talkative.
  - I see myself as someone who is outgoing, sociable.
  - I see myself as someone who is reserved.
- **Neuroticism:**
  - I see myself as someone who worries a lot.
  - I see myself as someone who gets nervous easily.
  - I see myself as someone who is relaxed, handles stress well.
- **Openness:**
  - I see myself as someone who is original, comes up with new ideas.
  - I see myself as someone who values artistic, aesthetic experiences.
  - I see myself as someone who has an active imagination.

Cleaning of the dataset

The main purpose of the cleaning strategy was to standardise text entries, in particular the name of the company, location, and descriptive variable of the advertisement. To create an occupation variable, the Turrell et al. (2019) algorithm was used. The code requires as inputs advertisement text, job title, and industry. For the latter, companies variables were matched with Companies House. Cleaned text was later used as input to retrieve job requirements, wages, and location.

Since the entire dataset was too large for the cleaning process to be conducted in one attempt (646 GB), it was first divided into 272 weekly snapshots, and each snapshot was subdivided into 1 per cent samples using the last 2 digits of the vacancy identifier. Each of these 27,200 datasets contained around 10,000 observations.

Companies names cleaning process

Adzuna and Companies House companies lists were cleaned using the same code. The process, in STATA, involved the use of regular expressions to identify characters and words. The process intended to minimise the potential discrepancies between the same name in both datasets, therefore eliminating all non-relevant information to the name and homogenising formats. It included lowering the case; removing consecutive internal blank spaces, quotes, and special characters; changing symbol characters for their word counterpart (& for *and*), dropping endings such as "*.sol*" or "*.com*"; removing "*The*" at

the beginning, and replacing text not encoded as American Standard Code for Information Interchange (ASCII) into ASCII.

In Adzuna, some companies' names identify the place where the job is located, such as *KPMG (Liverpool)*, or *Accenture (U.K.)*. This information is also available in the "location" variable, so its presence in the company name stops the algorithm from matching the company with Companies House. In consequence, the information inside the brackets at the end of a company was also removed. A considerable difference between Adzuna and Companies House is that the former frequently does not incorporate companies' categories, such as *Private Limited Company, Public Limited Company, Limited Liability Company* or its diverse acronyms and synonyms *(plc, llc, ltd, plc, lp, pcc, plc, scio, cio, cic, ukeig, limited, srl, inc, co, sp, group)*. These were removed if they appeared at the end of the company name.The cleaning process was repeated 3 times to ensure that companies with several inaccuracies in their names could be exhaustively cleaned ("*K.F.C (Liverpool) LTD++*" into "*kfc*").

Matching companies with Companies House

Since companies' names in Adzuna could have diverse variations, the cleaning process consisted of generating a dataset with all the existing entries of companies in Adzuna and applying a cleaning code. This *matching dataset* was then associated with the Companies House cleaned list of companies before removing public and dormant companies, merging into a unique data set both Adzuna's raw company name, the clean name common to Companies House, and all Company information matched from Companies House. Nearly half of the unique company entries in Adzuna (337,461) were matched. The remaining companies either present a fantasy name different from the legal name, which means they were not listed in Companies House or have missing company names.

Interestingly, out of the 4.5 million active and inactive companies listed in Companies House, less than 10 per cent were found in online job vacancies. This is explained by the high prevalence of Recruitment Agencies, whose growth since the global crisis in 2008 reshaped companies' strategy to recruit employees. Low barriers and increased digitalisation have benefited the industry over the past years. According to the Recruitment and Employment Confederation (REC), by 2021, more than 30,000 recruitment enterprises were operating across the UK. In the same year, nearly 550,000 permanent placements and 22.4 million temporary contracts were made by the recruitment industry (REC, 2022).

In addition, half of the companies in the Adzuna vacancy dataset could not be matched with a company in the CH data. The main reason is that companies do not necessarily publish their legal registration name when posting a job opening. These may appear with a shorter version, an acronym or a more known fantasy name. Even though the cleaning

algorithm aims to unify criteria in terms of format and accessory wording, if a company publishes job advertisements using a highly different name than its legal name, then it is impossible to match it with Companies House. As an example, a company may appear in online postings as *The Berliner*, but their legal name in Companies House incorporates their complete identification *The Berliner Beer House*.

Finally, some companies may share a similar part of their legal name, (*The Berliner Services, The Berliner International, The Berliner Central*). When companies have an almost identical name, it is impossible to systematically identify which version of the name in Companies House corresponds to the clean name in Adzuna. To avoid false-positive matching, whenever a company's clean name was identical to another, none of them was incorporated.

Finally, those vacancies with company names not matched and therefore not identifiable were removed. Therefore, all the analysis done was on vacancy data which contained a matched firm name.

Recruitment agencies and companies not matched removal

In the Adzuna Database, 45 per cent of vacancies were posted by recruitment agencies, hence making it impossible to identify either the company behind the search nor the job's sector. It is possible that many enterprises, in particular small businesses, subcontract the human resource task and pay recruitment agencies to be in charge of the first recruiting stages. In addition, many existing enterprises may still not use online job boards to advertise their search, still relying on traditional means such as newspapers, word-of-mouth, or other types of referrals. Robustness checks on published wages show no significant differences between jobs posted through Recruitment agencies and those published by companies themselves.

Recruitment agencies were identified and removed following three strategies; firstly, company names containing words such as "*recruit*", "*employ,*", "*staffing*", "*personnel*", "*resourcing*", "*careers*", "*talent*" or "*placement*" were removed. These words were mostly associated with companies in SIC codes 62090 (other information technology service activities), 7810 (activities of employment placement agencies), 78200 (temporary employment agency activities), 78300 (human resources provision) and 82990 (other business support service activities), so all vacancies in these industries were removed as well. A manual inspection of companies that had a high rate of job advertisements led to the removal of another 33 recruitment agencies, including CV Library, Reed and Brook Street.

To identify companies' sectors, Standard Industrial Classification (SIC) is needed. Only private companies in the Companies House are assigned this information. Consequently, only vacancies with the company were found in both Adzuna and Companies House, and

with no missing SIC, codes were kept. This led to a total of 64.6 million vacancies with associated company names and sectors not published by recruitment agencies.

Duplicates removal

Once companies' names were cleaned and matched with Companies House, and recruitment agencies were deleted, databases with the same 2 last digits of the ID variable were appended, to proceed with the removal of the duplicates.

One issue with vacancies aggregators is that an advertisement remaining open on a website does not necessarily reflect that the position has not been filled, nor vice-versa. While some companies have their advertisements posted for months or even years, constantly receiving applications, other companies may pay for a fixed advertisement time but fill the position before it ends, or not fill it at all. It becomes necessary then to deduplicate the database to only work with unique vacancies.

Using the ID variable provided by Adzuna, I kept the first observation of each vacancy that appears in the dataset. Two more rounds of duplicate removals were performed with other methods. First, by grouping variables with the same job title, company name, date created and source-id variable. In case some of the vacancies posted on different web pages were not initially captured by the Adzuna ID vacancy variable, a last duplication removal round was performed by deleting repeated observations with the same company, title, date of creation, description, and location. Since only a third of the total observations belong to unique vacancies, the number of vacancies was reduced to 20.7 million.

Cleaning text descriptions and generating new wage variables

The salary variable provided by Adzuna was coded from complementary information on web pages, and not necessarily from the text descriptions. Around 65 per cent of observations in the raw data had a non-missing salary. In comparison with other studies, Adzuna's wage variable is significantly high. For instance, wages are advertised in just 20 per cent of job advertisements in Marinescu and Wolthoff (2020) using CareerBuilder and 16.4 per cent of job advertisements in Kuhn et al. (2020) using a Chinese job portal. It is possible that, by only working with vacancies that have a posted wage, the sample becomes less representative of the general labour market. In particular, Brenčič (2012) shows that posted wages are more common in advertisements seeking less skilled workers.

When considering only the dataset with Companies House matching, recruitment agencies and duplicate removal, only 9.6 million vacancies had a non-missing wage. However, some of the entries in Adzuna's wage variable seemed to have errors or were coded in different time frames (daily/weekly/monthly rate instead of annual). I provide a robustness exercise of wage variable and revise Adzuna's entry, by retrieving salary information

from the advertisement text. This entailed, first of all, cleaning job descriptions.

Similar, to the cleaning process for companies' names, the descriptions of the vacancies were cleaned using regular expressions. This process involved lowering the case, replacing HTML entities for known terms (*£* for £ or *⋅* for "."), adding space to words besides special characters or numbers (changing *the salary will be:£35000.and the tasks will be* for *the salary will be: £ 35000 . and the task will be*) for easy identification of wages. Web pages, quotes, and special characters were removed. Finally, contractions such as "we're" or "you'd" were changed for "we are" and "you would". It followed a meticulous cleaning of the earnings text inside jobs' descriptions. Salaries in vacancies were posted in different formats. Table B.1 exemplifies the wide variety of arrangements found in job descriptions:

Table B.1: **Examples of wages differences in formats**

Differences	**Unique salaries**	**Range of salaries**
Punctuation and spaces	£35,000 £ 35000 £35, 000	£30,000-£40,000 £ 30.000 - £40000
The use of k for thousands or the lack of £ sign	£35k 35000 35k	£30k-£40,000 £ 30.000 - 40k 80 - 120000
Range separators		£30,000 to £40,000 £ 30000 - 40k £35. 000 and 40000
Monthly date	£1300 per month	
Weekly date	£500 per week	
Daily date	£80.5 per day	
Hourly date	£9 9 per hour	£8.46-10

Notes: Author's creation based on Adzuna.

The homogenising process involved replacing 2-digit annual wages for 5-digit numbers, adding pound signs when these were missing, removing empty spaces and characters in between large numbers, removing "k" when appearing after a number, and not considering numbers when referring to other items. A list of 25 common words that appear in text entries after 5-7 digit numbers were eliminated (for example, 30,000 *employees/organisations/passengers/people/clients, etc.*). An analogous cleaning exercise was done for monthly, weekly, daily and hourly information.

Once the text was cleaned, the wage was retrieved, both as unique payments and when it was offered between ranges. To construct a new earnings variable, wages which

referred to annual payments were considered first. If this information was not found, then the mean of the lower and upper-range payment offers was considered. This strategy was also used by Chaturvedi et al. (2021) and Kruyen et al. (2020).

If this was not retrieved, then monthly, weekly, daily and hourly rates were considered following the same strategy and taking into account whether the job was part-time or full-time. Finally, if no information could be retrieved from the text, then Adzuna's salary variable was considered, after being revised so that all salaries were computed annually. The creation of the new salary variable out of the text entries increased the dataset by 2,446,300 observations.

A last cleaning on duplicate vacancies was performed, by dropping observations with the same source-id and by dropping observations with the same company, job title, date created, description, and location. The final dataset, with a non-cero new salary, includes 11,430,457 observations. The loss of vacancies history is described in Table B.5.

Creating NUTS location variable

Adzuna presented two location variables: *location_raw*, which could contain entries such as "Newcastle, Staffordshire" and *location_path*, containing long entries such as "UK ∼ North West England ∼ Greater Manchester ∼ Manchester ∼ The Trafford Centre". The first variable is the raw information extracted from the job vacancy, sometimes only incorporating the town where the job is located and, in a few cases, it also incorporates the postcode. The second is coded by Adzuna, but on some occasions, it does not incorporate the full path but only "UK", or "UK∼London". When it does follow the path, the information is always accurate. The approach followed was to first use *location_path*, then complement those observations with a missing detailed location variable with *location_raw*. Ultimately, they were complemented using the information from postcodes. Table B.2 is an example of *location_path* and the resulting subdivision[21]:

Table B.2: **Location example**

Loc. 1	Loc. 2	Loc. 3	Loc. 4	Loc. 5	Loc. 6
UK	North West England	Greater Manchester	Manchester	The Trafford Center	-

Notes: Author's creation based on Adzuna.

Similar to the process for matching the companies, the strategy was to match locations with a nomenclature that allows to use known location categories. Locations were mapped with the Nomenclature of Territorial Units for Statistics (NUTS). The NUTS identification at one digit (UK + letter) corresponds to 12 UK regions, as shown in Table B.3. A list containing the 3,102 unique locations found in Adzuna's location variable was created and assigned a NUTS code at 2 digits (UK+letter+number). Then, each job vacancy was

[21] All examples are illustrative and do not represent any actual vacancy information.

associated with this list, to obtain a matching as the one in Table B.4.

Table B.3: **Nuts 1**

NUTS	Code
North East	UKC
North West	UKD
Yorkshire and The Humber	UKE
East Midlands	UKF
West Midlands	UKG
East of England	UKH
London	UKI
South East	UKJ
South West	UKK
Wales	UKL
Scotland	UKM
Northern Ireland	UKN

Table B.4: **Nuts 2**

NUTS	Code
Belper	UKF1
Ely	UKH1
Mortlake	UKI7
Porthmadog	UKL1
Newton-Le-Willows	UKD7
Barnacle	UKG3
Sea Mills	UKK1
Cross Keys	UKL2
Ickenham	UKI7
Maghera	UKN0
Darlington	UKC1

Notes: Author's creation based on ONS.

A similar approach was taken for all the components of *location_raw*: they were split and matched with their corresponding NUTS codes at 2 digits. For those observations with missing information on Location 3 onwards, postcodes were used to retrieve the corresponding NUTS. These belong to four sources: first, Adzuna provided a variable with postcodes, although it had 70 per cent missing values. For those missing, postcodes in the location variable or the title of the job were retrieved. Finally, Companies House provides the location and postcode of the companies' headquarters. For those observations with job locations set in London, Companies House postcodes were used. Postcodes were ultimately matched with NUTS. In total, 96 per cent of observations have non-missing values of NUTS at 1 digit and 85 per cent at 2 digits.

Adding Standard Occupational Classification (SOC)

The major disadvantage of Adzuna's database is that it does not come with official classification labels. I implemented a Natural Language Processing (NLP) matching algorithm that maps the free-form text of job descriptions into UK Standard Occupational Classification (SOC) at three digits, following Turrell et al. (2019). Their algorithm has an accuracy rate of 76 per cent. Python codes are available on GitHub. This algorithm was also implemented by Bellatin and Galassi (2022), who group occupations according to their involvement in the production and use of digital technologies.

By identifying occupations SOC categories, it is possible to provide a granular analysis, identifying the change in personality trait demand within similar occupations. The algorithm requires as inputs a job title, a text description, and a sector identification.

Table B.5: Cleaning of the database

Process	N vacancies	N no missing W*	Mean W*	SD W*	Δ
Adzuna raw database	276,325,386	178,229,590	31,120	18,524	-
Without recruitment agencies	150,918,209	83,727,680	29,882	18,558	-45%
Without missing companies match	66,833,773	32,647,950	28,365	18,392	-56%
Keep only Private companies with SIC	64,642,205	31,619,913	28,303	18,468	-3%
Remove duplicates vacancy-id, source-id, same description-location	20,783,111	10,490,006	28,342	18,830	-68%
Remove salaries missing	9,657,106	9,657,106	29,454	18,964	-54%
Incorporate new salaries from cleaned description	12,103,405	11,866,183	31,930	18,873	25%
Remove outliers in wages and text size	11,659,923	11,430,457	32,018	18,944	-4%

Notes: Author's creation based on Adzuna. W∗ = Salary considered after trimming 1% tails.

For that, being able to match Adzuna's company's name with Companies House was fundamental.

Classification Algorithm of Turrell et al. (2019)

Authors' contribution uses the job title, description and sector for job postings from the recruitment website Reed.com. In addition, they build dictionaries using the information published by the Office for National Statistics (ONS), for the four-digit UK SOC job titles and job descriptions. Inspired in Bellatin and Galassi (2022) description, the algorithm follows these steps:

1. Cleans the job posting text data using standard text analysis tools.
2. Creates a vector space for the dictionary of UK SOC text data provided by the ONS, using the term frequency-inverse document frequency (IF-IDF).
3. Searches for exact matches between the job titles in the job postings and those in the UK SOC list.
4. Combines the job posting title, the description and the sector into one string and expresses this string as a vector in the IF-IDF matrix of the UK SOC data.
5. For jobs that were not matched, it calculates the cosine similarity between the string obtained in the previous step, and the information in the UK SOC dictionary, and selects the five categories in the UK SOC list with the highest cosine similarity.
6. If the job posting title is empty, it returns the job posting with the highest cosine similarity. Instead, if there is text in the job posting title, the fuzzy-wuzzy Python

package is used to identify the best fuzzy match out of the top five ONS categories, following Levenshtein distance calculations.

Creation of new variables for the analysis

I built dictionaries of job requirements that were later searched on job advertisements. The thesaurus of both personality traits and requirements was expanded with germane variations, including words containing spelling variations and errors. To identify potential false positives, I did manual checks in 8 rounds of 100 advertisements, for each dictionary. By looking at the 10 closed words in the proximity of the text retrieved, I ensure to be grasping sentences with the correct meaning. When the searched word retrieved a different connotation, I modified the false-positive retrieval. For instance, *flexible hours* were replace by *flx hours* to avoid it being retrieved under the *flexible* matching. Terms with multiple words, such as *machine learning*, were joint (*machinelearning*), and the code only matched these when surrounded by blank spaces, commas or stops. In total, 1702 word replacements were done.

For the Personality Traits dictionary, I followed the *Knowledge engineering* approach, where a set of explicit rules to classify text are manually defined. A job posting is identified as requiring personality traits as long as it lists at least one keyword in its text description. The strategy of identifying a requirement by pinpointing its appearance in the text description and job titles has been used by Deming and Kahn (2018), Acemoglu et al. (2022), Alekseeva et al. (2021), among others. To construct the traits dictionary, adjectives belonging to the Big Five Personality schemes were used. In comparison with previous papers using rather arbitrary *character-related* language, the dictionary was created strictly following trait-related words used in the seminal psychology and economics papers Goldberg (1990), Costa Jr and McCrae (1992), Goldberg (1992), McCrae and John (1992), John (1990), Saucier (1994) and Heckman and Kautz (2012).

For Education, I use ONS education descriptions complemented with UK qualification levels[22]. I built 8 categories, later reduced to 5 following Nandi and Nicoletti (2014). The variable was constructed to consider the lowest educational level that was made explicit in the job search. Hence, if the advertisement is looking for someone with 'a bachelor's degree in Social Science, but a master's is preferable', the observation is marked as requiring at least a bachelor's level. Interestingly, the great majority of the vacancies did not mention an educational requirement: only 12 per cent of vacancies mentioned an education requirement.

For remote work, I inspired in Draca et al. (2022) dictionary of words. According to the authors, the pandemic of Covid-19 accelerated the adoption of remote work practices.

[22](https://www.gov.uk/what-different-qualification-levels-mean/list-of-qualification-levels).

The appearance of remote-work-related words in job vacancies in the UK increased by 300 per cent since 2019. Figure B.1 shows the weekly share of occupations that mention Remote Work in Adzuna, before and After the Covid-19 outburst.

Figure B.1: Weekly share of occupations that mention Remote Work



Notes: Author's creation based on Adzuna.

For Seniority, I followed Nandi and Nicoletti (2014) Nyhus and Pons (2005), who claim that the relevance of a personality trait may differ according to the position. The structure of hierarchy in some occupations may be more sensitive to personality characteristics. For the same SOC code, individuals applying for senior or managerial positions, *ceteris paribus* the experience and education, have different traits than for junior placements. I follow hierarchy descriptions available in Blogs, such as Indeed.

Adzuna's raw database incorporated information from the advertisements that were not necessarily available in the text entry. This was the case for Part and Full-time indicators, and Temporary and Permanent schedules. I created dictionaries for the four categories and complemented Adzuna's variables whenever they were missing.

Finally, I incorporated a dictionary on Experience and Bonuses. For the first one, I identify whether advertisements require any type of experience or explicitly say that this is not required. For the latter, I explore different variations in the text that signal applicants that the position could have a bonus or pay-for-performance reward.

Dictionaries

The skills used by Acemoglu et al. (2022) are *machine learning, computer vision, machine vision, deep learning, virtual agents, image recognition, natural language processing, speech recognition, pattern recognition, object recognition, neural networks, ai chatbot, supervised learning, text mining, unsupervised learning, image processing, mahout, recommender*

systems, support vector machines, random forests, latent semantic analysis (lsa), sentiment analysis, opinion mining, latent dirichlet allocation (lda), predictive models, kernel methods, keras, gradient boosting, openCV, xgboost, libsvm, word2vec, machine translation, and sentiment classification.

The skills used by Babina et al. (2024) that were not in Acemoglu et al. (2022) that were incorporated are *apache uima, artificial intelligence, autoencoders, bayesian networks/methods, caffe deep learning framework, classification algorithms, clustering algorithms, computational linguistics, convolutional neural network (cnn), dbscan (density-based spatial clustering), deeplearning4j, dimensionality reduction, dlib, expectation-maximization algorithm, hidden markov model, information extraction, jung framework, k-means, long short-term memory (lstm), matrix factorization, maximum entropy classifier, microsoft cognitive toolkit, mlpack, mxnet, naive bayes, natural language toolkit (nltk), nd4j, opennlp, pybrain, recurrent neural network (rnn), scikit-learn, semi-supervised learning, stochastic gradient descent (sgd), theano, unstructured information management architecture, vowpal, wabbit, weka.*

I expanded the list with the following terms: *a b testing, ai, bidirectional encoder (bert), chaid, cloud technology, cntk, cnns, convolutional, convolutional networks, cross domain learning, data augmentation, decision trees, deep embedding, discriminant analysis, dnn techniques, document embeddings, emotion detection, ensemble learning, ensemble methods, entity linking, fastai, feature engineering, federated learning, ffmpeg, gaussian processes, generative adversarial networks (gans), generative models, generative pre-trained transformer (gpt), hierarchical clustering, huggingface, hyperparameter tuning, image fusion, image segmentation, intelligent automation, k-nearest method, kalman filtering, knowledge distillation, kubernetes, mlflow, model compression, multi-task learning, named entity recognition, natural language generation (nlg), natural language understanding (nlu), object tracking, optical character recognition, pattern detection, pcl, predictive churn, predictive modelling, pytorch, quantum computing, recurrent network, reinforcement learning, resnets vaes, robotic process automation, rnns, sequence-to-sequence modelling, shot learning, siamese networks, similarity models, sound recognition, text classification, topic detection, topic modelling, tree based lerners, tree-based classifiers, t-sne, umap, uniform manifold, vision processing.*

Category	**Words searched**
Traits	
Agreeableness	affection, agreeableness, altruist, amiability, be gentle, considerate, co-operative, courteous, empathetic, forgiving, good nature, helpful, kind, naturalness, pleasant personality, sensitive, soft hearted, straightforward, sympathetic, trust, unselfish, warm

Continued on next page

Table B.6 – continued from previous page

Category	Words searched
Extraversion	adventurous, assertive, bold, candour, confident, courage, daring, energetic, enthusiastic, expressive, extraverted, forceful, friendly, funny, gregarious, humour, optimist, outgoing, outspoke, passion, playful, proactive, sociable, spontaneous, talkative
Conscientiousness	ambitious, be careful, competent, conscientious, decisive, dependable, dutiful, efficient, hard work, be logic, be neat, organised, painstaking, perseverance, persistent, planful, precise, prompt, punctual, reliable, responsible, striving, systematic, be thorough
Emotional Stability	at ease, calm, emotional, hostil, not fearful, tempered, fortitude, resilience, tense situations, unflappable, patience, pressure
Openness	artistic, be bright, be clever, creative, curious, flexible, ideas, imaginative, independent mind, ingenious, innovative, insightful, intellectual, intelligent, introspective, inventive, openness, be original, prefer variety, reflective, be sharp, unconventional, wide interests, witted
Skills	
Soft Skills	analytical, communication skills, critical thinking, cultural sensitivity, decision making, language skills, leadership, listening skills, mentoring, negotiation, networking abilities, presentation, problem solving, interpersonal, teamwork, time management, attention to detail, organisation skills, multitask, conflict resolution, soft skills, planning, questioning
Hard Skills	accounting, administrative, content creation, copywriting, data analysis, design skills, editing, engineering, legal analysis, marketing, microsoft office, project management, quality assurance, sales skills, software, statistics, user experience, video production, customer service, business development, computer literacy, commercial skills, budget, research, stakeholder

Continued on next page

Table B.6 – continued from previous page

Category	Words searched
AI Skills	machine learning, computer vision, machine vision, deep learning, virtual agents, image recognition, speech recognition, pattern recognition, object recognition, neural networks, supervised learning, text mining, support vector machines, unsupervised learning, image processing, image fusion, mahout, recommender systems, random forests, latent semantic analysis, sentiment analysis, opinion mining, latent dirichlet allocation, predictive models, kernel methods, keras, gradient boosting, gradient descent, opencv, xgboost, libsvm, word2vec, ai chatbot, machine translation, sentiment classification, generative adversarial networks, reinforcement learning, quantum computing, federated learning, natural language understanding, natural language processing, natural language generation, mlops, artificial intelligence, cloud technology, tensorflow, pytorch, scikit learn, natural language toolkit, bidirectional encoder, shot learning, ensemble methods, decision trees, k means clustering, autoencoders, robotic process automation, cnns, rnns, optical character recognition, named entity recognition, topic modelling, ab testing, dimensionality reduction, feature engineering, hyperparameter tuning, data augmentation, model compression, knowledge distillation, kalman filtering, kubernetes, azure ml, weka, clustering algorithms, classification algorithms, density based spatial clustering, caffe deep learning framework, generative pretrained transformer, expectation maximization, jung framework, theano, hugging face, ffmpeg, dlib, mxnet, vowpal wabbit, mlflow, bayesian methods, kubeflow, spark ml, discriminant analysis, long short term memory, hidden markov model, computational linguistics
Work Location	
Remote Work	conducted remote, fully remote, home based, homework, homeworking, job is remote, job is virtual, location is remote, location remote, position is remote, position remote, remote work, role will be remote, telecommuting, telework, virtual job, work from home, work home, work remote, working remote, flexible with office, remote position, role is remote, remote basis, hybrid work, hybrid remote
Seniority	
Level 1	chief, director, executive, vice president
Level 2 high	associate, coordinator, project lead, senior, supervisor, team lead, supervising a team, manager
Level 2 low	administrator, agent role, analyst, consultant, generalist, representative role, technician, secretary
Level 3	apprentice, assistant, clerk, internship, junior, receptionist, trainee, trial period, entry level
Education	

Continued on next page

Table B.6 – continued from previous page

Category	Words searched
No Formal Education	basic functional skills, basic skills qualification, entry level certificate, entry level functional skills, entry level qualifications, entry nqf level, essential skills wales, level 1 scqf, level 1 svq, level 2 scqf, level 2 svq, level 3 scqf, level 3 svq, no education requirements, no esol, no formal qualification, no previous experience or qualification, no previous qualification, no qualification, scottish national 1, scottish national 2, scottish national 3, scqf level 2, scqf level 3, skills for life, svq level 2, svq level 3, no formal education
GCSE Level	gcse, gnvq, btec award, btec first, btec level 1, certificate of secondary education, first certificate, foundation welsh baccalaureate, key skills level 1, level 1 award, level 1 certificate, level 1 diploma, level 1 esol, level 1 essential skills, level 1 functional skills, level 1 key skills, level 1 national vocational qualification, level 1 nvq, nqf level 1, nvq level 1, rqf level 1, scqf level 4, diploma level 1, esol 1, national level 1, qualification level 1, level 1 qualification, qcf level 1, level 1 qcf, functional skills level 1, essential skills level 1, level 1 apprenticeship, cipd level 1, level 1 cipd, level 4 scqf, level 4 svq, svq level 4
O-Level	o level, btec level 2, general certificate of education, intermediate apprenticeship, intermediate welsh baccalaureate, key skills level 2, level 2 award, level 2 certificate, level 2 diploma, level 2 esol, level 2 essential skill, level 2 functional skill, level 2 key skills, level 2 national, level 2 nvq, national level 2, nqf level 2, nvq level 2, rqf level 2, scqf level 5, secondary education, diploma level 2, award level 3, esol 2, qcf level 2, level 2 qcf, esol level 2, functional skills level 2, essential skills level 2, qualification level 2, level 2 qualification, level 2 apprenticeship, cipd level 2, level 2 cipd, level 5 scqf, level 5 svq, svq level 5, certificate of secondary education
A-Level	a level qualification, as level qualification, access to higher education diploma, advanced apprenticeship, advanced diploma, advanced extension, advanced welsh baccalaureate, baccalaureate, btec award level 3, btec certificate level 3, btec diploma level 3, btec level 3, btec national, diploma programme, extended certificate, extended diploma, key skills level 3, level 3 award, level 3 certificate, level 3 esol, level 3 national, level 3 nvq, national level 3, nqf level 3, nvq level 3, ocr national, progression diploma, rqf level 3, scottish advanced higher, scottish baccalaureate, scottish higher, scqf level 6, diploma level 3, award level 3, esol 3, nvqs level 3, qcf level 3, level 3 qcf, functional skills level 3, essential skills level 3, qualification level 3, level 3 qualification, level 3 apprenticeship, cipd level 3, level 3 cipd, level 6 scqf, level 6 svq, svq level 6

Continued on next page

Table B.6 – continued from previous page

Category	Words searched
Higher Certificate	hnc, certhe, certificate of higher education, higher apprenticeship, higher national certificate, level 4 award, level 4 certificate, level 4 diploma, level 4 national, level 4 nvq, national level 4, nqf level 4, nvq level 4, scqf level 7, diploma level 4, award level 4, nvqs level 4, qcf level 4, level 4 qcf, functional skills level 4, essential skills level 4, qualification level 4, level 4 qualification, level 4 apprenticeship, cipd level 4, level 4 cipd, level 7 scqf, level 7 svq, svq level 7
Higher Diploma	hnd, btec professional, diphe, diploma of further education, diploma of higher education, foundation degree, higher diploma, higher national diploma, level 5 award, level 5 certificate, level 5 diploma, level 5 national, level 5 nvq, national level 5, nqf level 5, nvq level 5, rqf level 4, scqf level 8, diploma level 5, award level 5, nvqs level 5, qcf level 5, level 5 qcf, qualification level 5, level 5 qualification, cipd level 5, level 5 cipd, level 8 scqf, level 5 apprenticeship, level 5 professional, level 8 svq, svq level 8
Bachelor's Degree	bachelor, bachelors degree, btec advanced professional, degree apprenticeship, first degree, grad cert, grad dip, bachelor certificate, bachelor diploma, level 6 award, level 6 certificate, level 6 diploma, level 6 national, level 6 nvq, national level 6, nqf level 6, nvq level 6, ordinary degree, oxbridge ma, rqf level 5, scottish ma, scqf level 10, scqf level 9, ugad, undergraduate, higher education, diploma level 6, award level 6, nvqs level 6, qcf level 6, level 6 qcf, qualification level 6, level 6 qualification, cipd level 6, level 6 cipd, level 9 scqf, level 10 scqf, level 6 apprenticeship, level 6 professional, level 10 svq, svq level 10, have a degree, bsc or masters, degree or master, bsc or msc, ba or ma, honours degree
Master's Degree	award level 7, ba and ma, cipd level 7, diploma level 7, have a master, have a masters, level 11 scqf, level 11 svq, level 7 apprenticeship, level 7 award, level 7 certificate, level 7 cipd, level 7 diploma, level 7 national, level 7 nvq, level 7 professional, level 7 qcf, master of arts, master of business administration, master of engineering, master of law, master of mathematics, master of pharmacy, master of philosophy, master of physics, master of professional studies, master of public health, master of research, master of science, master of social work, master of studies, masters, masters by research, masters degree, national level 7, nqf level 7, nvq level 7, nvqs level 7, postgrad, postgraduate certificate, postgraduate certificate in education, postgraduate diploma, professional masters, qcf level 7, rqf level 6, rqf level 7, scqf level 11, svq level 11
Doctorate	doctor of philosophy, doctoral degree, doctorate, doctors degree, level 8 award, level 8 certificate, level 8 diploma, level 8 national, level 8 nvq, national level 8, nqf level 8, nvq level 8, phd, rqf level 8, scqf level 12, diploma level 8, award level 8, nvqs level 8, qcf level 8, level 8 qcf, cipd level 8, level 8 cipd, level 12 scqf, level 8 apprenticeship, level 8 professional, svq level 12, level 12 svq
Schedule	

Continued on next page

Table B.6 – continued from previous page

Category	Words searched
Full-time	9 to 5, day shift, full time, standard hours, night shift
Part-time	4 hour, 6 hour, part time, 20 hour, 30 hour
Duration	
Temporary	apprenticeship, fixed term, freelance, interim, internship, project base, seasonal contract, self employed, short term contract, temporary, day to day contract, zero hours, casual contract
Permanent	long term, ongoing contract, permanent
Experience	
Required	3 to 5 years, experienced, can demonstrate, demonstrable, essential experience, experience a minimum, experience for, experience in, experience is also, experience needed, experience of, experience preferably, experience required, experience working, extensive experience, for experienced, has experience, have experience, have previous experience, have some knowledge, have worked, minimum experience, months experience, need experience, past experience, practical experience, previous, prior experience, proven, recent experience, relevant experience, right experience, solid experience, strong experience, suitable experience, three to five years, two to three years, with experience, work experience, worked in similar, years experience, years of experience, you got experience, required experience, experience essential, experience is preferred
Not Required	a novice, do not need any, do not need experience, don't need experience, don't need previous, experience is not, experience not, lots of experience, need any experience, no essential experience, no experience, no formal experience, no formal qualifications or experience, no past experience, no previous, no prior, no work experience, not essential, not necessary, not need previous, with no experience, without any experience, don't need experience, not needed
Other	
Bonus	bonus, commission, incentive, profit sharing, reward scheme, stock options, variable pay, performance related, rewards packages, pay for performance
Teaching	can teach, teaching
Negotiable	on experience, negotiable salary, salary negotiable, hour negotiable, salary is negotiable, salary up to
Communication	working directly with, reporting, report into, your team leader, working with, working closely, under the direction, speaking to, directly to, work alongside, required by, keeping your, work closely, reports, step to, shadowing, managing the, assistance to the, work with, managed by, approval of, communicate with, supporting

Dictionaries built following (1) Goldberg (1990); (2) Costa Jr and McCrae (1992); (3) Goldberg (1992); (4) McCrae and John (1992); (5) John (1990); (6) Saucier (1994); (7) Heckman and Kautz (2012) were scrutinised as not to retrieve false positive words. Each word's context was analysed by sampling the dataset in 8 rounds of 100 advertisements per

word. When the word's context was not related to personality traits and a straightforward amendment of the word was not possible, this was removed from the dictionary. Words that did not retrieve any match were also not considered.

Before searching for words, modifications were made to some words to avoid false positives and account for different variations of words. The list of changes done to the text is the following:

Personality Traits language:

"altruistic" for "altruist ", "altruism" for "altruist ", "amiable" for "amiability ", "empathy" for "empathetic ", "empathic" for "empathetic ", "extraversion" for "extraverted ", "extrovert" for "extraverted", "extrovertion" for "extraverted", "emotion" for "emotional", "emotionally" for "emotional", "reslient" for "recilience, "extroversion" for "extraverted ", "extroverted" for "extraverted ", "optimism" for "optimist ", "optimistic" for "optimist ", "outspoken" for "outspoke ", "spontaneity" for "spontaneous ", "ambition" for "ambitious", "competence" for "competent", "efficiency" for "efficient", "persevere" for "perseverance", "persevering" for "perseverance", "persistence" for "persistent", "hostile" for "hostil ", "hostility" for "hostil ", "creativity" for "creative", "creatively" for "creative", "innovation" for "innovative", "innovations" for "innovative", "innovating" for "innovative", "innovator" for "innovative", "intellect" for "intellectual", "intellectually" for "intellectual", "unconvention" for "unconventional", "pro active" for "proactive", "compassion" for "comps", "compassionate" for "comps", "an executive" for "an exec", "analytics" for "analysis", "and executive" for "and exec", "sense of adventure" for "adventurous", "humor" for "humour", "anger" for "angry", "anxiety" for "anxious", "agreeable personality" for "agreeableness", "being on time" for "punctual", "co operative" for "cooperative", "comfortable using" for "confoble using", "appreciative of" for "appreciativeof", "at ease" for "atease", "curios" for "curious", "owned cooperative" for "owned cop", "your energy" for "energetic", "with energy" for "energetic", "and energy" for "energetic", "encourages" for "encorages", "encourage" for "encorage", "encouraged" for "encoraged", "encouragement" for "encoragement", "discourage" for "discorage", "encouraging" for "encoraging", "confidential" for "confidntial", "self confidence" for "self confident", "confident employer" for "employer", "confident in" for "good in", "confidentiality" for "confid", "confidentially" for "confid", "confident at" for "good at", "confidently" for "good at", "originally posted" for "posted", "original proof" for "proof", "daring and apply" for "apply", "darington" for "daryngton", "daring to" for "daryng to", "our daring" for "our", "calendaring" for "calendar", "expressive arts" for "arts", "boldon" for "bolon", "nwebold" for "newbld", "not fearful" for "notfearful", "never fearful" for "notfearful", "be fearful" for "notfearful", "being fearful" for "notfearful", "careful to select" for "to select", " careful " for " becareful ", "spontaneous holiday" for "holiday", "to spontaneous" for "to fast", "from spontaneous" for "from fast", "spontaneous monthly" for "monthly", "spontaneous moments" for "moments", "various spontaneous" for "various", "spontaneous events" for "events", "are clever" for "beclever", "be clever" for "beclever", ", clever" for " beclever", "clever," for "beclever ", "with clever" for "beclever","some clever" for "beclever", "clever and" for "beclever", "and clever" for "beclever", "flexible shift" for "flx shift", "flexible shifts" for "flx shifts", "flexible working" for "flx work", "flexible work" for "flx work", "flexible with" for "flx with", "flexible basis" for "flx basis", "flexible hours" for "flx hours", "flexible benefit" for "flx benefit", "flexible benefits" for "flx benefits", "flexible about" for "flx about", "hours are flexible" for "hours are flx", "flexible annual" for "flx annual", "flexible to" for "flx to", "flexible for" for "flx for", "is flexible" for "is flx", "flexible hour" for "flx hour", "fully flexible" for "fully flx", "flexible schedule" for "flx schedule", "flexible salary" for "flx salary", "to flexible" for "tp flx", "flexible new" for "flx new", "flexible holiday" for "flx holiday", "flexible way" for "flx way", "flexible ways" for "flx ways", "pay flexible" for "flx pay", "work flexible" for "flx work", "flexible contract" for "flx contract", "flexible opportunities" for "flx opportunities", "flexible options" for "flx

options", "flexible subcontract" for "flx subcontract", "be flexible as" for "flx as", "flexible between" for "flx between", "flexible start" for "flx start", "flexible rota" for "flx rota", "weeks are flexible" for "flx week", "flexible days" for "flx days", "choose flexible" for "choose flx", "entirely flexible" for "entirely flx", "flexible within" for "flx within", "for flexible" for "for flx", "flexible additional" for "flx additional", "flexible fixed" for "flx fixed", "times flexible" for "times flx", "location flexible" for "location flx", "flexible on" for "flx on", "flexible schedules" for "flx schedules", "flexible reward" for "flx reward", "flexible rewards" for "flx rewards", "flexible 40" for "flx 40", "flexible weekend" for "flx weekend", "flexible vacancies" for "flx vacancies", "scheme flexible" for "scheme flx", "flexible in working" for "flx in working", "flexible menu" for "flx menu", "time flexible" for "time flx", "flexible pm" for "flx pm", "flexible saturdays" for "flx saturdays", "flexible home" for "flx home", "flexible team" : "flx team", "cleaner flexible" : "cleaner flx", "flexible and optional" : "flx and optional", "flexible self employed" : "flx self employed", "flexible in relation to" : "flx in relation to","flexible and able to" : "flx and able to", "flexible, full" : "flx, full", "flexible full" : "flx full", "flexible pattern" : "flx pattern", "hours of work can be flexible" : "flx hours", "flexible approach to remote working" : "flx remote working", "flexible work" for "flx work", "flexible and guaranteed hours" : "flx hours", "flexible networks" : "flx networks", "flexible job" : "flx job", "benefits flexible" : "flx benefits", "weekdays flexible" : "flx weekdays", "flexible guaranteed hours" : "flx hours", "freelance and flexible" : "freelance and flx", "flexible with office" : "flx with office", "options flexible" : "options flx", "flexible core" : "flx core", "flexiblejob" : "flx job", "workflexible" : "work flx", "flexbile interview" : "flx interview", "flexible temporary" : "flx temporary", "flexible, part" : "flx, part", "flexible part" : "flx part", "flexible and home" : "flx and home", "days are flexible" : "days are flx", "flexible employment" : "flx employment", "that flexible" : "that flx", "flexible location" : "flx location", "flexible rotas" : "flx rotas", "flexible and remote" : "flx and remote", "flexible pension" : "flx pension", "esteemed" for "estemed", "be bright " for "bebright ", " bright and" for " bebright", "and bright" for "bebright", ", bright " for " bebright ", " bright, " for " bebright ", " a bright " for " bebright ", "brightest" for "bebright", "bright welcoming" for "welcoming", "bright future" for "future", "bright sunny" for "sunny", "be helpful" for "be help", "is helpful" for "is help", "experience helpful" for "experience help", "independent minded" for "independentmind", "independent mind" for "independentmind", "prefer variety" for "prefervariety", "straight forward" for "straightforward", "straightforward for you" for "forward for you", "straight forward for you" for "forward for you", "straightforward questionnaire" for "questionnaire", "very straightforward" for "very", "provides straight forward" for "provides", "straightforward design" for "design", "is straight forward" for "is", "friendly policies" for "frend policies", "family friendly" for "family frend", "friendliness" for "friendly", "user friendly" for "user friend", "friendly local" for "local", "dogfriendly" for "dog", "ecofriendly" for "eco", " gentle " for " begentle ", "solver" for "solving", "wide interest" for "wideinterests", "wider interest" for "wideinterests", "otherwise" for "other", "likewise" for "like", "connectwise" for "connect", "deliberateness" for "deliberate", "beneath" for "beneth", " neat " for " beneat ", " be neat " for " beneat ", "helpful tool" for "tool", "effective communication" for "communication skill", "optimisation" for "optimization", "optimisations" for "optimizations", "optimising" for "optimizing", "optimise" for "optimize", "optimised" for "optimized", "optimises" for "optimizes", "optimiser" for "optimizer", "empathy" for "empathetic", "enthusiasm" for "enthusiastic", "open, " for "openness ", "open and" for "openness ","and open " for "openness ","openmind" for "openness", "open mind" for "openness ","open minded" for "openness ", "openness close " for "close", "open and close" for "close", "open source" for "source", "openness source" for "source", "enthusiast" for "enthusiastic", "fun " for "fun ", "fun day" for "fus day", "fun park" for "fus park", "fun " for "funny ", "fun," for "funny ", "fun." for "funny ", "grounded in" for "in", "be grounded in " for "in", "grounded established" for "established", "a grounded" for " a ","grounded knowledge" for "knowledge", "grounded programming" for "programming","grounded perspective" for "perspective","grounded reflected" for "reflected","grounded solutions" for "solutions", "grounded and" for "begrounded", "and grounded" for "begrounded", ", grounded" for "

begrounded", "grounded," for "begrounded ", "be grounded" for "begrounded", "grounded analysis" for "analysis", "well grounded" for "well", " a grounded," for " a ", ", grounded in" for " in", "begrounded in" for " in", ", grounded in " for "in ", "hard working" for "hardwork", "sociable hours" for "hours", "sociable working hours" for "working hours", "hard work" for "hardwork", "hardworker" for "hardwork", "hardworking" for "hardwork", "logic apps" for "apps", " logic " for " belogic ","‌ logical " for " belogic ", "thorough knowledge" for "knowledge", "bethorough knowledge" for "knowledge", "and thorough" for "bethorough", ",thorough" for "bethorough", "thorough and" for "bethorough","be thorough" for "bethorough", "home working" for "homeworking", "sensitive information" for "information", "sensitive or contentious information" for "contentious information", "sensitive and restricted information" for "restricted information", "sensitive data" for "data", "sensitive company" for "company", "intellectual property" for "property", "intellectual dis" for "dis", "track record" for "experienced", "background" for "experienced", "outgoing mail" for "mail", "outgoing call" for "call", "pleasant attitude" for "pleasantpersonality", "pleasant personality" for "pleasantpersonality", "pleasant manner" for "pleasantpersonality", "be pleasant" for "pleasantpersonality", "pleasant and" for "pleasantpersonality", "kind of" for "jind of", "kinds of" for "jinds of", "kindly" for "jindly", "kind to" for "jind to", "tender minded" for "tendermind", " tender mind" for "tendermind", "good nature" for "goodnature", "tense situations":"tensesituations", "under control":"undercontrol", "competencies" for "skills", "competence in" for "skilled in", "competence with" for "skilled with", "soft heart" for "warm heart", "soft hearted" for "warm hearted", "competency" for "skilled", "competently" for "skilled", "organized" for "organised","show consideration" for "considerate", "temperance" for "tempered", "temperament" for "tempered", "original solutions" for "beoriginal", "original thinking" for "beoriginal", "original ideas" for "beoriginal", "original research" for "beoriginal", "be original" for "beoriginal ", "original, " for "beoriginal", "original writing" for "beoriginal", "original content" for "beoriginal", "is original" for "beoriginal", "think originally" for "beoriginal", ", original" for " beoriginal", "original thought" for "beoriginal", "original choice" for "beoriginal", "originality" for "beoriginal", "creating original" for "beoriginal", "national trust" for "nationaltr", "trusts" for "trus", "trustee" for "trusee","trustees" for "trusees", "cooperative trust" for "coperative", "nhs trust" for "nhs", "trust us" for "trus us", "trusted" for "trused", " our trust" for " our trus", "local trust" for "local", "trust are" for "are", "academy trust" for "academy", "trust schools" for "schools", "the trust" for "the trus", "trust is" for "is", "foundation trust" for "foundation", "at trust" for "at trus", "entrusted" for "enrusted", "wildlife trust" for "wildlife", "castle trust" for "castle", "trust scheme" for "scheme", "trust duties" for "duties", "education trust" for "education", "adventure trust" for "adventure", "disabilities trust" for "disabilities", "affinity trust" for "affinity", "academies trust" for "academies", "whole trust" for "whole", "trust offers" for "offers", "trust within" for "within", "trust and corporate" for "corporate", "trust duties" for "duties", "trust and grants" for "grants", "leisure trust" for "leisure", "we trust" for "we", "trust manages" for "manages", "enterprise trust" for "enterprise", "trust keith" for "keith", "trust flagship" for "flagship", "woodland trust" for "woodland", "trust and fundraising" for "fundraising", "care trust" for "care", "health trust" for "health", "rivers trust" for "rivers", "transform trust" for "transform", "transformation trust" for "transformation", "are a trust " for "are ", "trust qvd" for "qvd", "educational trust" for "educational", "trust exposure" for "exposure", "trust has" for "has", "dean trust" for "dean", "trust delivers" for "delivers", "as a trust " for " ", "london trust" for "london", "trustpilot" for " ", "trust pilot" for " ", "trust education" for "education", "trust we believe" for "we believe", "northern trust" for "northern", "trust website" for "website", "trust. org" for ".org", "trust . org" for ".org", "trust.org" for ".org", "trust org" for "org", "persist" for "persistent", "prompt response" for "prom response", "as prompt" for "as prom", "prompting" for "promting", "precisely what" for "what", "as prompt" for "as prom", "warm welcome" for "welcome", "warm you" for "you", "reliable vehicle" for "rel veh", "representative of" for "represent of", "responsible to" for "res to", "responsible of" for "res of", "responsible in" for "res in", "responsible for" for "res for", "reflective of" for "ref of", "a reflective" for "a ref", " sharp " for "

besharp ", "socially responsible" for "socially res", "unsociable" for "unsoci", "are striving" for "are str", "sympathetically converted" for "converted", "sympathetically restored" for "restored", "sympathetic refurbishment" for "refurbishment", "sympathetic donors" for "donors".

Soft skills language:

"communication skills" for "communicationskills", "critical thinking" for "criticalthinking", "cultural sensitivity" for "culturalsensitivity", "decision making" for "decisionmaking", "language skills" for "languageskills", "listening skills" for "listeningskills", "problem solving" for "problemsolving", "time management" for "timemanagement", "attention to detail" for "attentiontodetail", "organisation skills" for "organisationskills", "conflict resolution" for "conflictresolution", "soft skills" for "softskills","maintain relationships" for "networkingabilities", "network abilities" for "networkingabilities", "working relationships" for "networkingabilities", "network ability" for "networkingabilities", "relationship building" for "networkingabilities", "headhunting" for "networkingabilities", "networking events" for "networkingabilities", "maintain contacts" for "networkingabilities", "external relationships" for "networkingabilities", "strong relationships" for "networkingabilities", "personal network" for "networkingabilities", "ability to network" for "networkingabilities", "attend events" for "networkingabilities", "business relationships" for "networkingabilities", "positive relationships" for "networkingabilities", "good listening" for "listening skills", "good listener" for "listening skills", "active listening" for "listening skills", "ability to listen" for "listening skills", "resolve issues" for "problemsolving", "problem solver" for "problemsolving", "fluency" for "fluent", "representation" for "reprsentation", "representations" for "reprsentation", "making decisions" for "decisionmaking", "make decisions" for "decisionmaking", "best decisions" for "decisionmaking", "good decisions" for "decisionmaking", "take decisions" for "decisionmaking","informed decisions" for "decisionmaking", "difficult decisions" for "decisionmaking", "appropriate decisions" for "decisionmaking", "resolving the complex problems" for "problemsolving", "resolving complex problems" for "problemsolving", "negotiating" for "negotiation", "negotiator" for "negotiation", "good spoken" for "communicationskills", "transparent communication" for "communicationskills", "good communication" for "communicationskills", "strong communication" for "communicationskills", "excellent communication" for "communicationskills", "written and verbal" for "communicationskills", "written and oral" for "communicationskills", "oral and written" for "communicationskills", "communication abilities" for "communicationskills", "great communication" for "communicationskills", "oral" for "communicationskills", "verbal, written" for "communicationskills", "technical writting" for "communicationskills", "verbal and written" for "communicationskills", "written" for "communicationskills", "verbal" for "communicationskills", "oral" for "communicationskills", "ability to communicate" for "communicationskills", "strong communicator" for "communicationskills", "excellent communicator" for "communicationskills", "good communicator" for "communicationskills", "communicator" for "communicationskills", "able to communicate" for "communicationskills", "transparent communication" for "communicationskills", "great communication" for "communicationskills", "can lead" for "leadership", "people management" for "leadership", "able to lead" for "leadership", "team leadership" for "supervisingateam", "team leadership" for "leading a team", "team management" for "leading a team", "cultural awareness" for "culturalsensitivity", "crosscultural" for "culturalsensitivity", "multicultural" for "culturalsensitivity", "multi cultural" for "culturalsensitivity", "culturally diverse" for "culturalsensitivity", "culturally aware" for "culturalsensitivity", "cultural diversity" for "culturalsensitivity", "asking questions" for "questioning", "european language" for "languageskills", "fluent in chinese" for "languageskills", "fluent in german" for "languageskills", "fluent in french" for "languageskills", "fluent in english" for "languageskills", "fluent in mandarin" for "languageskills", "fluent in spanish" for "languageskills", "fluent in portuguese" for "languageskills","fluent in polish" for "languageskills", "fluent speaking" for "languageskills", "fluent in dutch" for "languageskills", "fluent in russian" for "languageskills", "two languages" for "languageskills", "fluent in both" for "languageskills", "fluent in hebrew" for "languageskills", "fluent in speaking" for "languageskills", "fluent in italian" for "languageskills", "fluent in welsh" for "languageskills",

“fluent in japanese” for “languageskills”, “fluent in danish” for “languageskills”, “fluent in english and” for “languageskills”, “communicating in two langauges” for “languageskills”, “fluent both english and” for “languageskills”, “fluent in english as well as” for “languageskills”, “fluent in another” for “languageskills”, “fluent in the german” for “languageskills”, “number of languages” for “languageskills”, “fluent in turkish” for “languageskills”, “coaching” for “mentoring”, “provide guidance” for “mentoring”, “multiple projects” for “multitask”, “multiple tasks” for “multitask”, “training and mentoring” for “training”, “mentoring and training” for “training”, “multi projects” for “multitask”, “multitasker” for “multitask”, “multitasking” for “multitask”, “ability to motivate” for “mentoring”,“able to motivate” for “mentoring”, “couching” for “mentoring”, “problem resolution” for “problemsolving”, “complex problems” for “problemsolving”, “other languages” for “languageskills”, “additional languages” for “languageskills”, “multilingual” for “languageskills”, “office skills” for “microsoftoffice”, “excel skills” for “microsoftoffice”, “ms packages” for “microsoftoffice”, “ms package” for “microsoftoffice”, “organisational skills” for “organisationskills”, “organizational skills” for “organisationskills”, “organisational capacity” for “organisationskills”, “organisational abilities” for “organisationskills”, “organisational ability” for “organisationskills”, “excellent organisation” for “organisationskills”, “organisation, coordination” for “organisationskills”, “organisational and managements skills” for “organisationskills”, “excellent organisational” for “organisationskills”, “good organisational” for “organisationskills”, “strong organisational” for “organisationskills”, “organisational, ” for “organisationskills ”, “organisational and ” for “organisationskills ”, “strong managements skills” for “organisationskills”, “excellent managements skills” for “organisationskills”, “competent managements skills” for “organisationskills”, “strong organisational” for “organisationskills”, “prioritisation” for “organisationskills”, “able to prioritise” for “organisationskills”, “excellent organisation” for “organisationskills”, “mentoring of the junior” for “mentoring”, “mentoring junior” for “mentoring”, “able to mentor the more junior members” for “mentoring”, “mentoring our junior” for “mentoring”, “mentoring our more junior” for “mentoring”, “able to mentor junior” for “mentoring”, “especially junior” for “mentoring”, “mentor junior” for “mentoring”, “mentoring and training of more junior” for “mentoring”, “mentoring to junior” for “mentoring”, “mentoring of junior” for “mentoring”, “mentoring and coaching junior” for “mentoring”, “mentor more junior” for “mentoring”, “training of junior” for “mentoring”, “eye for detail” for “attentiontodetail”, “detail oriented” for “attentiontodetail”, “detailed approach” for “attentiontodetail”, “mentoring more junior” for “mentoring”, “support junior” for “mentoring”, “managing junior” for “mentoring”, “developing junior” for “mentoring”, “developing more junior” for “mentoring”, “support to more junior” for “mentoring”, “support more junior” for “mentoring”, “development of junior” for “mentoring”, “work of junior” for “mentoring”, “coaching more junior” for “mentoring”, “coaching of junior” for “mentoring”, “training junior” for “mentoring”, “supervising junior” for “mentoring”, “work as a team” for “teamwork”, “work in a team” for “teamwork ”, “work in teams” for “teamwork”, “work well in a team” for “teamwork”, “within a team” for “teamwork”, “ intern ” for “ internship ”, “team work” for “teamwork”, “team working” for “teamwork”, “team environment” for “teamwork”, “able to relate” for “teamwork”, “team player” for “teamwork”, “collaborative work” for “teamwork“.

Hard skills language:

“content creation” for “contentcreation”, “design skills” for “designskills”, “data analysis” for “dataanalysis”, “legal analysis” for “legalanalysis”, “microsoft office” for “microsoftoffice”, “project management” for “projectmanagement”, “quality assurance” for “qualityassurance”, “sales skills” for “salesskills”, “user experience” for “userexperience”, “video production” for “videoproduction”, “customer service” for “customerservice”, “business development” for “businessdevelopment”, “computer literacy” for “computerliteracy”, “commercial skills” for “commercialskills”, “business opportunities” for “businessdevelopment”, “database management” for “dataanalysis”, “data analytics” for “dataanalysis”, “commercial abilities” for “commercialskills”, “commercially aware” for “commercialskills”, “commercially awareness” for “commercialskills”, “focus on commercial” for “commercialskills”, “commercially astute” for “commercialskills”,

“timekeeper” for “timemanagement”, “time keeping” for “timemanagement”, “commercially minded” for “commercialskills”, “commercially focused” for “commercialskills”, “commercially astute” for “commercialskills”, “commercially driven” for “commercialskills“ , “commercially focussed” for “commercialskills”, “word outlook” for “microsoftoffice”, “excel and outlook” for “microsoftoffice”, “microsoft outlook” for “microsoftoffice”, “prefficient in word” for “microsoftoffice”,“microsoft word” for “microsoftoffice”, “proficient in excel” for “microsoftoffice”, “microsoft excel” for “microsoftoffice”, “power point” for “microsoftoffice”, “powerpoint” for “microsoftoffice”, “competent in microsoft” for “microsoftoffice”, “microsoft packages” for “microsoftoffice”, “office 365” for “microsoftoffice”, “ms excel” for “microsoftoffice”, “command of microsoft” for “microsoftoffice”, “within budget” for “within budg”, “word excel” for “microsoftoffice”, “word and excel” for “microsoftoffice”, “excel and word” for “microsoftoffice”, “word, excel” for “microsoftoffice”, “excel, word” for “microsoftoffice”, “outlook and” for “microsoftoffice”, “and outlook” for “microsoftoffice”, “using outlook” for “microsoftoffice”, “ms office” for “microsoftoffice”, “microsoft teams” for “call”, “microsoft programmes” for “microsoftoffice”, “positive outlook” for “positive”, “working with team” for “teamwork”, “statistical” for “statistics”, “pc literate” for “computerliteracy”, “computer literate” for “computerliteracy”, “using a computer” for “computerliteracy”, “computer skills” for “computerliteracy”, “pc skills” for “computerliteracy”, “analytics” for “analysis”, “investigative skills” for “analytical skills”, “research” for “research skills”, “cancer research” for “cancer”,“cancer research” for “cancer”, “research organisations” for “organisations”, “researching” for “research skills”, “research organisation” for “organisations”, “research programmes” for “programmes”, “researcher” for “research”, “good numeracy” for “analytical skills”, “numeracy skills” for “analytical skills”, “numeracy litarate” for “analytical skills”, “conflict management” for “conflictresolution”, “collaborative working” for “teamwork”, “supervising a” for “supervisingateam”, “managing a team” for “supervisingateam”, “team management” for “supervisingateam”, “part of a team” for “teamwork”, “sunrise senior” for “sunrise”, “seniors” for “old”, “associated” for “asociated”, “associates” for “asociates”, “their representatives” for “their clients”, “client representatives” for “client”, “sales representative” for “representativerole salesskills”, “sales representatives” for “representativerole salesskills”, “customerservice” for “representativerole customerservice”, “customer” for “representativerole customer”, “marketing representative” for “marketing representativerole”, “representative role” for “representativerole”, “our consultants” for “our consult”, “a consultant” for “a consult”, “consultants available” for “consult available”, “our qualified consultants” for “our consult”, “agent role” for “agentrole”, “sales agent” for “agentrole”, “support agent” for “agentrole”, “customerservice agent” for “agentrole”, “customerservice agent” for “agentrole”, “service agent” for “agentrole”, “accountant” for “accounting”, “directory” for “directo”, “directorate” for “directo”, “sales experience” for “salesskills experienced”, “sales person” for “salesskills”, “experience in sales” for “salesskills experienced”, “identify key sales” for “salesskills”, “sales oriented” for “salesskills”, “sales culture” for “salesskills”, “sales representative” for “sales representative salesskills”, “of sales” for “salesskills”, “sales plan” for “salesskills”, “sales plans” for “salesskills”, “sales associate” for “salesskills”, “sales role” for “salesskills”, “sales team” for “salesskills team”, “sales manager” for “salesskills manager”, “sales job” for “salesskills”, “sales target” for “salesskills”, “direct sales” for “salesskills”, “success in sales” for “salesskills”, “success as sales” for “salesskills”, “develop sales” for “salesskills”, “sales strategies” for “salesskills”, “sales targets” for “salesskills”, “target driven sales” for “salesskills”, “digital design” for “designskills”, “design engineer” for “designskills”, “graphic design” for “designskills”, “motion design” for “designskills”, “design tools” for “designskills”, “photoshop” for “designskills”, “motion graphics” for “designskills”, “digital art” for “designskills”, “adobe” for “designskills”, “design management” for “designskills”, “technical designs” for “designskills”, “photo shop” for “designskills”, “creating content” for “contentcreation”, “generate content” for “contentcreation”, “content generation” for “contentcreation”, “management of content” for “contentcreation”, “illustrator” for “designskills”, “content strategy” for “contentcreation”, “content development” for “contentcreation”, “content management” for “contentcreation”, “content creators” for “contentcreation”, “social content”

for “contentcreation”, “managing content” for “contentcreation”, “website content” for “contentcreation”, “content writer” for “contentcreation”, “editorial management” for “editing”, “editorial skills” for “editing”, “editorial tasks” for “editing”, “editorial role” for “editing”, “editorial roles” for “editing”, “publishing skills” for “editing”, “editorial judgment” for “editing”, “editorial experience” for “editing experienced”, “editorial style” for “editing”, “editorial work” for “editing”, “proofreading” for “editing”, “proof reading” for “editing”, “drafting” for “editing”, “copy writing” for “editing”, “copywriting” for “editing”, “copywriter” for “editing”, “copywrite” for “editing”, “good editorial” for “editing”, “excellent editorial” for “editing”, “editorial assistant” for “editing assistant”, “editorial ability” for “editing”, “editorial judgment” for “editing”, “creating engaging content” for “contentcreation”, “develop content” for “contentcreation”, “developing content” for “contentcreation”, “writing, editorial” for “editing”, “legal text” for “legalanalysis”, “legal role” for “legalanalysis”, “legal abilities” for “legalanalysis”, “paralegal” for “legalanalysis”, “legal background” for “legalanalysis”, “legal qualifications” for “legalanalysis”, “legal qualification” for “legalanalysis”, “legal excecutive” for “legal excecutive legalanalysis”, “lawyer” for “lawyer legalanalysis”, “reading leases” for “legalanalysis”, “lawyers” for “lawyers legalanalysis”, “reading leases” for “legalanalysis”, “legal clerk” for “legalanalysis”, “provide legal” for “legalanalysis”, “legal billing” for “legalanalysis”, “ engineer ” for “ engineering ”, “legal documentation” for “legalanalysis”, “legal approach” for “legalanalysis”, “caseload” for “legalanalysis”, “understand our customers” for “customerservice”, “customer focused” for “customerservice”, “customer focus” for “customerservice”, “identify customer needs” for “customerservice”, “understanding the customers” for “customerservice”, “needs of the customers” for “customerservice”, “manage customer” for “customerservice”, “customer orders” for “customerservice”, “interact with customers” for “customerservice”, “quality checking” for “qualityassurance”, “quality checks” for “qualityassurance”, “quality check” for “qualityassurance”, “ensuring quality” for “qualityassurance”, “search engine” for “software”, “sem ” for “software”, “seo ” for “software ”, “web development” for “software”, “programming” for “software”, “mobile app development” for “software”, “coding” for “software”, “web design” for “software”, “cybersecurity” for “software”, “data mining” for “dataanalysis”, “data visualization” for “dataanalysis”, “database skills” for “dataanalysis”, “data to” for “dataanalysis”, “quantitative analysis” for “dataanalysis ”, “data modelling” for “dataanalysis”, “datasets” for “dataanalysis”, “interpret data” for “dataanalysis”, “data reporting” for “dataanalysis”, “it skills” for “dataanalysis”, “it literate” for “dataanalysis”, “it literacy” for “dataanalysis”, “variety of data” for “dataanalysis”, “data skills” for “dataanalysis”, “data models” for “dataanalysis”, “data sets” for “dataanalysis”, “data analyst” for “dataanalysis analyst”, “data monitoring” for “dataanalysis”, “data scientist” for “dataanalysis”, “data skill” for “dataanalysis”, “handling data” for “dataanalysis”, “complex data” for “dataanalysis”, “big data” for “dataanalysis”, “data science” for “dataanalysis”, “data engineer” for “dataanalysis”, “collate data” for “dataanalysis”, “data interpretation” for “dataanalysis”, “interpretation of data” for “dataanalysis”, “creation of data” for “dataanalysis”, “relational databases” for “dataanalysis”, “database management” for “dataanalysis”, “database utilisation” for “dataanalysis”, “data management” for “dataanalysis”, “database administration” for “dataanalysis”, “comparative data” for “dataanalysis”, “process data” for “dataanalysis”, “raw data” for “dataanalysis”, “data insight” for “dataanalysis”, “customer data” for “dataanalysis”, “customer databases” for “dataanalysis”, “company databases” for “dataanalysis”, “accurate databases” for “dataanalysis”, “accurate data” for “dataanalysis”, “data driven” for “dataanalysis”, “data entry” for “dataanalysis”, “data reviews” for “dataanalysis”, “review data” for “dataanalysis”, “project coordination” for “projectmanagement”, “project planning ” for “projectmanagement”, “contract drafting” for “projectmanagement”, “producing video” for “videoproduction”, “photography” for “videoproduction”, “video editing” for “videoproduction”, “video creation” for “videoproduction”, “videographer” for “videoproduction”, “film maker” for “videoproduction”, “creating videos” for “videoproduction”, “video experience” for “videoproduction experienced”, “filming” for “videoproduction”, “social interaction” for “interpersonal”, “social skills” for “interpersonal”, “people

skills” for “interpersonal”, “people person” for “interpersonal”, “clerical” for “administrative”,“ ux ” for “userexperience”, “ ux, ” for “userexperience”, “ ux.” for “userexperience”, “statistics” for “statistical”, “stake holder” for “stakeholder”, “client management” for “stakeholder management”, “supply chain management” for “product management“.

AI language:

“ ai ” for “ artificialintelligence ”, “ xai ” for “ artificialintelligence ”, “ ai,” for “ artificialintelligence ”, ”, ai ” for “ artificialintelligence ”, “artificial intelligence” for “ artificialintelligence ”, “ gpt ” for “generativepretrainedtransformer”, “generative models” for “generativepretrainedtransformer”, “generative pretrained” for “generativepretrainedtransformer”, “generative pre” for “generativepretrainedtransformer”, “ai chatbox” for “aichatbox”, “artificialintelligence chatbox” for “aichatbox”, “machine learning” for “machinelearning”, “ ml ” for “ machinelearning ”, ”, ml ” for “ machinelearning ”, “ ml,” for “ machinelearning ”, “cross domain learning” for “machinelearning”, “cross - domain learning” for “machinelearning”, “reinforcement learning” for “machinelearning”, “mlpack” for “machinelearning”, “federated learning” for “machinelearning”, “pybrain” for “machinelearning”, “shot learning” for “machinelearning”, “hugging face” for “huggingface”, “multi task learning” for “machinelearning”, “multi - task learning” for “machinelearning”, “residual learning” for “machinelearning”, “mediapipe” for “machinelearning”, “media pipe” for “machinelearning”, “computer vision” for “computervision”, “image segmentation” for “computervision”, “vision processing” for “computervision”, “robotic process automation” for “roboticprocessautomation”, “ rpa ” for “ roboticprocessautomation ”, “autonomous robotics” for “roboticprocessautomation”, “intelligent automation” for “roboticprocessautomation”, “machine vision” for “machinevision”, “ bert ” for “bidirectionalencoder”, “bidirectional encoder” for “bidirectionalencoder”, ”, bert” for “ bidirectionalencoder”, “ bert,” for “ bidirectionalencoder”, “deep learning” for “deeplearning”, “residual network” for “deeplearning”, “ resnet ” for “deeplearning”, “fastai” for “deeplearning”, “nd4j” for “deeplearning”, “microsoft cognitive toolkit” for “deeplearning”, “ cntk ” for “deeplearning”, “deep embeddings” for “deeplearning”, “deep embedding” for “deeplearning”, “kalman filtering” for “kalmanfiltering”, “kalman filter” for “kalmanfiltering”, “virtual agents” for “virtualagents”, “support vector machines” for “supportvectormachines”, “ svm ” for “ supportvectormachines ”, “ svm,” for “ supportvectormachines ”, ”, svm ” for “ supportvectormachines ”, “image recognition” for “imagerecognition”, “point cloud library” for “imagerecognition”, “image fusion” for “imagefusion”, “natural language toolkit“ : “naturallanguagetoolkit”, “nltk“ : “naturallanguagetoolkit”, “ maximum entropy ” for “ naturallanguageprocessing ”, “ nlp ” for “naturallanguageprocessing”, ”, nlp ” for “ naturallanguageprocessing ”, “ nlp, ” for “ naturallanguageprocessing ”, “ corenlp” for “ naturallanguageprocessing ”, “ opennlp” for “ naturallanguageprocessing ”,“ apache uima” for “ naturallanguageprocessing ”, “natural language processing” for “naturallanguageprocessing”, “ gensim ” for “ naturallanguageprocessing ”, “document embeddings” for “naturallanguageprocessing”, “doc2vec” for “naturallanguageprocessing”, “speech recognition” for “speechrecognition”, “sound recognition” for “speechrecognition”, “scikit learn” for “scikitlearn”, “scikit - learn” for “scikitlearn”, “pattern recognition” for “patternrecognition”, “pattern detection” for “patternrecognition”,“discriminant” for “discriminantanalysis”, “object recognition” for “objectrecognition”, “object tracking” for “objectrecognition”, “neural networks” for “neuralnetworks”, “convolutional” for “neuralnetworks”, “ lstm ” for “longshorttermmemory”, “ lstms ” for “longshorttermmemory”, “ , lstm ” for “longshorttermmemory”, “ , lstms ” for “longshorttermmemory”, “ lstm, ” for “longshorttermmemory”, “ lstms, ” for “longshorttermmemory”, “recurrent network” for “neuralnetworks”, “recurrent networks” for “neuralnetworks”, “siamese networks” for “neuralnetworks”, “ rnn ” for “ rnns ”, “dnn techniques” for “neuralnetworks”, “supervised learning” for “supervisedlearning”, “gaussian processes” for “supervisedlearning”, “text mining” for “textmining”, “information extraction” for “textmining”, “unstructured information management” for “textmining”, “support vector machines” for “supportvectormachines”, “unsupervised learning” for “unsupervisedlearning”, “image processing” for “imageprocessing”,“jung framework” for “jungframework”,

“recommender systems” for “recommendersystems”, “random forests” for “randomforests”, “random forest” for “randomforests”, “latent semantic analysis” for “latentsemanticanalysis”, “sentiment analysis” for “sentimentanalysis”, “emotion detection” for “sentimentanalysis”, “tensor flow” for “tensorflow”, “opinion mining” for “opinionmining”, “hidden markov model“:“hiddenmarkovmodel”, “latent dirichlet allocation” for “latentdirichletallocation”, “ lda ” for “ latentdirichletallocation ”, “ lda,” for “ latentdirichletallocation ”, ”, lda” for “ latentdirichletallocation ”, “predictive models” for “predictivemodels”, “predictive modelling” for “predictivemodels”, “kernel methods” for “kernelmethods”, “kernel method” for “kernelmethods”, “gradient descent” for “gradientdescent”, “ sgd ” for “gradientdescent”, “gradient boosting” for “gradientboosting”, “gradient boost” for “gradientboosting”, “gradient boosted” for “gradientboosting”, “machine translation” for “machinetranslation”, “seq2seq” for “machinetranslation”, “sequence - to - sequence” for “machinetranslation”, “sequence to sequence” for “machinetranslation”, “sentiment classification” for “classificationalgorithms”, “classification algorithms” for “classificationalgorithms”, “classification algorithm” for “classificationalgorithms”, “ gans ” for “generativeadversarialnetworks”, ”, gans ” for “generativeadversarialnetworks”, “ gans, ” for “generativeadversarialnetworks”, “ generative adversarial networks” for “generativeadversarialnetworks”, “quantum computing” for “quantumcomputing”, “ nlu ” for “naturallanguageunderstanding”, ”, nlu ” for “ naturallanguageunderstanding ”, “ nlu, ” for “ naturallanguageunderstanding ”, “natural language understanding” for “naturallanguageunderstanding”, “ nlg ” for “naturallanguagegeneration”, ”, nlg ” for “ naturallanguagegeneration ”, “ nlg, ” for “ naturallanguagegeneration ”, “ natural language generation ” for “naturallanguagegeneration”, “nlg health ” for “health”, “nlg are” for “are”, “cloud technology” for “cloudtechnology”, “cloud technologies” for “cloudtechnology”, “ensemble methods” for “ensemblemethods”, “ensemble learning” for “ensemblemethods”, “decision trees” for “decisiontrees”, “ chaid” for “decisiontrees”, “tree based lerners” for “decisiontrees”, “tree - based lerners” for “decisiontrees”, “tree - based classifiers” for “decisiontrees”, “tree based classifiers” for “decisiontrees”, “k means clustering” for “clusteringalgorithms”, “hierarchical clustering” for “clusteringalgorithms”, “k - means clustering” for “clusteringalgorithms”, “clustering algorithms” for “clusteringalgorithms”, “optical character recognition” for “opticalcharacterrecognition”, “optical flow” for “opticalcharacterrecognition”, “named entity recognition” for “namedentityrecognition”, “entity linking” for “ naturallanguageprocessing ”, “topic modeling” for “topicmodelling”, “topic modelling” for “topicmodelling”,“topic detection” for “topicmodelling”, “text classification” for “classificationalgorithms”, “similarity models” for “topicmodelling”, “a b testing” for “abtesting”, “a . b testing” for “abtesting”, “dimensionality reduction” for “dimensionalityreduction”, “ umap ” for “dimensionalityreduction”, “uniform manifold” for “dimensionalityreduction”, “dimension reduction” for “dimensionalityreduction”, “ tsne ” for “dimensionalityreduction”, “t - sne” for “dimensionalityreduction”, “ t sne ” for “dimensionalityreduction”, “feature engineering” for “featureengineering”, “hyperparameter tuning” for “hyperparametertuning”, “hyper parameter tuning” for “hyperparametertuning”, “data augmentation” for “dataaugmentation”, “model compression” for “modelcompression”, “naive bayes” for “bayesianmethods”, “naïve bayes” for “bayesianmethods”, “naï ve bayes” for “bayesianmethods”, “bayesian” for “bayesianmethods”, “knowledge distillation” for “knowledgedistillation”, “density based spatial clustering” for “densitybasedspatialclustering”,“density - based spatial clustering” for “densitybasedspatialclustering”, “ dbscan ” for “densitybasedspatialclustering”, “computational linguistics“ : “computationallinguistics”, “expectation maximization“ : “expectationmaximization”, “caffe deep learning framework“:“caffedeeplearningframework“.

Contract language:

“conducted remote” for “conductedremote”, “fully remote” for “fullyremote”, “home based” for “homebased”, “home work” for “homework”, “job is virtual” for “jobisvirtual”, “location is remote” for “locationisremote”, “location remote” for “locationremote”, “position is remote” for “positionisremote”, “position remote” for “positionremote”, “remote work” for “remotework”, “role will be remote” for “rolewil-

lberemote", "virtual job" for "virtualjob", "work from home" for "workfromhome", "work home" for "workhome", "work remote" for "workremote", "working remote" for "workingremote", "flx with office" for "flxwithoffice", "remote position" for "remoteposition", "role is remote" for "roleisremote", "remote basis" for "remotebasis", "hybrid work" for "hybridwork", "hybrid remote" for "hybridremote", "vice president" for "vicepresident", "project lead" for "projectlead", "team lead" for "teamlead", "supervising a team" for "supervisingateam", "trial period" for "trialperiod", "entry level" for "entrylevel", "tele commuting" for "telecommuting", "tele work" for "telework ", "no telecommuting" for "no", "does not offer teleworker" for "no", "no taking work home" for "no", "take work home" for "no", "teaching assistant" for "academic assistant", "secure position" for "permanent position", "temporary to permanent" for "temporary", "not permanent" for "temporary", "lead to permanent" for "temporary", "not temporary" for "permanent", "getting a permanent" for "temporary", "seasonal contract" for "seasonal contract", "seasonal hours" for "seasonal contract", "decommissions" for "decom", "commissioning" for "comis", "rewards and benefits" for "rewards packages", "pay - for - performance" for "pay for performance", "share options" for "stock options","based on performance" for "pay for performance", "short term, long term" for "short term contract, long term contract", "short term , long term" for "short term contract, long term contract", "short term and long term" for "short term contract and long term contract", "short term basis" for "short term contract","short term or long term" for "short term contract or long term contract", "contract is short term" for "short term contract", "or short term" for "or short term contract", " temp " for " temporary ", "long term conditions" for "conditions", "long term condition" for "condition", "long term strategy" for "strategy", "long term strategic" for "strategic", "long term satisfaction" for "satisfaction", "long term residential" for "residential", "long term financial" for "financial", "long term happiness" for "happiness", "long term need" for "need", "long term health" for "health", "temporary ongoing" for "temporary contract", "ongoing contract subject" for "temporary contract", " perm " for " permanent ", " perm, " for " permanent ", " ,perm " for " permanent ", " temp, " for " temporary ", " ,temp " for " temporary ", "experience is needed" for "experience needed", "years of experience" for "years experience", " yrs " for " years ", "don" t need experience" for "dont need experience", "don' t need previous" for "dont need experience" , "experience is essential" for "experience essential", "not essential" for "no essential", "not essential" for "no essential", "no formal qualifications" for "no formal qualification", "casual work" for "casual contract", "casual basis" for "casual contract", "casual work" for "casual contract", "casual hours" for "part time", "casual shifts" for "casual contract", "40 hours" for "full time", "40 hour" for "full time", "38.5 hour" for "full time", "38.5 hours" for "full time","38 hours" for "full time", "38 hour" for "full time","37 hours" for "full time", "37 hour" for "full time", "9 to 17" for "9 to 5", "9 17" for "9 to 5", "9 . 17" for "9 to 5", "profit sharing" for "profitsharing", "reward scheme" for "rewardscheme", "stock options" for "stockoptions", "variable pay" for "variablepay", "performance related" for "performancerelated", "rewards packages" for "rewardspackages", "pay for performance" for "payforperformance", "fixed term" for "fixedterm", "short term contract" for "shorttermcontract", "day to day contract" for "daytodaycontract", "zero hours" for "zerohours", "casual contract" for "casualcontract", "long term" for "longterm", "ongoing contract" for "ongoingcontract", "4 hour" for "4hour", "4 hours" for "4hour","6 hour" for "6hour", "6 hours" for "6hour","20 hour" for "20hour", "20 hours" for "20hour","30 hour" for "30hour", "9 to 5" for "9to5", "full time" for "fulltime", "standard hours" for "standardhours", "day shift" for "dayshift", "night shift" for "nightshift", "3 to 5 years" for "3to5years", "experience a minimum" for "experienceaminimum", "experience for" for "experiencefor", "experience in" for "experiencein", "experience is also" for "experienceisalso", "experience needed" for "experienceneeded", "experience of" for "experienceof", "experience preferably" for "experiencepreferably", "experience required" for "experiencerequired", "experience working" for "experienceworking", "extensive experience" for "extensiveexperience", "have experience" for "haveexperience", "have previous experience" for "havepreviousexperience", "have some knowledge" for "havesomeknowledge", "have worked" for "haveworked", "can teach" for "canteach", "minimum experience"

for “minimumexperience”, “months experience” for “monthsexperience”, “need experience” for “needexperience”, “past experience” for “pastexperience”, “practical experience” for “practicalexperience”, “prior experience” for “priorexperience”, “recent experience” for “recentexperience”, “relevant experience” for “relevantexperience”, “right experience” for “rightexperience”, “solid experience” for “solidexperience”, “strong experience” for “strongexperience”, “suitable experience” for “suitableexperience”, “three to five years” for “threetoFiveyears”, “two to three years” for “twotothreeyears”, “work experience” for “workexperience”, “worked in similar” for “workedinsimilar”, “required experience” for “requiredexperience”, “experience essential” for “experienceessential”, “experience is preferred” for “experienceispreferred”, “do not need any” for “donotneedany”, “do not need experience” for “donotneedexperience”, “experience is not” for “experienceisnot”, “no essential experience” for “noessentialexperience”, “no formal experience” for “noformalexperience”, “no formal qualifications or experience” for “noformalqualificationsorexperience”, “no past experience” for “nopastexperience”, “no work experience” for “noworkexperience”, “with no experience” for “withnoexperience”, “without any experience” for “withoutanyexperience”, “working directly with” for “workingdirectly”, “don t need experience” for “dontneedexperience”, “not needed” for “notneeded”, “working directly with” for “workingdirectlywith”, “report into” for “reportinto”, “your team leader” for “yourteamleader”, “working with” for “workingwith”, “working closely” for “workingclosely”, “under the direction” for “underthedirection”, “speaking to” for “speakingto”, “directly to” for “directlyto”, “work alongside” for “workalongside”, “required by” for “requiredby”, “keeping your” for “keepingyour”, “work closely” for “workclosely”, “step up to” for “stepto”, “managing the” for “managingthe”, “assistance to the” for “assistancetothe”, “work with” for “workwith”, “managed by” for “managedby”, “approval of” for “approvalof”, “communicate with” for “communicatewith”, “on experience” for “onexperience”, “negotiable salary” for “negotiablesalary”, “salary negotiable” for “salarynegotiable”, “hour negotiable” for “hournegotiable”, “salary is negotiable” for “salaryisnegotiable”, “salary up to” for “salaryupto“.

Education language:

“at level” for “level”, “at either level” for “level”, “ masters of ” for “ masterof ”,“ master in ” for “ masterof ”, “ masters in ” for “ masterof ”, “hour ” for “hours”, “skill ” for “skills ”, “cipd qualified” for “cipd”, “qualified to level” for “level”, “qualifications” for “qualification”, “mcipd” for “cipd”, “scqf level 1 ” for “scqflevel2 ”, “svq level 1 ” for “svqlevel2 ”, “ elc ” for “entrylevelcertificate”, “ cse “:“ certificateofsecondaryeducation ”, “a levels” for “alevelqualification”, “as levels” for “aslevelqualification”, “a level science” for “alevelqualification”, “a level in math” for “alevelqualification”, “a level in english” for “alevelqualification”, “as level in math” for “aslevelqualification”, “as level in english” for “aslevelqualification”, “as level in science” for “aslevelqualification”, “ certhe ” for “ certificateofhighereducation ”, “post graduate” for “postgraduate”, “ graduate ” for “ bachelor ”, “ bsc ” for “ bachelor ”, “ b . ed ” for “ bachelor ”, “ b. ed ” for “ bachelor ”, “ ba hons ” for “ bachelor ”, “ b eng ” for “ bachelor ”, “ b. eng ” for “ bachelorofengineering ”, “ beng ” for “ bachelorofengineering”, “ mba ” for “ masterofbusinessadministration ”, “ ll. b ” for “ bacheloroflaws”, “ meng ” for “ mastersofengineering ”, “ mmath ” for “ mastersofmathematics ”, “ mphil ” for “masterofphilosophy ”, “ mpsych ” for “ masterofphysics ”, “ mprof ” for “ masterofprofessionalstudies ”, “ mres ” for “ masterofresearch ”, “ msc ” for “ masterofscience ”, “ msc, ” for “ masterofscience ”, “ msci ” for “ masterofscience ”, “ msci, ” for “ masterofscience ”, “ mst ” for “ masterofstudies ”, “ m. sc ” for “ masterofscience ”, “ m . sc ” for “ masterofscience ”, “ llm ” for “ masteroflaw ”, “ ll . m ” for “ masteroflaw ”, “ mpharm ” for “ mastersofpharmacy ”, “ mchem ” for “ masterofchemistry ”, “ ms ” for “ masterofscience ”, “ msw ” for “ masterofsocialwork ”, “ pcge ” for “ postgraduatecertificateineducation ”, “ pcg ” for “ postgraduatecertificate ”, “ pgd ” for “ postgraduatediploma ”, “ pgcert ” for “ postgraduatecertificate ”, “ pg cert ” for “ postgraduatecertificate ”, “ph. d” for “ phd ”, “ph d” for “ phd ”, “ ma programme ” for “ masterofart ”, “ ma programme ” for “ masterofart ”, “ ma in english ” for “ masterofart ”, “hold an ma ” for “ masterofart ”, “ ma degree ” for “ masterofart ”, “honors degree” for “honoursdegree”, “self employed” for “selfemployed”, “self employment” for “selfemployed”,

“seasonal contract” for “seasonalcontract”, “basic functional skills” for “basicfunctionalskills”, “basic skills qualification” for “basicskillsqualification”, “entry level certificate” for “entrylevelcertificate”, “entry level functional skills” for “entrylevelfunctionalskills”, “entry level qualifications” for “entrylevelqualifications”, “entry nqf level” for “entrynqflevel”, “essential skills wales” for “essentialskillswales”, “level 1 scqf” for “level1scqf”, “level 1 svq” for “level1svq”, “level 2 scqf” for “level2scqf”, “level 2 svq” for “level2svq”, “level 3 scqf” for “level3scqf”, “level 3 svq” for “level3svq”, “no education requirements” for “noeducationrequirements”, “no formal qualification” for “noformalqualification”, “no previous experience or qualification” for “nopreviousexperienceorqualification”, “no previous qualification” for “nopreviousqualification”, “no qualification” for “noqualification”, “scottish national 1” for “scottishnational1”, “scottish national 2” for “scottishnational2”, “scottish national 3” for “scottishnational3”, “scqf level 2” for “scqflevel2”, “scqf level 3” for “scqflevel3”, “skills for life” for “skillsforlife”, “svq level 2” for “svqlevel2”, “svq level 3” for “svqlevel3”, “no formal education” for “noformaleducation”, “btec award” for “btecaward”, “btec first” for “btecfirst”, “btec level 1” for “bteclevel1”, “certificate of secondary education” for “certificateofsecondaryeducation”, “first certificate” for “firstcertificate”, “foundation welsh baccalaureate” for “foundationwelshbaccalaureate”, “key skills level 1” for “keyskillslevel1”, “level 1 award” for “level1award”, “level 1 certificate” for “level1certificate”, “level 1 diploma” for “level1diploma”, “level 1 esol” for “level1esol”, “level 1 essential skills” for “level1essentialskills”, “level 1 functional skills” for “level1functionalskills”, “level 1 key skills” for “level1keyskills”, “level 1 national vocational qualification” for “level1nationalvocationalqualification”, “level 1 nvq” for “level1nvq”, “nqf level 1” for “nqflevel1”, “nvq level 1” for “nvqlevel1”, “rqf level 1” for “rqflevel1”, “scqf level 4” for “scqflevel4”, “diploma level 1” for “diplomalevel1”, “national level 1” for “nationallevel1”, “qualification level 1” for “qualificationlevel1”, “level 1 qualification” for “level1qualification”, “level 1 qcf” for “level1qcf”, “functional skills level 1” for “functionalskillslevel1”, “essential skills level 1” for “essentialskillslevel1”, “level 1 apprenticeship” for “level1apprenticeship”, “cipd level 1” for “cipdlevel1”, “level 1 cipd” for “level1cipd”, “level 4 scqf” for “level4scqf”, “level 4 svq” for “level4svq”, “svq level 4” for “svqlevel4”, “btec level 2” for “bteclevel2”, “general certificate of education” for “generalcertificateofeducation”, “intermediate apprenticeship” for “intermediateapprenticeship”, “intermediate welsh baccalaureate” for “intermediatewelshbaccalaureate”, “key skills level 2” for “keyskillslevel2”, “level 2 award” for “level2award”, “level 2 certificate” for “level2certificate”, “level 2 diploma” for “level2diploma”, “level 2 esol” for “level2esol”, “level 2 essential skill” for “level2essentialskill”, “level 2 functional skill” for “level2functionalskill”, “level 2 key skills” for “level2keyskills”, “level 2 national” for “level2national”, “level 2 nvq” for “level2nvq”, “national level 2” for “nationallevel2”, “nqf level 2” for “nqflevel2”, “nvq level 2” for “nvqlevel2”, “rqf level 2” for “rqflevel2”, “scqf level 5” for “scqflevel5”, “secondary education” for “secondaryeducation”, “diploma level 2” for “diplomalevel2”, “award level 3” for “awardlevel3”, “esol level 2” for “esollevel2”, “functional skills level 2” for “functionalskillslevel2”, “essential skills level 2” for “essentialskillslevel2”, “qualification level 2” for “qualificationlevel2”, “level 2 qualification” for “level2qualification”, “level 2 apprenticeship” for “level2apprenticeship”, “cipd level 2” for “cipdlevel2”, “level 2 cipd” for “level2cipd”, “level 5 scqf” for “level5scqf”, “level 5 svq” for “level5svq”, “svq level 5” for “svqlevel5”, “certificate of secondary education” for “certificateofsecondaryeducation”, “a level qualification” for “alevelqualification”, “as level qualification” for “aslevelqualification”, “access to higher education diploma” for “accesstohighereducationdiploma”, “advanced apprenticeship” for “advancedapprenticeship”, “advanced diploma” for “advanceddiploma”, “advanced extension” for “advancedextension”, “advanced welsh baccalaureate” for “advancedwelshbaccalaureate”, “btec award level 3” for “btecawardlevel3”, “btec certificate level 3” for “bteccertificatelevel3”, “btec diploma level 3” for “btecdiplomalevel3”, “btec level 3” for “bteclevel3”, “btec national” for “btecnational”, “diploma programme” for “diplomaprogramme”, “extended certificate” for “extendedcertificate”, “extended diploma” for “extendeddiploma”, “key skills level 3” for “keyskillslevel3”, “level 3 award” for “level3award”, “level 3 certificate” for “level3certificate”, “level 3 esol” for “level3esol”, “level 3 national” for “level3national”, “level 3 nvq” for “level3nvq”, “national

level 3” for “nationallevel3”, “nqf level 3” for “nqflevel3”, “nvq level 3” for “nvqlevel3”, “ocr national” for “ocrnational”, “progression diploma” for “progressiondiploma”, “rqf level 3” for “rqflevel3”, “scottish advanced higher” for “scottishadvancedhigher”, “scottish baccalaureate” for “scottishbaccalaureate”, “scottish higher” for “scottishhigher”, “scqf level 6” for “scqflevel6”, “diploma level 3” for “diplomalevel3”, “award level 3” for “awardlevel3”, “functional skills level 3” for “functionalskillslevel3”, “essential skills level 3” for “essentialskillslevel3”, “qualification level 3” for “qualificationlevel3”, “level 3 qualification” for “level3qualification”, “level 3 apprenticeship” for “level3apprenticeship”, “cipd level 3” for “cipdlevel3”, “level 3 cipd” for “level3cipd”, “level 6 scqf” for “level6scqf”, “level 6 svq” for “level6svq”, “svq level 6” for “svqlevel6”, “certificate of higher education” for “certificateofhighereducation”, “higher apprenticeship” for “higherapprenticeship”, “higher national certificate” for “highernationalcertificate”, “level 4 award” for “level4award”, “level 4 certificate” for “level4certificate”, “level 4 diploma” for “level4diploma”, “level 4 national” for “level4national”, “level 4 nvq” for “level4nvq”, “national level 4” for “nationallevel4”, “nqf level 4” for “nqflevel4”, “nvq level 4” for “nvqlevel4”, “scqf level 7” for “scqflevel7”, “diploma level 4” for “diplomalevel4”, “award level 4” for “awardlevel4”, “functional skills level 4” for “functionalskillslevel4”, “essential skills level 4” for “essentialskillslevel4”, “qualification level 4” for “qualificationlevel4”, “level 4 qualification” for “level4qualification”, “level 4 apprenticeship” for “level4apprenticeship”, “cipd level 4” for “cipdlevel4”, “level 4 cipd” for “level4cipd”, “level 7 scqf” for “level7scqf”, “level 7 svq” for “level7svq”, “svq level 7” for “svqlevel7”, “btec professional” for “btecprofessional”, “diploma of further education” for “diplomaoffurthereducation”, “diploma of higher education” for “diplomaofhighereducation”, “foundation degree” for “foundationdegree”, “higher diploma” for “higherdiploma”, “higher national diploma” for “highernationaldiploma”, “level 5 award” for “level5award”, “level 5 certificate” for “level5certificate”, “level 5 diploma” for “level5diploma”, “level 5 national” for “level5national”, “level 5 nvq” for “level5nvq”, “national level 5” for “nationallevel5”, “nqf level 5” for “nqflevel5”, “nvq level 5” for “nvqlevel5”, “scqf level 8” for “scqflevel8”, “diploma level 5” for “diplomalevel5”, “award level 5” for “awardlevel5”, “qualification level 5” for “qualificationlevel5”, “level 5 qualification” for “level5qualification”, “cipd level 5” for “cipdlevel5”, “level 5 cipd” for “level5cipd”, “level 8 scqf” for “level8scqf”, “level 5 apprenticeship” for “level5apprenticeship”, “level 5 professional” for “level5professional”, “level 8 svq” for “level8svq”, “svq level 8” for “svqlevel8”, “bachelors degree” for “bachelorsdegree”, “btec advanced professional” for “btecadvancedprofessional”, “degree apprenticeship” for “degreeapprenticeship”, “first degree” for “firstdegree”, “bachelor certificate” for “bachelorcertificate”, “bachelor diploma” for “bachelordiploma”, “level 6 award” for “level6award”, “level 6 certificate” for “level6certificate”, “level 6 diploma” for “level6diploma”, “level 6 national” for “level6national”, “level 6 nvq” for “level6nvq”, “national level 6” for “nationallevel6”, “nqf level 6” for “nqflevel6”, “nvq level 6” for “nvqlevel6”, “ordinary degree” for “ordinarydegree”, “oxbridge ma” for “oxbridgema”, “scottish ma” for “scottishma”, “scqf level 10” for “scqflevel10”, “scqf level 9” for “scqflevel9”, “higher education” for “highereducation”, “diploma level 6” for “diplomalevel6”, “award level 6” for “awardlevel6”, “qualification level 6” for “qualificationlevel6”, “level 6 qualification” for “level6qualification”, “cipd level 6” for “cipdlevel6”, “level 6 cipd” for “level6cipd”, “level 9 scqf” for “level9scqf”, “level 10 scqf” for “level10scqf”, “level 6 apprenticeship” for “level6apprenticeship”, “level 6 professional” for “level6professional”, “level 10 svq” for “level10svq”, “svq level 10” for “svqlevel10”, “have a degree” for “haveadegree”, “bsc or masters” for “bscormasters”, “degree or master” for “degreeormaster”, “bsc or msc” for “bscormsc”, “ba or ma” for “baorma”, “honours degree” for “honoursdegree”, “award level 7” for “awardlevel7”, “ba and ma” for “baandma”, “cipd level 7” for “cipdlevel7”, “diploma level 7” for “diplomalevel7”, “have a master” for “haveamaster”, “have a masters” for “haveamasters”, “level 11 scqf” for “level11scqf”, “level 11 svq” for “level11svq”, “level 7 apprenticeship” for “level7apprenticeship”, “level 7 award” for “level7award”, “level 7 certificate” for “level7certificate”, “level 7 cipd” for “level7cipd”, “level 7 diploma” for “level7diploma”, “level 7 national” for “level7national”, “level 7 nvq” for “level7nvq”, “level 7 professional” for “level7professional”, “level 7 qcf” for “level7qcf”, “master of arts” for “masterofarts”,

“master of business administration” for “masterofbusinessadministration”, “master of engineering” for “masterofengineering”, “master of law” for “masteroflaw”, “master of mathematics” for “masterofmathematics”, “master of pharmacy” for “masterofpharmacy”, “master of philosophy” for “masterofphilosophy”, “master of physics” for “masterofphysics”, “master of professional studies” for “masterofprofessionalstudies”, “master of public health” for “masterofpublichealth”, “master of research” for “masterofresearch”, “master of science” for “masterofscience”, “master of social work” for “masterofsocialwork”, “master of studies” for “masterofstudies”, “masters by research” for “mastersbyresearch”, “masters degree” for “mastersdegree”, “national level 7” for “nationallevel7”, “nqf level 7” for “nqflevel7”, “nvq level 7” for “nvqlevel7”, “postgraduate certificate” for “postgraduatecertificate”, “postgraduate certificate in education” for “postgraduatecertificateineducation”, “postgraduate diploma” for “postgraduatediploma”, “professional masters” for “professionalmasters”, “scqf level 11” for “scqflevel11”, “svq level 11” for “svqlevel11”, “doctor of philosophy” for “doctorofphilosophy”, “doctoral degree” for “doctoraldegree”, “doctors degree” for “doctorsdegree”, “level 8 award” for “level8award”, “level 8 certificate” for “level8certificate”, “level 8 diploma” for “level8diploma”, “level 8 national” for “level8national”, “level 8 nvq” for “level8nvq”, “national level 8” for “nationallevel8”, “nqf level 8” for “nqflevel8”, “nvq level 8” for “nvqlevel8”, “rqf level 8” for “rqflevel8”, “scqf level 12” for “scqflevel12”, “diploma level 8” for “diplomalevel8”, “award level 8” for “awardlevel8”, “qcf level 8” for “qcflevel8”, “level 8 qcf” for “level8qcf”, “cipd level 8” for “cipdlevel8”, “level 8 cipd” for “level8cipd”, “level 12 scqf” for “level12scqf”, “level 8 apprenticeship” for “level8apprenticeship”, “level 8 professional” for “level8professional”, “svq level 12” for “svqlevel12”, “level 12 svq” for “level12svq“.

The Value of Personality: Exploring the Wage Penalty Paradox in the UK

Matías Golman*

Abstract

This study examines the relationship between personality traits and wage offers. Using text analysis of 11.7 million UK online job postings from 2017 to 2022, I identify Big Five personality traits in vacancy descriptions. Jobs mentioning personality traits offer 4.6 per cent lower wages on average, creating a paradox: why do employers advertise for traits if these are penalised? I test four explanations: correlation with low-pay job characteristics, compensating differentials, linguistic intensity effects, and demographic targeting. Trait penalties are significantly reduced by the inclusion of occupation fixed effects. Penalties are larger in high-skill occupations but smaller in female-dominated roles. Trait penalties decrease when combined with traditional benefits and workplace culture signals, but increase with remote work. Linguistic intensity analysis shows that low-intensity mentions carry minimal penalties, but high-intensity mentions show substantial penalties. Machine learning analysis of gender targeting contradicts simple demographic targeting predictions. The findings suggest employers use personality-traits for multiple purposes rather than a single mechanism: signalling requirements at low intensity, facilitating compensating differentials with appropriate amenity bundles, and serving organisational signalling functions at high intensity.

Keywords: personality traits, Big Five, job vacancy data, wage determination, compensating differentials, UK labour market, text analysis
JEL Codes: J31, J23, J24, C55

*University of Nottingham. Email: matias.golman@gmail.com.

Contents

1 Introduction **128**

2 Literature Review **130**

2.1 Individual-Level Evidence: Personality Traits and Wage Premiums 130

2.1.1 Direct Wage Effects 130

2.1.2 Mechanisms and Heterogeneity 131

2.2 Job Advertisement Evidence: Personality Traits and Wage Penalties . . . 132

2.2.1 The Demand-Side Paradox 132

3 Data **133**

3.1 Data Source and Construction Process 134

3.2 Occupation and Firm Characteristics 135

3.3 Retrieving Personality Traits and Job Characteristics 136

4 Baseline Analysis: The Personality Traits Wage Penalty **137**

4.1 Empirical Methodology 138

4.2 Results 138

4.2.1 Evidence of the Wage Penalty 138

4.2.2 Controlling for Job Characteristics and Fixed Effects 139

4.2.3 Occupation-Specific Trait Effects 141

5 Correlation with potential low-pay characteristics **142**

5.1 Results 145

5.1.1 Occupation-level heterogeneity 145

5.1.2 Job-requirement heterogeneity 148

5.1.3 Correlation with firm characteristics 151

5.1.4 Assessment of the Low-Pay Correlation Hypothesis 152

6 Compensating Differentials **154**

6.1 Empirical Methodology 155

6.1.1 Identifying Advertised Amenities 155

6.1.2 Empirical Models 158

6.2 Results 159

6.2.1 Trait-Amenity Correlations 159

6.2.2 Amenity Controls and Trait Wage Penalties 161

6.2.3 Interaction between Amenities and Traits 162

6.2.4 Assessment of the Compensating Differentials Hypothesis 166

7 Language Intensity **166**

7.1 Empirical Methodology 167

7.1.1 Empirical models 168

7.2 Results 169

7.2.1 Intensity scores and linear effects 169

7.2.2 Non-linear effects across intensity quintiles 170
7.2.3 Interaction with firm characteristics 172
7.2.4 Frequency-Based Intensity Analysis 172
7.2.5 Trait Position Effects 173
7.2.6 Assessment of the Linguistic Intensity Hypothesis 174
8 Demographic Targeting **175**
8.1 Empirical Methodology 175
8.1.1 Empirical models 177
8.2 Results 178
8.2.1 Gender Targeting and Personality Trait Prevalence 178
8.2.2 Wage Effects by Gender Targeting 178
8.2.3 Trait Intensity and Gender Targeting Interactions 180
8.2.4 Assessment of the Demographic Targeting Hypothesis 181
9 Conclusion **183**
10 Appendices **194**

1 Introduction

Individuals' personality traits are highly demanded in UK labour markets. Most British employers use personality tests during recruitment, showing that these characteristics matter in employment decisions (McGee and McGee, 2020). Individual-level studies consistently find that people with higher scores on traits like Conscientiousness, Openness, Extraversion, and Emotional Stability typically earn more, although high Agreeableness is associated with a wage penalty (Alderotti et al., 2023). Yet, recent research on online job advertisements reveals a puzzling contradiction: positions mentioning personality traits or character attributes are associated with *lower* wages (Acemoglu et al., 2022; Brenčič and McGee, 2023b; Deming and Kahn, 2018).

This negative association contradicts human capital theory. Skills and attributes that enhance productivity should be rewarded in the labour market. When employers explicitly request certain characteristics, economic theory suggests they value these attributes enough to pay for them. This expectation is strengthened by research on other job requirements such as education, experience, and technical skills, which show positive wage premiums when mentioned in advertisements (Alekseeva et al., 2021; Deming and Kahn, 2018; Hershbein and Kahn, 2018). Why would employers emphasise personality traits in job advertisements if these are not economically rewarded?

Analysing personality traits through job advertisements captures employers' strategies to attract candidates rather than workers' actual endowments, offering a distinct perspective from individual-level data. I identify four potential explanations from labour and personnel economics literature which can be tested using this data.

Correlation with other low-pay characteristics: Personality trait mentions might cluster systematically in lower-paying contexts. Trait mentions could correlate with job, occupation, or firm characteristics that determine wages (Brenčič and McGee, 2023b; Deming and Kahn, 2018; Woodcock, 2008). For instance, customer service or care-related occupations may mention more traits and simultaneously pay lower wages due to low market value. I first test this by including high-dimensional fixed effects. Then I examine how trait-wage relationships vary across occupations and interact traits with job and firm characteristics.

Compensating differentials through person-organisation fit: Following the theory of equalising differences, workers might trade monetary compensation for non-monetary benefits (Rosen, 1986). When jobs signal rewarding features such as flexible work, supportive culture, or meaningful impact, employers can offer lower wages whilst still attracting qualified candidates (Burbano, 2016; Escudero et al., 2024). Similarly, environments emphasising personality traits may attract individuals who value being valued for their

personality (Cable and Judge, 1996; Kristof-Brown et al., 2005; Stevens and Szmerekovsky, 2010). I test this by examining whether amenity provision explains trait penalties and whether traits interact with specific amenities.

Demographic targeting and statistical discrimination: Drawing on models of statistical discrimination (Aigner and Cain, 1977; Arrow, 1971), employers might use personality trait language to target demographic groups they believe will accept lower wages. If employers associate certain traits with specific demographics, they might use trait language strategically to induce self-selection (Gaucher et al., 2011; Kuhn et al., 2020). Recent experimental evidence shows that trait language in advertisements affects application patterns differently across demographic groups (Abraham et al., 2024; Flory et al., 2015; Oldford and Fiset, 2021). I test this using machine learning to classify implicit gender targeting in advertisements.

Linguistic intensity and signalling: The intensity and placement of trait mentions may alter their interpretation and value signalling. Information structure theory suggests that message frequency, prominence, and contextual placement shape how receivers process and value information (Koch and Zerback, 2013; Lambrecht, 1996; Rosenfeld et al., 2015). I test this by measuring trait mention intensity using TF-IDF text analysis and examining whether high-intensity mentions show different wage effects than low-intensity mentions.

I begin by establishing a baseline model documenting the negative relationship between personality trait mentions and wages. Mentioning personality-related language in advertisements is associated with 4.6 per cent lower wages on average. The penalty is largest for Extraversion (3.2 per cent) and Emotional Stability (3.1 per cent), followed by Conscientiousness (2.0 per cent) and Agreeableness (1.4 per cent). Only Openness shows a small positive effect (0.7 per cent).

I then test each hypothesis using complementary methodological approaches. For correlation with low-pay characteristics, I use occupation, job, and firm fixed effects and interaction terms. For compensating differentials, I examine amenity controls and trait-amenity interactions. For linguistic intensity, I use TF-IDF text analysis. For demographic targeting, I employ BERT machine learning models to classify implicit gender preferences in advertisements.

The findings reveal that the relationship between personality trait mentions and wages reflects multiple overlapping employer strategies rather than a single mechanism. Trait language varies by intensity, context, and organisational characteristics. The persistence of penalties even with extensive controls, combined with context-dependent effects, suggests employers strategically deploy personality trait language for different purposes depending on job characteristics, linguistic emphasis, amenity bundles, and organisational positioning.

This paper proceeds as follows. Section 2 reviews the literature on personality traits' effect on wages. Section 3 describes the Adzuna job vacancy data. Section 4 presents the baseline analysis of the personality trait wage penalty. Sections 5–8 examine each of the four explanations presented in turn, providing both methodological approaches and results. Section 9 concludes.

2 Literature Review

2.1 Individual-Level Evidence: Personality Traits and Wage Premiums

Economists recognise that personality traits explain differences in economic outcomes (Borghans et al., 2008; Heckman and Kautz, 2012). However, the literature is not clear on the distinction about personality traits, soft skills, social skills or non-cognitive attributes, with authors often mixing their meanings and scope (Woessmann, 2024). I define personality traits as stable individual characteristics that remain relatively fixed after early adulthood, following psychology literature (Cobb-Clark and Schurer, 2012; McCrae and John, 1992).

Unlike soft skills or social skills, which are dynamic and learnable, personality traits are largely uncorrelated with cognitive abilities and difficult to modify through training (Grugulis and Vincent, 2009; Matteson et al., 2016). This stability gives employers strong incentives to select for required traits during hiring rather than developing them later. I use the Big Five framework (Openness, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability), which has become the standard taxonomy in economics.

2.1.1 Direct Wage Effects

Individual-level studies consistently document positive wage returns to personality traits, with the exception of Agreeableness. Flinn et al. (2025) show that higher Conscientiousness and Emotional Stability, and lower Agreeableness increase earnings and job stability, with effects ranging between 1-5 per cent according to trait and gender. Mueller and Plug (2006) find that low Agreeableness increases earnings by 10 per cent for men and 5 per cent for women. Nyhus and Pons (2005) show that Emotional Stability and Extraversion correlate with higher wages, while Agreeableness reduces wages, particularly for women. Heineck (2011) reports 2-3 per cent wage premiums for Openness but similar penalties for Agreeableness in the UK.

The most comprehensive evidence comes from Alderotti et al. (2023)'s meta-analysis of 62 studies. They find consistent patterns across countries: Conscientiousness, Openness, and Extraversion show positive wage effects averaging 2-3 per cent, while Agreeableness

and Neuroticism[1] show negative effects of similar magnitude. These effects persist after controlling for education, experience, and cognitive ability.

2.1.2 Mechanisms and Heterogeneity

Individual-level evidence mostly uses Bowles et al. (2001) "incentive-enhancing preferences" to explain wage effects. This theory posits that workers with certain characteristics, particularly Conscientiousness, require less monitoring and show greater reliability, making them more productive. Empirical support comes from several studies. Cubel et al. (2016) show in a laboratory experiment that conscientious individuals are more productive when completing tasks. Judge and Cable (1997) and Barrick and Mount (1991) find that Conscientiousness predicts job performance across occupations, while Extraversion particularly benefits management and sales roles.

Returns to personality traits vary across contexts. Gensowski (2018) shows that personality effects peak between ages 40-60, with larger returns for highly-educated workers. Collischon (2020) uses quantile regressions that show increased returns at higher wage levels. In contrast, Nyhus and Pons (2005) find that Conscientiousness affects wages primarily at career start, suggesting that employers reward the trait to increase the probability of recruiting and retaining the conscientious worker, but stop this once they learn about the workers' productivity.

Personality differences can also partly explain gender inequalities. Using German data, Braakmann (2009) finds that personality differences account for 18 per cent of the gender wage gaps and 7 per cent of full-time employment gaps. Nyhus and Pons (2012) show that Agreeableness and Intellect explain 11.5 per cent of the Dutch gender wage gap. Risse et al. (2018) find that men's lower Agreeableness and higher confidence contribute to Australian gender wage gap.

Occupation-specific returns matter considerably. John and Thomsen (2014) show identical traits have different returns across occupations. Cobb-Clark and Tan (2011) find that personality affects both occupation selection and wages within occupations. Denissen et al. (2018) show that personality-job fit, the congruence between individual traits and job demands, influences income.

Recent evidence suggests these patterns evolve with labour market changes. Deming (2017) and Weinberger (2014) document increasing returns to social skills and character attributes over time, particularly in team-based work. Rohrbach-Schmidt et al. (2023) find that trait-wage relationships depend on job task requirements, with stronger effects for non-routine and interactive tasks.

[1] The opposite of Emotional Stability.

2.2 Job Advertisement Evidence: Personality Traits and Wage Penalties

The availability of online job posting data enables direct observation of employer demands. However, this literature finds the opposite pattern: personality trait mentions correlate with *lower* wages.

2.2.1 The Demand-Side Paradox

Deming and Kahn (2018) pioneered using job postings to measure job requirements across firms and labour markets, finding that social skills are positively correlated with wages and firm performance, particularly when combined with cognitive skills. The authors use a *character* feature inspired by broad personality attributes from Heckman and Kautz (2012). Similar to later studies by Alekseeva et al. (2021) and Acemoglu et al. (2022), they find negative wage correlations between this *character* control variable and posted wages.

Only two studies examine wage effects from Big Five traits specifically. Brenčič and McGee (2023b) analyse 140,193 job advertisements from Monster.com posted over a two-week period in 2006, using dictionaries based on trait-descriptive adjectives from Goldberg (1981) and John (1990). They find that 54 per cent of advertisements mention at least one personality trait, with employers primarily demanding Extroversion (31 per cent), Conscientiousness (26 per cent), and Openness (21 per cent).

Their analysis reveals negative wage penalties for most traits. Among advertisements posting hourly wages, they find negative coefficients for Extroversion (4.2 per cent with firm fixed effects), Agreeableness (6.3 per cent), and Emotional Stability (6.3 per cent), while only Openness shows positive effects (5 per cent). These findings establish a clear wage penalty paradox: despite employers explicitly demanding personality traits in advertisements, most traits are associated with lower posted wages.

Beyond wage effects, Brenčič and McGee (2023b) make a key theoretical contribution by examining how personality traits relate to compensation structure rather than wage levels. They develop a model where workers differ in intrinsic motivation and firms choose between fixed wages and incentive-based compensation. In their framework, conscientious workers supply effort even without extrinsic incentives, making fixed wage contracts more attractive to firms hiring such workers. Consistent with this model, advertisements demanding conscientious workers are 3.7 percentage points less likely to offer incentive pay and 2.7 percentage points less likely to reference promotion opportunities. These findings suggest firms use personality requirements as substitutes for extrinsic motivation mechanisms when selecting workers who are intrinsically motivated. While this compensation structure mechanism differs from the compensating differentials and demographic targeting hypotheses examined in the present study, it provides important context for understanding how personality trait language functions in job advertisements

beyond simple wage determination.

In a subsequent study, Brenčič and McGee (2023a) investigate whether task requirements can explain the personality-wage relationship using the same Monster.com dataset. They examine employers' demands for personality traits across different job tasks and test for task-specific wage returns using the National Longitudinal Survey of Youth 1997. Their job advertisement analysis shows that trait demands correlate with general tasks (communication, teamwork, problem-solving) but are negatively correlated with routine, manual, and mathematics tasks.

Using longitudinal data with respondent and occupation fixed effects, they find limited evidence of task-specific wage returns to personality traits. Most interaction coefficients between traits and tasks are insignificant. The few significant results include a 1.2 per cent wage increase for Extroversion in teamwork-intensive occupations and a 2.3 per cent increase for Openness in communication-intensive roles. However, these effects are modest and inconsistent. Brenčič and McGee (2023a) conclude that task-specific wage returns do not provide compelling economic incentives for personality-based sorting, suggesting the wage penalties observed in job advertisements cannot be explained by differential returns to personality traits across tasks.

The present study builds on Brenčič and McGee (2023b)'s establishment of the wage penalty paradox but advances the literature in important ways. While their analysis uses a two-week sample from 2006, this research examines 11.7 million job postings spanning 2017-2022, providing contemporary relevance and statistical power to investigate underlying mechanisms. Where Brenčič and McGee (2023b) identify the paradox and test one potential explanation (incentive pay substitution for conscientiousness) and Brenčič and McGee (2023a) examine task-specific returns, this study provides the first systematic investigation of multiple competing explanations: correlation with occupation characteristics through extensive fixed effects and interaction analysis, compensating differentials through comprehensive amenity provision measures, linguistic intensity effects through TF-IDF analysis, and demographic targeting through machine learning classification of implicit gender preferences. Additionally, this study advances methodologically by implementing TF-IDF measures to capture trait mention intensity and employs BERT models to identify implicit demographic targeting patterns not examined in prior work.

3 Data

This study uses the Adzuna job vacancy dataset employed in Chapter 1. This chapter focuses on wage-related aspects and firm characteristics that influence wage determination. For details on general data characteristics and personality trait identification, see Chapter 1.

3.1 Data Source and Construction Process

The data construction process began with 276.3 million raw job advertisements from Adzuna, an aggregator whose algorithms continuously collect information on open vacancies from online job boards. The Office for National Statistics (ONS) of the United Kingdom uses Adzuna to derive its weekly vacancy indicators, reflecting the dataset's broad coverage of the UK labour market.

In the raw dataset, approximately 65 per cent of vacancies contained wage information, higher than other job vacancy sources used in the literature, which typically report wage information in only 13-25 per cent of postings (Banfi and Villena-Roldan, 2019; Brenčič, 2012; Kuhn and Shen, 2023; Marinescu and Wolthoff, 2020). This high wage coverage is a result of Adzuna's efforts in identifying wages from the different advertisement sections. It makes the dataset suitable for analysing the relationship between job requirements and offered wages.

To create a dataset appropriate for wage analysis, I applied a filtering and cleaning process shown in Table 1. First, I removed advertisements posted by recruitment agencies (45 per cent of vacancies), as these intermediaries make it impossible to identify the actual employing firm or industry sector. Without industry information, advertisements could not be accurately classified into occupational categories.

Table 1: Cleaning of the database

Process	**N vacancies**	**N no missing W***	**Mean W***	**SD W***	**Δ**
Adzuna raw database	276,325,386	178,229,590	31,120	18,524	-
Without recruitment agencies	150,918,209	83,727,680	29,882	18,558	-45%
Without missing companies match	66,833,773	32,647,950	28,365	18,392	-56%
Keep only Private companies with SIC	64,642,205	31,619,913	28,303	18,468	-3%
Remove duplicates vacancy-id, source-id, same description-location	20,783,111	10,490,006	28,342	18,830	-68%
Remove salaries missing	9,657,106	9,657,106	29,454	18,964	-54%
Incorporate new salaries from cleaned description	12,103,405	11,866,183	31,930	18,873	25%
Remove outliers in wages and text size	11,659,923	11,430,457	32,018	18,944	-4%

Notes: Author's creation based on Adzuna. W∗ = Salary considered after trimming 1% tails.

Next, I matched company names from the remaining advertisements with the Companies House database, a registry of UK businesses. This matching process required text normalisation to account for variations in company name formats and allowed me to identify Standard Industrial Classification (SIC) codes for each vacancy. After restricting to private companies with valid SIC codes and removing duplicates through multiple

deduplication strategies, 20.8 million unique vacancies remained.

From this set, 9.7 million vacancies had structured wage fields in Adzuna's data. To expand wage information coverage, I implemented a text analysis procedure to extract additional wage information from unstructured job descriptions. This process involved:

1. Standardising salary formats through regular expressions, addressing variations such as annual figures (e.g., "£35,000"), ranges (e.g., "£30,000-£40,000"), abbreviations (e.g., "£35k"), and hourly/weekly/monthly rates.
2. For range values, calculating the mean of the lower and upper bounds.
3. Converting non-annual rates to annual equivalents based on the job's full-time/part-time status.
4. Applying a hierarchical decision rule that prioritised annual figures from the text, followed by structured wage fields after verification.

This wage extraction process recovered an additional 2.4 million vacancies with wage information not considered in Adzuna's wage fields. After removing outliers in both wages and text length, the final dataset contained 11.7 million vacancies, all with valid wage information, advertised by 128,559 companies.

The median nominal salary in online advertisements posted on Adzuna for the period 2017-2022 was £25,960, which is 15 per cent lower than the median income reported by the Annual Survey of Hours and Earnings (ASHE) for the UK during the same period. There are three potential explanations. First, Adzuna reports starting offered wages rather than realised wages. Second, these posted salaries do not include bonuses, performance-related rewards, overtime or other variable income. Third, the Adzuna dataset has more low-paid occupations compared to the wider labour market. For example, Caring Personal Services and Customer Services jobs, which typically offer below-median wages, make up about 18 per cent of vacancies in the Adzuna database but only 10 per cent of occupations in ASHE. Finally, the difference partly reflects the potential selection bias noted by Batra et al. (2023) that posted wages are more common in advertisements seeking less-skilled workers.

3.2 Occupation and Firm Characteristics

Since Adzuna does not provide standard occupational classifications, I implemented a machine learning algorithm developed by Turrell et al. (2019) to classify advertisements into 3-digit Standard Occupation Codes (SOC). This algorithm assigns SOC codes based on job title, description, and industry information derived from the Companies House matching, with a reported accuracy rate of 76 per cent.

To investigate mechanisms through which personality trait demands might affect wages, I complemented the vacancy data with information at both occupation and firm levels.

At the occupation level, I incorporated data from the UK Household Longitudinal Study (UKHLS) on gender composition, educational attainment, migrant proportion, and average worker age for each SOC code. The UKHLS also includes self-reported Big Five personality assessments based on a 7-point scale, enabling comparison of employer demands for traits with actual trait distributions across occupations. These occupation-level characteristics are essential for testing whether the trait-wage relationship reflects sorting or demographic targeting mechanisms.

At the firm level, I matched companies with the Financial Analysis Made Easy (FAME) database, which provided information on firm size (employee count), annual profit, public listing status, and directors' demographic characteristics. This firm-level information allows testing whether organisational characteristics influence the relationship between personality traits and wages.

Together, these data linkages enable analysis of how personality trait demands relate to wages across different types of jobs, occupations, and firms, providing the foundation for testing the four mechanisms regarding why employers use personality-related language despite potential wage penalties.

3.3 Retrieving Personality Traits and Job Characteristics

This paper builds on the personality trait identification methodology developed in Chapter 1. I use the established Big Five personality framework widely used in economic and psychological research (Almlund et al., 2011; Borghans et al., 2008; Heckman and Kautz, 2012). This taxonomy categorises personality into five dimensions: Conscientiousness, Openness to experience, Extraversion, Agreeableness, and Emotional Stability.

Using a knowledge engineering approach, I identify personality traits through dictionaries compiled from authoritative sources including Goldberg (1990), Costa Jr and McCrae (1992), Goldberg (1992), McCrae and John (1992), John (1990), Saucier (1994) and Heckman and Kautz (2012). This approach allows for direct application of established taxonomic work while avoiding ambiguous or context-dependent terms that might lead to false positives. This process identified words with multiple meanings, such as *flexible* or *trust*, which could refer to work schedules rather than personality traits.

To minimise this risk, words were only included after being validated through context analysis: 8 samples of 100 randomly selected sentences containing each potential trait word were examined to ensure the term genuinely referred to a personality characteristic

in the context of job requirements. I modified terms in contexts not referring to worker or team characteristics before dictionary matching. For instance, replace "flexible hours" with "flx hours". I removed words that, regardless of context, did not refer to personality traits in job vacancies, such as *modest*, *conventional*, *deliberate*, *order*, *aesthetic*, *deep*, *relaxed* or *comfortable*.

A job posting is labelled as requiring a specific personality trait if it includes at least one keyword from the corresponding trait dictionary in its job requirements or job title.[2] This boolean approach generates binary indicators for each of the five personality traits, which are then related to wage offers to test the hypothesis that trait mentions correlate with lower wages even after controlling for various job, occupation, and firm characteristics.

In addition to personality traits and wages, I created several other variables through text analysis of job advertisements. For location identification, I mapped Adzuna's location variables to the Nomenclature of Territorial Units for Statistics (NUTS), achieving 96 per cent coverage at the one-digit level and 85 per cent at the two-digit level. Education requirements were classified into five categories following Nandi and Nicoletti (2014), though notably only 12 per cent of vacancies explicitly mentioned education requirements.

I created indicators for hard and soft skills based on expanded dictionaries from Deming and Kahn (2018); Khaouja et al. (2021); Lyu and Liu (2021). I built 16 categories of amenities, inspired by the methodology and dictionaries from Escudero et al. (2024) and Boschetti Adamczyk et al. (2025). Remote work indicators were based on Draca et al. (2022)'s dictionary, capturing the increase in remote work mentions following the COVID-19 pandemic. Experience requirements were coded as binary indicators, while seniority levels (entry, mid, senior) were identified following hierarchical descriptions from previous literature. Job arrangement variables (part-time/full-time, temporary/permanent) combined Adzuna's structured fields with text-derived information when the former were missing. All dictionaries were subjected to validation through manual checks of word context to minimise false positives, with over 1,700 specific word replacements implemented to ensure accuracy.

For details on the dictionaries and validation process, see Chapter 1.

4 Baseline Analysis: The Personality Traits Wage Penalty

This section establishes the baseline relationship between personality trait mentions in job advertisements and offered wages. I first present the empirical methodology for

[2] Marinescu and Wolthoff (2020) show that job titles usually contain relevant hierarchical, experience and specialisation information which can be omitted in the body of the advertisement. Both were merged for the text analysis retrieval.

identifying this relationship, then document the evidence of the wage penalty across different personality traits.

4.1 Empirical Methodology

To identify the relationship between personality traits and wages, I employ a high-dimensional fixed-effects (HDFE) model that controls for multiple sources of unobserved heterogeneity across job characteristics, occupations, and firms. This approach isolates the relationship between personality traits and wages while controlling for a large set of observable and fixed unobservable factors. The baseline model is:

$$\ln(wage_i) = \beta_0 + \sum_{t=1}^{5} \beta_t Trait_{ti} + \beta_6 X_i + \delta_l + \theta_m + \epsilon_i \tag{1}$$

where $\ln(wage_i)$ is the log of the posted wage for vacancy i, $Trait_{ti}$ represents binary indicators for each of the five personality traits (Conscientiousness, Openness, Extraversion, Agreeableness, and Emotional Stability), and X_i is a vector of job posting characteristics including education requirements, experience, seniority, contract type, and other job attributes. The terms δ_l and θ_m capture location at two digits of NUTS and month-year fixed effects, respectively.

I progressively augment this specification by adding industry fixed effects (α_s), occupation fixed effects (γ_k), and finally firm fixed effects (ϕ_j):

$$\ln(wage_i) = \beta_0 + \sum_{t=1}^{5} \beta_t Trait_{ti} + \beta_6 X_i + \delta_l + \theta_m + \alpha_s + \gamma_k + \phi_j + \epsilon_i \tag{2}$$

This stepwise approach allows me to assess how much of the observed wage penalty for personality traits can be explained by within-industry, within-occupation, and within-firm variation. If the trait coefficients (β_t) remain overall negative after including these fixed effects, this suggests the wage penalty persists even within narrowly defined industries, occupations, and firms, pointing to a relationship that requires further investigation.

To further explore how the trait-wage relationship varies across different occupations, I also estimate the baseline specification separately for each occupation at 1-digit, 2-digit, and 3-digit SOC levels, excluding the occupation fixed effects from these models.

4.2 Results

4.2.1 Evidence of the Wage Penalty

Table 2 shows the relationship between personality trait mentions and posted wages. Panel A presents the effect of mentioning any personality trait, while Panel B shows the

coefficients for each of the Big Five traits separately. The first column in both panels gives the raw relationship without controls, showing a wage penalty of 17.6 per cent for job advertisements that mention any personality trait. Looking at individual traits in Panel B, all traits except Openness have large wage penalties in the raw data. Agreeableness has the largest penalty (21.4 per cent), followed by Extraversion (18.5 per cent), Emotional Stability (16.8 per cent), and Conscientiousness (6.7 per cent). Openness-to-experience has a small positive association with wages (1.6 per cent). On average, jobs mentioning personality traits offer lower wages than those that do not.

Table 2: Log posted wage and Personality Traits

	(1) Baseline	(2) +Controls	(3) +Month	(4) +Location	(5) +Industry	(6) +Occupation	(7) +Firm
Panel A: Effect of Any Personality Trait							
Any Personality Trait	-0.176***	-0.157***	-0.157***	-0.151***	-0.126***	-0.069***	-0.045***
	(0.000)	(0.000)	(0.000)	(0.005)	(0.009)	(0.010)	(0.007)
R^2	0.024	0.187	0.191	0.212	0.267	0.450	0.600
Panel B: Effects of Individual Personality Traits							
Conscientiousness	-0.067***	-0.034***	-0.034***	-0.030***	-0.029***	-0.022**	-0.020***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.009)	(0.009)	(0.007)
Openness	0.016***	0.026***	0.027***	0.017***	0.021**	0.005	0.007*
	(0.000)	(0.000)	(0.000)	(0.003)	(0.008)	(0.007)	(0.004)
Extraversion	-0.185***	-0.140***	-0.140***	-0.135***	-0.117***	-0.060***	-0.032***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.007)	(0.009)	(0.007)
Agreeableness	-0.214***	-0.106***	-0.109***	-0.104***	-0.083***	-0.022***	-0.014*
	(0.000)	(0.000)	(0.000)	(0.005)	(0.015)	(0.008)	(0.008)
Emotional Stability	-0.168***	-0.091***	-0.092***	-0.086***	-0.079***	-0.043***	-0.031***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.013)	(0.010)	(0.008)
Observations	11,662,124	11,662,124	11,662,124	9,923,701	9,923,697	9,923,697	9,888,315
R^2	0.040	0.201	0.206	0.226	0.277	0.452	0.601
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	No	No	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	No	Yes	Yes	Yes	Yes
Industry FE	No	No	No	No	Yes	Yes	No
Occupation FE	No	No	No	No	No	Yes	Yes
Firm FE	No	No	No	No	No	No	Yes
N firms	128,559	128,559	128,559	128,559	128,559	128,559	81,843

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a different specification with progressively added controls and fixed effects. All regressions with controls include hard and soft skills, 16 sets of amenities, advertisement length, contract type, work schedule, seniority level, education, experience, bonus availability, and negotiable salary indicators. Standard errors are clustered at the region level in models (4)-(5), and at the occupation and firm level in models (6)-(7).

4.2.2 Controlling for Job Characteristics and Fixed Effects

To examine whether these penalties might be due to other factors, the subsequent columns add controls for job characteristics (column 2), month (column 3), location (column 4), industry (column 5), occupation (column 6), and firm fixed effects (column 7). The covariates include education, experience, seniority, contract types, skill requirements, and 16 categories of amenities that capture workplace benefits and non-wage compensation[3].

[3]Amenities specifications are listed in Appendix B and the rest are thoroughly described in Chapter 1.

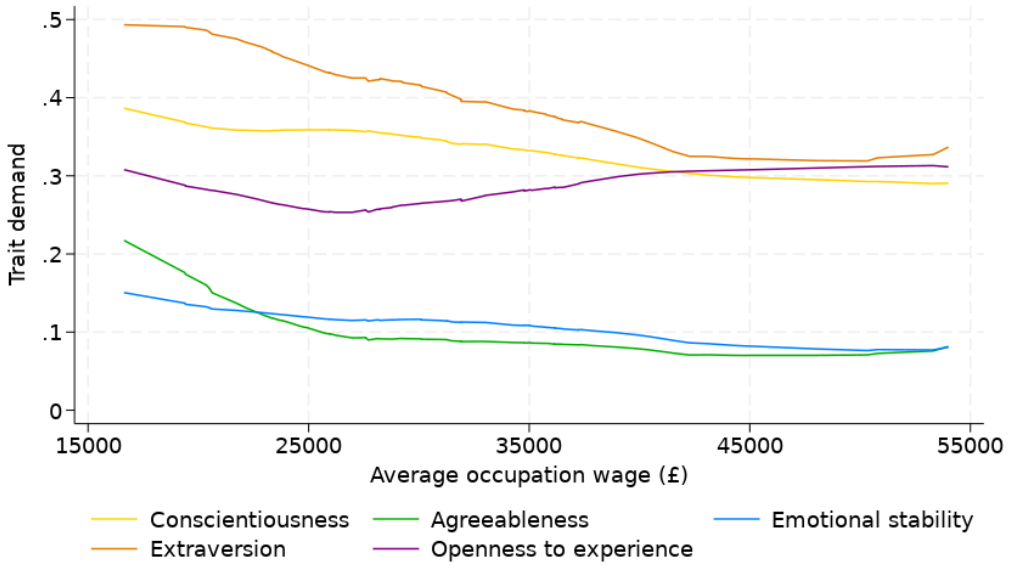
Even with the most extensive set of controls (column 7), the penalty for mentioning any personality trait remains statistically significant at 4.5 per cent. When examining individual traits, Conscientiousness (2.0 per cent), Extraversion (3.2 per cent), Agreeableness (1.4 per cent), and Emotional Stability (3.1 per cent) all have statistically significant negative coefficients. Openness-to-experience shows a small positive coefficient (0.7 per cent), though it is only significant at a 10 per cent confidence level.

The wage penalties decrease when fixed effects are added, especially occupation fixed effects (column 6), which reduce the penalty for any trait from 12.6 to 6.9 per cent. This indicates that part of the observed penalty comes from selection into occupations—personality traits appear more often in advertisements for lower-paying occupations. However, because significant penalties remain even after controlling for occupation and firm characteristics, selection alone cannot fully explain this relationship. The inclusion of job requirements and job amenities controls also reduce the size of the wage penalty in the baseline specification (column 1), but these differences are ultimately captured by the subsequent fixed effects.

To illustrate how trait demand varies across occupation-level average wages, Figure 1 plots, at the occupation level, the proportion of advertisements which mention a trait against the average wage in that occupation. In total, there are 90 occupations at the 3-digit level of the SOC. Extraversion and Conscientiousness show strong negative trends across the wage distribution, with mentions decreasing as average occupation wages increase. For example, Extraversion appears in nearly 50 per cent of advertisements for occupations at the lowest end of the wage distribution but only in about 32 per cent for the highest-paying occupations. Agreeableness and Emotional Stability follow similar but less pronounced trends, with both traits mentioned more often in lower-wage occupations. Openness to experience shows a different trend, with lower mention rates in the middle of the wage distribution and slight increases at both extremes. This pattern aligns with findings from Chapter 1, where Openness was most prevalent in professional and technical occupations, but also in sales and customer service.

These differences across the wage distribution help explain why occupation fixed effects reduce the wage penalties seen in Table 2. Personality traits appear more often in lower-paying occupations, creating a negative correlation between trait mentions and wages in the overall data. However, significant negative coefficients remain even after controlling for occupation, so this selection mechanism does not fully explain the wage penalties.

Figure 1: Personality trait demand across the wage distribution of occupations



Notes: 90 Occupations considered at the 3-digits level of the SOC.

4.2.3 Occupation-Specific Trait Effects

To assess occupational differences, I estimated the baseline specification for each occupation at 1, 2 and 3 digits, excluding the occupation fixed effects from the model. Figure 2 presents the coefficient estimates and 95 per cent confidence intervals for each personality trait across the nine major occupation groups. The results show substantial heterogeneity in how personality traits affect wages across occupation categories.

Extraversion has negative wage effects across almost all occupation types, with especially large penalties in Administrative and Secretarial roles (5.7 per cent), Associate Professional and Technical occupations (4.9 per cent), and Sales and Customer Service roles (3.6 per cent). Conscientiousness also has negative coefficients across most occupation groups, with the largest penalties in Managers, Directors and Senior Officials (3 per cent) and Associate Professional and Technical occupations (3 per cent).

Openness to experience stands out as the only trait with predominantly positive effects. It has positive and significant coefficients in several occupation groups, including Skilled Trades (2.1 per cent), Sales and Customer Service (2.2 per cent), and Caring, Leisure and Other Service occupations (1.8 per cent). This confirms that the small positive effect of Openness in the overall analysis represents positive valuation of this trait in specific occupations.

Agreeableness and Emotional Stability show markedly different effects across occupations. Agreeableness has strong negative effects in Administrative and Secretarial roles (4.7 per cent) and Sales and Customer Service positions (4.3 per cent), but shows a large

positive effect in Elementary occupations (5.5 per cent), although this latter effect has a wide confidence interval. Emotional Stability shows the most pronounced variation, with substantial penalties in Associate Professional and Technical occupations (5.6 per cent) and Managers (3.9 per cent), but no significant penalty in Skilled Trades or Elementary occupations.

These patterns suggest that the wage penalty for personality traits varies with the specific requirements and characteristics of different types of work. The positive effects of Agreeableness in Elementary occupations and the absence of penalties for most traits in Skilled Trades may reflect productivity benefits in these contexts. Conversely, the consistent penalties for Extraversion across high-skill occupations suggest that this trait may serve different signalling functions in professional and managerial roles than in service-oriented positions.

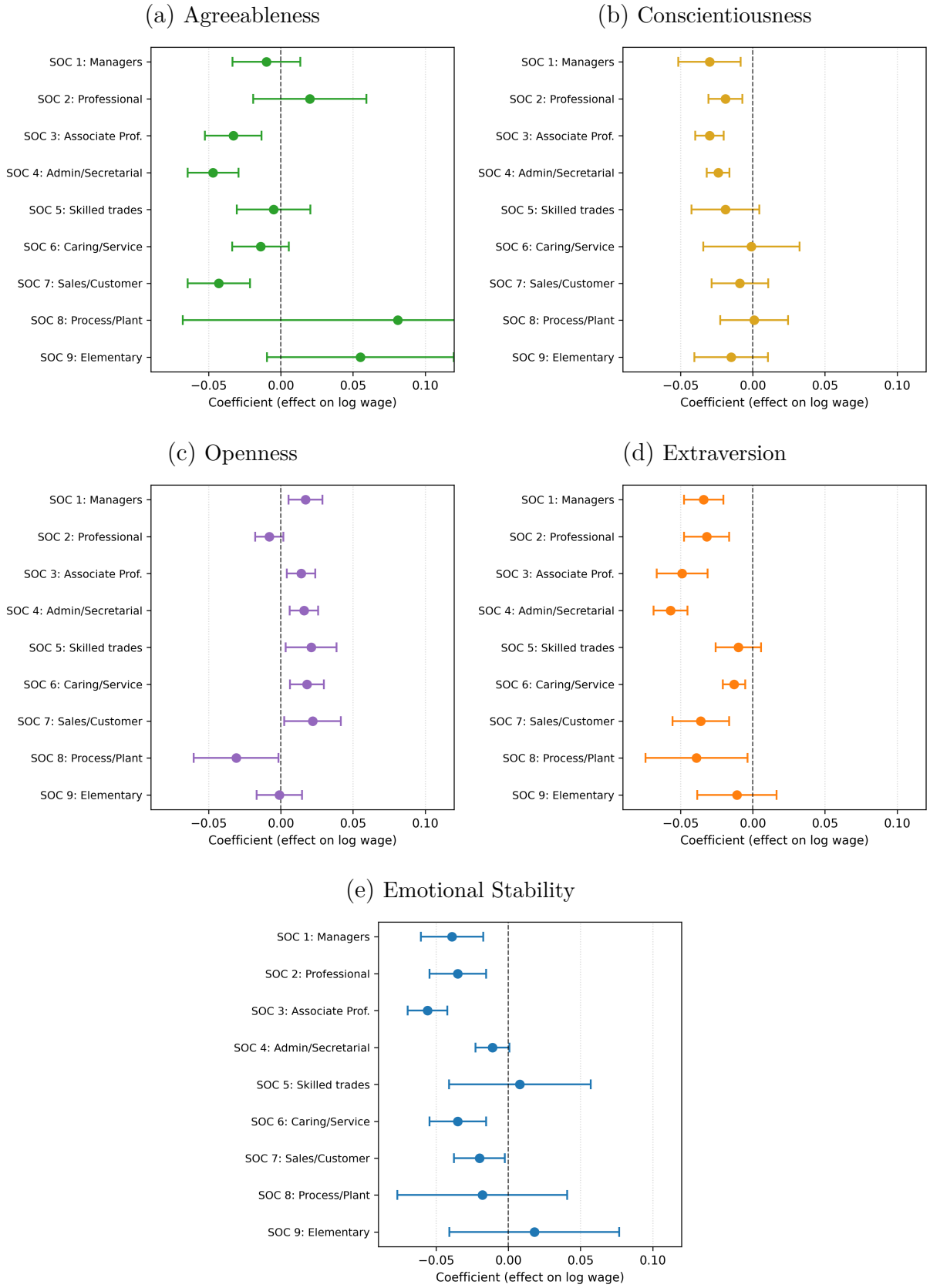
More detailed analyses at the 2-digit and 3-digit SOC level in Table A.2 and Table A.3 in Appendix A show even greater differences, indicating complex relationships between personality traits and wages that cannot be explained by simple occupational sorting. The granular occupation-specific patterns present heterogeneity that strengthens the argument against uniform explanations for trait penalties. At the 2-digit level, Extraversion penalties range from non-significant effects in Teaching and Educational Professionals to substantial penalties of 6.1 per cent in Science, Research, Engineering and Technology Professionals. Similarly, Conscientiousness shows positive effects in Caring Personal Service Occupations (1.7 per cent) but large negative effects in Business, Media and Public Service Professionals (3.5 per cent).

This fine-grained heterogeneity suggests that personality trait language serves fundamentally different functions across occupational contexts, varying not just by broad occupational categories but by specific job roles within them. The variation in trait effects across the occupational hierarchy also provides insight into potential mechanisms beyond those directly tested—the differential effects of traits in entry-level elementary occupations versus senior managerial positions may reflect different stages of career development where personality characteristics are valued differently, relating to theories of employer learning about worker productivity over time.

5 Correlation with potential low-pay characteristics

In the previous subsection, I showed that wage penalties varied between occupations. In this section, I consider whether the variation in wage penalties can be explained by specific occupational characteristics, such as workers' average demography, education level, or income. In addition, I estimate whether wage penalties are moderated or increased when vacancies demand different skills, seniority levels, or types of contracts. Finally, I consider

Figure 2: Personality Trait Coefficients on Log Posted Wage, by 1-digit SOC Occupation



Notes: Each panel shows coefficients from separate regressions for each 1-digit SOC occupation. Points represent coefficient estimates; horizontal bars show 95 per cent confidence intervals. Effects crossing zero are not statistically significant at the 5 per cent level. All regressions include controls for soft skills, hard skills, remote work, advertisement length, contract type, work schedule, seniority level, education, experience, amenities availability, and negotiable salary. Fixed effects for location, month-year, and firm are included. Standard errors are clustered at the firm level. Estimations are reported in Table A.1 in Appendix A.

whether there are different effects by type of firms. The baseline model is extended by incorporating interaction terms between each personality trait and occupation, job, or firm characteristics.

For occupation-level characteristics, I estimate separate regressions for each trait t and each occupation characteristic c in turn, while controlling for the other four traits:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \beta_t Trait_i^t + \sum_{j \neq t}^{5} \beta_j Trait_i^j + \gamma_c Occ\_Char_i^c \\ + \delta_{tc} \left(Trait_i^t \times Occ\_Char_i^c \right) + \boldsymbol{\beta}_X' X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{3}$$

where $Trait_i^j$ represents the other four traits included as controls, and $Occ\_Char_i^c$ represents the standardised value of occupation characteristic c, estimated separately for each $c \in$ {women's share, educational attainment, migrants' share, mean worker age, mean hourly wage, share of managerial positions}. By standardising the occupation characteristics, I interpret the interaction coefficients δ_{tc} as the change in the trait effect associated with a one standard deviation increase in characteristic c.

The results are presented in three components: (1) the effect when the characteristic is at its mean, (2) the change in effect for a one standard deviation increase, and (3) the total effect when the characteristic is one standard deviation above the mean. These regressions control for occupation-level variables while including occupation fixed effects, so the effect accounts for the trait-occupation-characteristic variation after controlling for occupation-specific features.

For job-requirement interactions, I estimate separate regressions for each trait t and each job requirement r in turn:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \beta_t Trait_i^t + \sum_{j \neq t}^{5} \beta_j Trait_i^j + \gamma_r Job\_Req_i^r \\ + \lambda_{tr} \left(Trait_i^t \times Job\_Req_i^r \right) + \boldsymbol{\beta}_X' X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{4}$$

where $Job\_Req_i^r$ is a binary indicator for job requirement r, estimated separately for each $r \in$ {soft skills, hard skills, amenities, negotiable salary, seniority levels, work schedule, contract period, experience, education}. The coefficients λ_{tr} show how the wage effect of trait t differs when job requirement r is present. These regressions include occupation fixed effects (γ_k) and firm fixed effects (ϕ_j).

For firm-level characteristics, I estimate separate regressions for each trait t and each

firm characteristic f in turn:

$$\ln(wage_i) = \beta_0 + \beta_t Trait_i^t + \sum_{j \neq t}^{5} \beta_j Trait_i^j + \gamma_f Firm\_Char_i^f + \mu_{tf}\left(Trait_i^t \times Firm\_Char_i^f\right) + \boldsymbol{\beta}_X' X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \tag{5}$$

where $Firm\_Char_i^f$ represents characteristic f of the firm posting vacancy i, estimated separately for each $f \in \{$size (employee count), age, profit, public listing status$\}$. Continuous firm characteristics are standardised, so the coefficients μ_{tf} measure the change in trait t's wage effect for a one standard deviation increase in f. These regressions include occupation fixed effects (γ_k) and firm fixed effects (ϕ_j).

If the wage penalty for personality traits is primarily driven by correlation with these observable characteristics, we would expect the interaction terms to be significant and the main effects of traits to become smaller or insignificant once these interactions are included. Alternatively, finding persistent main effects alongside significant interaction terms would suggest that while context matters, there is still an underlying relationship between personality trait mentions and wages that cannot be fully explained by these observable factors.

5.1 Results

5.1.1 Occupation-level heterogeneity

To test whether personality trait penalties reflect correlation with other occupation characteristics, I interact each trait with standardised occupation-level variables from the UK Household Longitudinal Study (UKHLS). Table 3 presents these interaction effects, showing how trait penalties vary with occupation characteristics.

The demographic composition of occupations influences trait-wage relationships. In female-dominated occupations, personality trait penalties are generally smaller or even become premiums. For instance, as women's representation increases by one standard deviation, the Conscientiousness penalty (2.1 per cent) becomes statistically insignificant (0.8 per cent), while Openness shifts from a marginally significant effect to a significant positive premium (1.6 per cent). Similarly, the Extraversion penalty is reduced from 3.3 to 2.3 per cent in occupations with higher female representation. The same pattern appears in occupations with higher migrant shares, where penalties for Conscientiousness, Extraversion, and Emotional Stability all diminish. Conscientiousness shows the largest shift, with its penalty (2.0 per cent) becoming statistically insignificant (0.8 per cent) as migrant representation increases by one standard deviation.

In contrast, increases in the human capital associated with an occupation generally

amplifies trait penalties. As the proportion of degree holders increases, the penalty for Extraversion worsens (3.1 to 4.9 per cent), and Openness loses its positive effect, shifting from 0.8 per cent to becoming statistically insignificant (0.2 per cent). Only Agreeableness shows a different pattern, with its penalty remaining relatively stable around 1.6 to 2.3 per cent across education levels, suggesting this trait may operate through different mechanisms. Similarly, in higher-paying occupations, most traits face stronger penalties. This is pronounced for Extraversion, where the penalty increases from 3.2 to 4.6 per cent in occupations with wages one standard deviation above the mean. Emotional Stability also shows amplified penalties, from 3.3 to 4.5 per cent in higher-paying occupations.

Occupational age structure has limited effects on most traits, with only Emotional Stability showing a larger penalty (3.1 to 4.1 per cent) in occupations with older workers. However, managerial content has larger effects on traits. In management-intensive occupations, the Extraversion penalty increases (3.2 to 4.9 per cent), as does the Emotional Stability penalty (3.2 to 4.2 per cent). These results contradict traditional expectations that leadership roles would value traits like Extraversion, suggesting that high-level positions may use personality-traits for different purposes beyond identifying productive characteristics. Similarly, occupations with higher hourly wages tend to penalise all traits but Agreeableness more.

These findings provide partial support for the correlation with low-pay characteristics hypothesis. The smaller penalties in female-dominated and migrant-heavy occupations suggest that demographic composition partly explains trait penalties. However, the amplified penalties in high-skill, high-wage occupations indicate that correlation with observable low-pay characteristics cannot fully account for the wage penalty paradox. The persistence and even intensification of penalties in contexts where these traits should be most valued suggests additional mechanisms beyond simple correlation with occupation characteristics operate in determining the relationship between personality trait mentions and wages.

Table 3: Effects of personality traits with standardised occupation characteristics

Occ. Characteristic		Consc.	Open.	Extra.	Agree.	Emot.
Age (years)	At mean (0 SD)	-0.020***	0.007*	-0.032***	-0.014*	-0.031***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	Per 1 SD increase	-0.005	0.002	-0.003	0.007	-0.010*
		(0.005)	(0.003)	(0.005)	(0.006)	(0.005)
	Total effect at +1 SD	-0.025**	0.009*	-0.035***	-0.006	-0.041***
		(0.010)	(0.006)	(0.009)	(0.012)	(0.010)
Degree holder share	At mean (0 SD)	-0.020***	0.008**	-0.031***	-0.016**	-0.033***
		(0.006)	(0.004)	(0.006)	(0.008)	(0.008)
	Per 1 SD increase	-0.008	-0.010**	-0.018***	-0.007	-0.020***
		(0.006)	(0.004)	(0.006)	(0.009)	(0.007)
	Total effect at +1 SD	-0.029***	-0.002	-0.049***	-0.023**	-0.053***
		(0.005)	(0.005)	(0.008)	(0.011)	(0.010)
Female share	At mean (0 SD)	-0.021***	0.007*	-0.033***	-0.010	-0.028***
		(0.006)	(0.004)	(0.006)	(0.011)	(0.009)
	Per 1 SD increase	0.013**	0.009**	0.010	-0.009	-0.010
		(0.005)	(0.004)	(0.006)	(0.011)	(0.008)
	Total effect at +1 SD	-0.008	0.016***	-0.023***	-0.019**	-0.038***
		(0.008)	(0.005)	(0.009)	(0.007)	(0.007)
Hourly income	At mean (0 SD)	-0.020***	0.008**	-0.032***	-0.013	-0.033***
		(0.006)	(0.004)	(0.006)	(0.008)	(0.008)
	Per 1 SD increase	-0.009*	-0.008*	-0.015***	0.006	-0.012*
		(0.005)	(0.005)	(0.006)	(0.006)	(0.007)
	Total effect at +1 SD	-0.030***	-0.001	-0.046***	-0.007	-0.045***
		(0.005)	(0.006)	(0.008)	(0.011)	(0.012)
Managerial share	At mean (0 SD)	-0.020***	0.007*	-0.032***	-0.014*	-0.032***
		(0.006)	(0.004)	(0.006)	(0.008)	(0.008)
	Per 1 SD increase	-0.009*	-0.001	-0.018***	-0.002	-0.010**
		(0.005)	(0.005)	(0.005)	(0.008)	(0.005)
	Total effect at +1 SD	-0.029***	0.006	-0.049***	-0.016	-0.042***
		(0.005)	(0.007)	(0.007)	(0.012)	(0.009)
Migrant share	At mean (0 SD)	-0.020***	0.007*	-0.032***	-0.015*	-0.032***
		(0.006)	(0.004)	(0.006)	(0.008)	(0.008)
	Per 1 SD increase	0.012**	0.002	0.012**	0.006	0.004
		(0.006)	(0.004)	(0.005)	(0.005)	(0.007)
	Total effect at +1 SD	-0.008	0.009	-0.021**	-0.009	-0.028***
		(0.008)	(0.007)	(0.010)	(0.009)	(0.010)

Notes: Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. "At mean" shows the effect when the standardised occupation characteristic equals zero (at its mean). "Per 1 SD" shows the change in effect for a one standard deviation increase in the occupation characteristic. "Total effect" shows the combined effect when the Regressions control for job requirements, location, time-month, occupation and firm fixed effects. 9,888,315 observations.

5.1.2 Job-requirement heterogeneity

Job-specific attributes also modify how personality traits affect wages. Table 4 presents interaction coefficients showing how trait-wage relationships change when specific job requirements are mentioned, with the baseline representing jobs that do not mention these requirements.

When soft skills are mentioned alongside personality traits, the interaction effects are generally positive, indicating that soft skills reduce trait penalties. The interaction is statistically significant for Extraversion (1.2 per cent) and Agreeableness (1.8 per cent), suggesting these traits are viewed more favourably when soft interpersonal skills are explicitly valued. This aligns with research showing personality traits and soft skills are complementary (Deming, 2017; Heckman and Kautz, 2012).

Conversely, hard skills requirements create negative interactions with personality traits. The interaction coefficients are negative for Openness (1.7 per cent) and Extraversion (3.3 per cent), indicating that technical skill requirements amplify penalties for these traits. This suggests employers may view personality characteristics as less valuable when technical competencies are the primary focus.

Low seniority roles have a significant positive interaction with Conscientiousness (2.1 per cent), suggesting this trait may entail productivity-enhancing demand in junior positions, as Bowles et al. (2001) theory predicts. Similarly, when no experience is needed, Conscientiousness shows a large positive interaction (8.5 per cent), higher than when experience is explicitly required (0.9 per cent). Interestingly, the negative effect of Agreeableness seems to be driven by its interaction with Experience needed (2.7 per cent).

Work schedule requirements create notable interaction effects. Full-time positions show positive interactions with Openness (1.8 per cent) and Extraversion (1.3 per cent). However, part-time positions have even stronger positive interactions with Openness (4.7 per cent) and Agreeableness (5.0 per cent), indicating these may be more valued in non-standard employment arrangements. Results for contract types produce negative effects for temporary positions, with a significant interaction for Emotional Stability (4.2 per cent). Only Openness shows a positive interaction with permanent positions (1.4 per cent).

Negotiable salary arrangements create a positive interaction with Openness (1.4 per cent) but a negative interaction with Agreeableness (3.1 per cent). The latter aligns with individual-level data findings, which argue that individuals with high Agreeableness scores may have lower risk preferences and bargaining power (Flinn et al., 2025).

Education requirements generally show positive interactions, particularly for A-Level requirements, which have significant positive interactions with Conscientiousness (2.0

per cent) and Emotional Stability (4.1 per cent). Bachelor's degree requirements show a positive interaction with Agreeableness (3.1 per cent) and GCSE with Extraversion (2.8 per cent).

These interaction patterns provide mixed evidence for the correlation with low-pay characteristics hypothesis. The positive interactions between personality traits and entry-level positions (low seniority, no experience required) suggest that trait penalties may partly reflect concentration in junior roles that typically offer lower wages. Similarly, the positive interactions with educational requirements indicate that when jobs explicitly require qualifications, personality traits face smaller penalties or even premiums. However, the negative interactions with hard skills and temporary contracts suggest that certain job characteristics amplify rather than explain trait penalties. The heterogeneity across different job contexts indicates that while some trait penalties may reflect correlation with low-paying job characteristics, this mechanism cannot fully account for the wage penalty across all job types.

Table 4: Interaction term coefficients and SE by Job characteristics

		Cons.	Open.	Extra.	Agree.	Emot.
Soft skills	Baseline	-0.027***	0.006	-0.040***	-0.027***	-0.033**
		(0.007)	(0.007)	(0.008)	(0.006)	(0.013)
	Interaction	0.010	0.002	0.012*	0.018**	0.002
		(0.007)	(0.007)	(0.006)	(0.009)	(0.015)
Hard skills	Baseline	-0.016	0.017***	-0.013	-0.004	-0.029**
		(0.013)	(0.005)	(0.010)	(0.013)	(0.013)
	Interaction	-0.006	-0.017***	-0.033***	-0.021	-0.005
		(0.012)	(0.006)	(0.009)	(0.014)	(0.013)
Negotiable	Baseline	-0.026**	0.000	-0.026***	-0.001	-0.025**
		(0.010)	(0.005)	(0.009)	(0.012)	(0.010)
	Interaction	0.012	0.014*	-0.013	-0.031**	-0.013
		(0.009)	(0.008)	(0.008)	(0.013)	(0.013)
Seniority	Baseline	-0.032***	0.004	-0.022**	-0.007	-0.048**
		(0.011)	(0.011)	(0.011)	(0.018)	(0.024)
	I. High	0.066	-0.009	-0.018	0.167	-0.001
		(0.042)	(0.035)	(0.035)	(0.105)	(0.030)
	I. Middle	-0.002	-0.013	-0.019	-0.023	0.010
		(0.013)	(0.018)	(0.012)	(0.018)	(0.024)
	I. Low	0.021*	0.009	-0.011	-0.007	0.025
		(0.010)	(0.010)	(0.012)	(0.020)	(0.025)

Continued on next page

Table 4 – Continued from previous page

		Cons.	Open.	Extra.	Agree.	Emot.
Work schedule	Baseline	-0.021***	-0.003	-0.036***	-0.022***	-0.035***
		(0.007)	(0.005)	(0.007)	(0.008)	(0.008)
	I. Full-time	0.010	0.018***	0.013**	0.005	-0.001
		(0.007)	(0.006)	(0.006)	(0.009)	(0.013)
	I. Part-time	-0.030	0.047**	0.008	0.050*	0.045
		(0.027)	(0.023)	(0.017)	(0.028)	(0.039)
Contract period	Baseline	-0.022**	0.004	-0.031***	-0.008	-0.025***
		(0.009)	(0.005)	(0.009)	(0.010)	(0.009)
	I. Temporary	-0.014	-0.016	-0.012	-0.023	-0.042*
		(0.010)	(0.010)	(0.012)	(0.016)	(0.024)
	I. Permanent	0.009	0.014***	0.002	-0.007	-0.002
		(0.008)	(0.005)	(0.006)	(0.011)	(0.009)
Experience	Baseline	-0.035***	0.001	-0.032***	0.005	-0.042*
		(0.011)	(0.011)	(0.011)	(0.016)	(0.023)
	I. Needed	0.019*	0.009	-0.000	-0.027*	0.013
		(0.010)	(0.013)	(0.009)	(0.014)	(0.024)
	I. Not needed	0.085***	0.015	0.029	-0.006	0.057
		(0.031)	(0.037)	(0.036)	(0.033)	(0.035)
Education	Baseline	-0.022***	0.007*	-0.032***	-0.017*	-0.035***
		(0.007)	(0.004)	(0.007)	(0.009)	(0.008)
	GCSE	0.005	-0.003	0.028*	0.030	0.023
		(0.013)	(0.014)	(0.015)	(0.025)	(0.017)
	A-Level	0.020*	-0.011	0.000	0.025	0.041***
		(0.010)	(0.010)	(0.012)	(0.019)	(0.010)
	Bachelor	0.012	0.003	-0.014	0.031*	0.031
		(0.018)	(0.014)	(0.014)	(0.018)	(0.026)
	Postgraduate	0.003	0.010	0.001	0.005	-0.025
		(0.018)	(0.018)	(0.015)	(0.012)	(0.016)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Regressions control for job requirements, location, time-month, occupation and firm fixed effects. 9,888,315 Observations. *I.* correspond to Interaction terms (Baseline + Interaction).

5.1.3 Correlation with firm characteristics

This section examines how firm characteristics affect the relationship between personality traits and wages. Three mechanisms may explain these effects. First, according to Bowles et al. (2001), larger firms with higher monitoring costs may benefit more from employees with traits like Conscientiousness and Emotional Stability. Second, firm reputation may matter: established firms can rely on their reputation, while smaller or younger firms might use personality language to attract candidates, possibly offering lower wages but better perceived fit (Kristof, 1996; Rosen, 1986). Third, different organisational structures may value certain traits based on their business models (Nordman et al., 2019).

To test these mechanisms, I matched Adzuna data with the FAME database to get information on firm age, size, profit, and public listing status. Table 5 shows how each firm characteristic affects the relationship between personality traits and wages. These regressions include firm fixed effects, meaning the coefficients capture how trait-firm characteristic interactions vary within firms after controlling for firm-specific factors.

Firm age shows significant effects only for Extraversion, where penalties increase from 3.0 per cent at mean firm age to 4.7 per cent in older firms. This suggests that established firms may use Extraversion-related language more selectively or for different signalling purposes than younger firms.

Firm size and profitability show no statistically significant interaction effects for any personality trait. This contradicts monitoring cost theories that predict larger firms should value traits like Conscientiousness and Emotional Stability more highly due to greater supervision challenges. The absence of size effects also indicates that firm scale does not explain trait penalties through correlation with lower or higher wages.

Public listing status has significant effects for Agreeableness, where publicly quoted firms show larger penalties (7.5 per cent total effect) compared to other private firms (1.3 per cent). The effect is also larger for the rest of the traits when publicly quoted. This represents the strongest firm-level interaction observed, suggesting that organisational governance structures may influence how interpersonal traits are valued in compensation decisions. However, if we consider publicly quoted firms as more widely known, the results contradict the argument that firms with little reputation use traits to attract individuals while paying lower wages.

These results provide limited support for firm-based explanations of personality trait wage penalties within the correlation with low-pay characteristics hypothesis. The absence of significant interactions for firm size and profitability indicates that trait penalties do not simply reflect concentration in smaller, less profitable firms. The monitoring cost mechanism receives no empirical support, as larger firms do not show reduced penalties

Table 5: Effects of personality traits with standardised firm characteristics

Firm Characteristic		Consc.	Open.	Extra.	Agree.	Emot.
Firm Age (years)	At mean (0 SD)	-0.024**	0.009*	-0.030***	-0.006	-0.017
		(0.011)	(0.005)	(0.010)	(0.014)	(0.012)
	Per 1 SD increase	-0.001	0.002	-0.017***	-0.013	0.004
		(0.009)	(0.005)	(0.006)	(0.011)	(0.007)
	Total effect at +1 SD	-0.026*	0.011	-0.047***	-0.019	-0.014
		(0.015)	(0.008)	(0.011)	(0.018)	(0.009)
Firm Size (employees)	At mean (0 SD)	-0.024**	0.009	-0.031***	-0.002	-0.013
		(0.011)	(0.006)	(0.010)	(0.014)	(0.011)
	Per 1 SD increase	-0.006	0.001	0.008	0.050	0.018
		(0.008)	(0.003)	(0.009)	(0.033)	(0.018)
	Total effect at +1 SD	-0.029*	0.009	-0.023	0.048	0.005
		(0.016)	(0.007)	(0.016)	(0.043)	(0.026)
Profit (£)	At mean (0 SD)	-0.016**	0.002	-0.029***	-0.011	-0.015
		(0.008)	(0.006)	(0.010)	(0.010)	(0.010)
	Per 1 SD increase	0.000	0.005	0.001	0.014	0.004
		(0.007)	(0.005)	(0.008)	(0.012)	(0.010)
	Total effect at +1 SD	-0.016	0.007	-0.028**	0.003	-0.012
		(0.011)	(0.008)	(0.011)	(0.018)	(0.014)
Publicly Quoted	At mean (0 SD)	-0.020***	0.006	-0.031***	-0.013	-0.031***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	When publicly quoted = 1	-0.007	0.026	-0.010	-0.062***	-0.007
		(0.015)	(0.032)	(0.018)	(0.018)	(0.021)
	Total effect when quoted	-0.027*	0.032	-0.042**	-0.075***	-0.038*
		(0.014)	(0.032)	(0.018)	(0.017)	(0.021)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. "At mean" shows the effect when the standardised firm characteristic equals zero (at its mean). "Per 1 SD increase" shows the combined effect when the firm characteristic is one standard deviation above its mean. Regressions control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 3,619,726 observations for Firm Age estimations, 1,897,228 observations for Profit, 3,009,216 observations for Firm Size and 9,888,315 observations for Publicly Quoted.

for traits that should facilitate supervision. Only public listing status shows meaningful interaction effects, and these amplify rather than explain trait penalties. Overall, these findings suggest that firm characteristics cannot account for the personality trait wage penalty through correlation with low-pay firm types, indicating that other mechanisms must explain this relationship.

5.1.4 Assessment of the Low-Pay Correlation Hypothesis

The correlation with low-pay characteristics hypothesis receives partial support but cannot fully explain the personality trait wage penalty. Three key findings emerge.

First, demographic composition at the occupation level matters. Personality traits have smaller penalties or even premiums in occupations with higher proportions of women and migrants, suggesting that some observed trait penalties reflect sorting into lower-paying occupational contexts. However, trait penalties generally increase with occupation education level, wage level, and managerial content—precisely the contexts where human

capital theory would predict stronger returns to productivity-enhancing traits.

Second, job-level characteristics provide mixed evidence. The positive interactions between personality traits and entry-level positions (low seniority, no experience required) suggest that trait penalties partly reflect concentration in junior roles. However, some job characteristics amplify rather than explain trait penalties, particularly hard skills and temporary contracts.

Third, firm-level characteristics provide limited explanatory power. The absence of significant interactions for firm size and profitability indicates that trait penalties do not simply reflect concentration in smaller, less profitable firms. Only public listing status shows meaningful effects, and these amplify rather than explain trait penalties.

Overall, while demographic composition and career stage characteristics provide some explanatory power, the persistence and amplification of penalties in high-skill, high-wage contexts indicates that additional mechanisms beyond correlation with observable low-pay characteristics must account for this relationship.

6 Compensating Differentials

A second explanation proposes that employers use personality-traits to signal non-pay amenities valued by workers, which may result in lower posted wages through a compensating differentials mechanism. According to Rosen (1986)'s equalising differences theory, when jobs offer attractive non-wage characteristics, employers can offer lower wages while still attracting qualified candidates. Personality-traits in job advertisements may signal a workplace where employees will experience better person-organisation fit, which candidates may value alongside monetary compensation.

Firms benefit from attracting workers whose personalities align with their organisational identity (Kristof, 1996; Ployhart et al., 2006; Roberson et al., 2005). This alignment can enhance job satisfaction, organisational commitment, team cohesion, and reduce turnover (Judge and Cable, 1997; Krueger and Schkade, 2008). By signalling these benefits through personality-related language, employers may effectively lower their wage offers while maintaining their ability to attract suitable candidates.

Building on signalling theory (Spence, 1973), research shows that job-seekers evaluate not just compensation but also fit with organisational culture when making application decisions (Breaugh, 2013; Chapman et al., 2005; Stevens and Szmerekovsky, 2010). This is supported by person-organisation fit theory, which suggests applicants are more likely to pursue positions aligned with their values (Judge and Cable, 1997; Kristof, 1996; Roberson et al., 2005). In this framework, job seekers may prioritise intrinsic rewards over purely monetary compensation (Akerlof and Kranton, 2005; Cassar and Meier, 2018).

This compensating differentials approach parallels Handy and Katz (1998)'s model of non-profit compensation, where firms offer lower monetary compensation but attract workers who value the organisation's mission. Indeed, Burbano (2016) found that information about an employer's social responsibility reduces candidates' wage requirements, providing direct evidence for how non-monetary aspects can allow firms to offer lower wages.

Empirical studies confirm that workers are attracted to organisations that align with their personality traits. Stevens and Szmerekovsky (2010) demonstrated that students with high extraversion scores showed greater interest in job advertisements seeking outgoing candidates. Similarly, Van Hoye and Turban (2015) found extraverted applicants were more attracted to recruiters displaying similar traits, while Judge and Cable (1997) observed that conscientious individuals preferred organisations with achievement-oriented cultures. Conversely, candidates who perceive a poor fit with an organisation's identity are less likely to apply (Breaugh, 2013; Chapman et al., 2005; Wille and Derous, 2018).

6.1 Empirical Methodology

To test whether personality-traits serves as a signal for job quality and person-organisation fit, I examine the relationship between personality traits, advertised amenities, and wages. The analysis proceeds in three tests: first examining correlations between traits and amenities, then testing whether amenity controls explain trait penalties, and finally investigating interaction effects.

6.1.1 Identifying Advertised Amenities

The identification of job amenities builds on the established methodology developed by Escudero et al. (2024) and Boschetti Adamczyk et al. (2025), which provides a taxonomical foundation validated across multiple labour market contexts. Their approach initially drew from Maestas et al. (2023)'s nine job attributes and was expanded using Sockin (2022)'s categorisation of 48 non-wage amenities derived through topic-modelling machine learning algorithms.

I adapted this framework to suit the UK labour market context and the Adzuna dataset characteristics. Following Escudero et al. (2024)'s approach, I reorganised and expanded their keyword dictionaries by conducting extensive context analysis of the terms they used, adding synonyms where appropriate, and implementing over 600 specific word replacements to eliminate false positives identified through manual verification of randomly sampled advertisements. This process involved examining multiple instances of each keyword to ensure it genuinely referred to the intended amenity rather than unrelated concepts that might share similar terminology.

The final taxonomy incorporates several modifications to better capture the UK labour market context. I excluded the physical effort and pace of work category used in the original frameworks, as preliminary analysis suggested they may not be considered as positive amenities. I split the original flexible work arrangements category into two distinct measures: flexible work arrangements (covering flexible hours, part-time options, and work-life balance) and remote work possibilities (capturing teleworking, home-based work, and hybrid arrangements), reflecting the increased prominence and distinct nature of remote work following the COVID-19 pandemic. Additionally, I introduced a general benefits category to capture miscellaneous perks and benefits that did not fit clearly into the established categories but were frequently mentioned in UK job advertisements. This adapted framework resulted in 16 distinct amenity categories, each validated through the rule-based natural language processing approach described in the source methodology, with binary indicators created for each category based on the presence of relevant keywords in job descriptions.

For each category, I expand Escudero et al. (2024)'s keyword lists through contextual

analysis, examining 100 randomly selected sentences containing each potential keyword to verify usage and identify synonyms. The final dictionaries contain between 6 keywords (overtime paid) and 145 keywords (food subsidies and discounts) per category. A full list of keywords and replacements done to the advertisements to facilitate amenities retrieval is available in Appendix B

Table 6 provides definitions of each amenity category, organised into four broad groups reflecting different aspects of compensation and working conditions.

Table 6: Categorisation of amenities and definitions

Panel (a): Variable Earnings	
Subcategory	**Definition**
A01–General benefits	Encompasses reward schemes, flexible benefit platforms, comprehensive benefit packages, staff benefits, recognition schemes, referral programmes, tax benefits, relocation support, and seasonal benefits offered by employers.
A02–Bonuses and commissions	Encompasses various forms of financial incentives and rewards aimed at motivating and compensating employees based on their performance, achievements, or specific goals within an organisation, including equity compensation, profit sharing, performance pay, sales compensation, and premium pay arrangements.
Panel (b): Fringe Benefits	
A03–Overtime paid	Encompasses aspects related to flexible earnings, reflecting the compensation and conditions associated with working beyond regular hours, including extra shifts, double pay, and unsocial hour premiums within an employment arrangement.
A04–Paid time off	Reflects provisions for employees to take time away from work while receiving compensation in specific circumstances, including vacation and holidays, sick leave, medical leave, maternity and paternity benefits, family leave, personal time, special leave, volunteer time, and wellbeing time arrangements.
A05–Health insurance	Includes provisions offered by employers to support employees' healthcare needs, ensuring access to medical services and providing financial protection, including private health insurance, dental and vision coverage, life insurance, accident insurance, free healthcare, health savings plans, wellness programmes, childcare support, eye care, and specialized coverage.
A06–Retirement contributions	Encompasses provisions designed to assist employees in saving for their retirement and securing financial stability during their later years, including pension schemes, employer contribution matching, savings accounts, and life benefits.
A07–Subsidies and discounts	Encompasses benefits related to food, housing, transportation, and various subsidies or discounts offered to employees, including meal provisions, drinks and beverages, fruit and snacks, gym and fitness benefits, transport benefits, discount schemes, clothing and uniform provision, accommodation support, vouchers and rewards, wellness perks, vehicle benefits, platform benefits, leisure entertainment, and insurance extras.

Continued on next page

Table 6 – continued from previous page

Subcategory	Definition
Panel (c): Non-Wage Job Attributes	
A08–Office space and amenities	Covers workplace-related benefits, including facilities and amenities provided by the employer such as fitness facilities, recreational spaces, dining facilities, modern office environments, specialised rooms, work facilities, outdoor spaces, service facilities, and storage amenities.
A09–Location and commuting	Focuses on factors related to the workplace's geographical location and how employees commute to and from work, including location quality, city center access, accessibility, commuting arrangements, parking facilities, public transport links, walking distance convenience, cycling facilities, local transport options, and nearby amenities.
A10–Work equipment and allowances	Sheds light on how the organisation assists employees in ensuring they have the necessary tools and technology, including remote work support, mobile devices, computing equipment, company vehicles, vehicle expenses, company cards, clothing and protective equipment, and general work equipment provision.
Panel (d): Working Conditions	
A11–Work schedule flexibility	Includes various aspects related to the flexibility of work schedules and arrangements, such as flexible shifts and hours, options for better work-life balance, and modern working practices that support family-friendly policies and employee autonomy over working time.
A12–Workplace safety and ergonomics	Pertains to all aspects related to ensuring a safe working environment for employees, including workplace safety culture, security measures, health and safety compliance, safety regulations and policies, and ergonomic office arrangements that prioritise employee wellbeing.
A13–Job security	Encompasses all aspects related to ensuring job security, stability, and financial protection for employees, including company stability indicators, severance benefits, unemployment support, and income protection measures in various employment scenarios.
A14–Remote work opportunities	Encompasses various arrangements that allow employees to work outside traditional office settings, including fully remote positions, home-based work, hybrid arrangements, telecommuting, and virtual job opportunities that became particularly prominent following the COVID-19 pandemic and changing work patterns.
A15–Work environment and impact on society	Provides insights into the organization's commitment to creating a positive workplace environment and contributing positively to the community and society, including workplace culture, work satisfaction, company values, social activities, company reputation, employee support, work quality, impact and purpose, diversity and inclusion, and team environment characteristics.
A16–Human capital development	Assesses the opportunities for personal and professional growth and development within the organisation, including career growth prospects, skill development opportunities, learning programmes, company growth potential, training and education provision, promotion and advancement pathways, merit-based progression, and personal growth support.

6.1.2 Empirical Models

The analysis employs three complementary approaches to test the compensating differentials hypothesis.

Test 1: Trait-Amenity Correlations

First, I examine whether amenities predict personality trait mentions, which would suggest these characteristics cluster together in job advertisements. For each personality trait t, I estimate:

$$Trait_{ti} = \beta_0 + \sum_{a=1}^{16} \beta_a Amenity_{ai} + \sum_{j \neq t}^{5} \gamma_j Trait_{ji} + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \quad (6)$$

where $Trait_{ti}$ is a binary indicator for whether advertisement i mentions personality trait t, $Amenity_{ai}$ represents the 16 amenity categories, and $\sum_{j \neq t}^{5} \gamma_j Trait_{ji}$ controls for the other four personality traits. Positive coefficients β_a indicate that advertisements mentioning amenity a are more likely to also mention trait t. If employers use personality language alongside amenities to signal attractive job characteristics, we would expect positive correlations between traits and amenities that facilitate compensating differential mechanisms.

Test 2: Amenity Controls and Trait Penalties

Second, I test whether controlling for amenities reduces personality trait wage penalties. I compare two specifications:

Model without amenity controls:

$$\ln(wage_i) = \beta_0 + \sum_{t=1}^{5} \beta_t Trait_{ti} + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \quad (7)$$

Model with amenity controls:

$$\ln(wage_i) = \beta_0 + \sum_{t=1}^{5} \beta_t^a Trait_{ti} + \sum_{a=1}^{16} \lambda_a Amenity_{ai} + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \quad (8)$$

If personality traits serve as proxies for non-monetary compensation, we would expect $|\beta_t^a| < |\beta_t|$, indicating that amenity controls explain part of the trait penalty through systematic correlation with amenity provision.

Test 3: Trait-Amenity Interactions

Third, I examine whether personality traits and amenities interact in determining wages, testing whether employers strategically combine trait requirements with specific amenity packages. This approach moves beyond testing simple correlation between traits and amenities (captured in Tests 1–2) to examine whether the wage effect of trait mentions

depends on the presence of specific amenities.

For each trait t and amenity category a, I estimate:

$$\ln(wage_i) = \beta_0 + \beta_t Trait_{ti} + \lambda_a Amenity_{ai} + \eta_{ta}(Trait_{ti} \times Amenity_{ai}) + \sum_{b \neq a}^{16} \lambda_b Amenity_{bi} + \sum_{j \neq t}^{5} \gamma_j Trait_{ji} + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \quad (9)$$

The interaction coefficient η_{ta} captures how the wage effect of trait t changes when amenity a is present. The economic interpretation depends on both the amenity main effect (λ_a) and the interaction effect (η_{ta}). Under pure compensating differentials theory, amenities should reduce wages ($\lambda_a < 0$) as employers offer lower monetary compensation when providing valuable non-wage benefits. However, some amenities may command wage premiums ($\lambda_a > 0$) due to productivity effects, selection requirements, or signalling functions. For instance, Escudero et al. (2024) document that bonuses and commissions are associated with lower posted wages, as is schedule flexibility, while working in teams and human capital development opportunities show positive wage associations.

When $\lambda_a < 0$ (compensating differentials), negative interactions ($\eta_{ta} < 0$) suggest trait requirements amplify the compensating differential effect, whilst positive interactions ($\eta_{ta} > 0$) suggest traits reduce or reverse the compensating differential mechanism. When $\lambda_a > 0$ (amenity premiums), negative interactions ($\eta_{ta} < 0$) suggest trait requirements eliminate the amenity premium, possibly indicating employers use trait language to justify lower wages despite offering attractive amenities.

The total wage effect when both trait and amenity are present equals $\beta_t + \lambda_a + \eta_{ta}$. If the baseline trait effect β_t becomes statistically insignificant whilst the interaction η_{ta} remains significant, this suggests trait language primarily functions as part of strategic compensation packages rather than reflecting standalone productivity requirements.

6.2 Results

6.2.1 Trait-Amenity Correlations

To test whether employers use personality trait language alongside amenity provision to signal attractive job characteristics, I first examine whether amenity mentions predict personality traits in job advertisements. Table 7 presents results from linear probability models where each personality trait serves as the dependent variable, with all 16 amenity categories as predictors.

Table 7: Amenities and personality trait demand

Amenity Category	Consc.	Open.	Extra.	Agree.	Emot.
A01 General Benefits	0.001	0.001	0.020*	0.018	0.005
	(0.013)	(0.009)	(0.011)	(0.011)	(0.009)
A02 Bonuses and Comm.	-0.006	-0.017**	0.015	-0.001	0.006
	(0.012)	(0.007)	(0.010)	(0.008)	(0.007)
A03 Overtime	-0.021**	-0.001	-0.014*	0.007	0.007
	(0.009)	(0.018)	(0.008)	(0.011)	(0.014)
A04 Paid Time Off	0.006	0.021**	-0.006	0.002	-0.005
	(0.009)	(0.009)	(0.011)	(0.006)	(0.007)
A05 Health Insurance	0.003	-0.000	-0.015	0.011	-0.004
	(0.011)	(0.012)	(0.013)	(0.008)	(0.011)
A06 Retirement	-0.030***	-0.015	0.001	-0.005	0.009
	(0.008)	(0.009)	(0.014)	(0.009)	(0.009)
A07 Subsidies and discounts	-0.002	-0.008	0.037***	-0.014	-0.023**
	(0.014)	(0.011)	(0.012)	(0.009)	(0.010)
A08 Office Space	-0.006	0.003	0.016*	-0.002	-0.016**
	(0.015)	(0.013)	(0.008)	(0.007)	(0.007)
A09 Location Amenities	-0.003	-0.006	0.016*	-0.005	-0.001
	(0.007)	(0.008)	(0.008)	(0.005)	(0.007)
A10 Work Equipment	0.018*	-0.025**	-0.000	-0.006	0.006
	(0.010)	(0.012)	(0.013)	(0.007)	(0.012)
A11 Work Schedule Flex	0.027	0.053***	0.023	-0.021	-0.032**
	(0.018)	(0.011)	(0.016)	(0.013)	(0.014)
A12 Workplace Safety	0.037***	-0.001	-0.058***	-0.001	0.018
	(0.011)	(0.011)	(0.013)	(0.010)	(0.012)
A13 Job Security	0.023	-0.028	0.023	0.021	0.019
	(0.036)	(0.024)	(0.031)	(0.023)	(0.023)
A14 Remote Work	-0.015	0.006	-0.011	-0.005	-0.002
	(0.014)	(0.016)	(0.012)	(0.006)	(0.005)
A15 Work Environment	0.007	0.025***	0.038***	0.021***	-0.003
	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)
A16 Human Capital Dev	0.019**	-0.004	0.036***	-0.006	-0.010
	(0.008)	(0.005)	(0.007)	(0.004)	(0.006)

Notes: Net coefficients show coefficient β_a from Eqn (6) of each amenity category on trait demand, controlling for all other amenities. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Regressions control for other traits, job requirements, location, time-month, occupation and firm fixed effects. Number of observations 9,888,315.

The results show selective clustering between personality traits and amenities that varies across trait types. Extraversion has the strongest positive associations with ameni-

ties, showing significant correlations with food subsidies and discounts (3.7 per cent), work environment features (3.8 per cent), human capital development (3.6 per cent), and location amenities (1.6 per cent). This clustering suggests that employers seeking extraverted workers bundle these positions with workplace amenities, which could facilitate compensating differential mechanisms.

Conscientiousness displays contradictory associations: positive correlations with workplace safety (3.7 per cent), work equipment (1.8 per cent), and human capital development (1.9 per cent), but negative correlations with overtime pay (2.1 per cent) and retirement benefits (3.0 per cent). These opposing patterns suggest that Conscientiousness requirements appear in different job contexts—some emphasising safety and development, others involving shift work with limited traditional benefits.

Emotional Stability shows predominantly negative associations with amenities, correlating negatively with food subsidies (2.3 per cent), office space (1.6 per cent), and work schedule flexibility (3.2 per cent). Openness has selective positive correlations with amenities that enhance work flexibility: paid time off (2.1 per cent), work environment (2.5 per cent), and flexible work schedules (5.3 per cent), but negative associations with bonuses and commissions (1.7 per cent) and work equipment (2.5 per cent). Agreeableness shows the most limited amenity associations, with only work environment showing a significant positive correlation (2.1 per cent).

These patterns indicate that trait-amenity clustering is trait-specific rather than uniform, suggesting that compensating differentials mechanisms, if they operate, must function differently across personality characteristics and job contexts.

6.2.2 Amenity Controls and Trait Wage Penalties

The second test examines whether controlling for amenities as additive terms reduces personality trait wage penalties, which would indicate that part of the penalty operates through systematic correlation with amenity provision. Table 8 compares trait coefficients with and without amenity controls in the full specification including all fixed effects.

The results provide limited evidence that amenities explain trait wage penalties through additive mechanisms. Controlling for amenities has minimal impact on trait coefficients: Agreeableness shows identical penalties (1.4 per cent), Conscientiousness penalties decrease marginally from 2.1 to 2.0 per cent, Emotional Stability penalties remain stable (3.0 to 3.1 per cent), Extraversion penalties decrease slightly from 3.4 to 3.2 per cent, and Openness effects remain small and positive (0.6 to 0.7 per cent). The R^2 increases minimally from 0.598 to 0.601, indicating that amenities explain very little additional variation in wages.

This stability suggests that the baseline wage penalties for personality traits do not

Table 8: Personality trait wage effects: with vs without amenities controls

Personality Trait	Without Amenities	With Amenities
Agreeableness	-0.014*	-0.014*
	(0.008)	(0.008)
Conscientiousness	-0.021***	-0.020***
	(0.007)	(0.007)
Emotional Stability	-0.030***	-0.031***
	(0.008)	(0.008)
Extraversion	-0.034***	-0.032***
	(0.008)	(0.007)
Openness	0.006	0.007*
	(0.004)	(0.004)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Both specifications include all personality traits together, plus job requirements, location, time-month, occupation and firm fixed effects. Without amenities: N= 9,888,315, R^2=0.598. With amenities: N= 9,888,315, R^2=0.601.

primarily operate through employers offering different amenity packages in trait-demanding jobs. The absence of additive effects indicates that compensating differentials, if they exist, must operate through interactive rather than additive mechanisms.

6.2.3 Interaction between Amenities and Traits

The interaction analysis tests whether employers strategically combine personality trait requirements with specific amenity packages, examining whether the wage effect of trait mentions varies depending on the compensation context. Table 9 presents both the baseline effects (when amenities are absent) and the total effects (baseline plus interaction, when amenities are present).

The analysis reveals evidence for context-dependent employer strategies operating through multiple mechanisms beyond pure compensating differentials. Amenities modify trait wage effects, with patterns varying across trait-amenity combinations, indicating that employers use personality trait language differently depending on the broader compensation package offered.

Table 9: Baseline and total wage effects of personality traits by amenity presence

		Consc.	Open.	Extra.	Agree.	Emot. Stab.
A01 General benefits	Baseline	-0.024***	0.004	-0.036***	-0.023***	-0.043***
		(0.007)	(0.006)	(0.007)	(0.008)	(0.009)
	Total	-0.013	0.014	-0.022**	-0.002	-0.013
		(0.011)	(0.009)	(0.010)	(0.012)	(0.015)

Continued on next page

Table 9 – continued from previous page

		Consc.	Open.	Extra.	Agree.	Emot. Stab.
A02 Bonuses and commissions	Baseline	-0.021**	0.003	-0.032***	-0.011	-0.026***
		(0.008)	(0.005)	(0.007)	(0.009)	(0.007)
	Total	-0.018**	0.018**	-0.030***	-0.023**	-0.044***
		(0.008)	(0.008)	(0.010)	(0.011)	(0.016)
A03 Overtime	Baseline	-0.019***	0.008*	-0.035***	-0.019***	-0.037***
		(0.006)	(0.004)	(0.007)	(0.006)	(0.006)
	Total	-0.051	-0.002	0.041	0.063	0.060
		(0.042)	(0.016)	(0.037)	(0.067)	(0.059)
A04 Paid time off	Baseline	-0.014*	0.006	-0.037***	-0.018**	-0.027***
		(0.007)	(0.005)	(0.007)	(0.007)	(0.005)
	Total	-0.037***	0.012	-0.016	-0.006	-0.041*
		(0.014)	(0.008)	(0.012)	(0.016)	(0.021)
A05 Health insurance	Baseline	-0.018***	0.005	-0.038***	-0.017***	-0.032***
		(0.006)	(0.005)	(0.006)	(0.006)	(0.007)
	Total	-0.034	0.019*	0.005	-0.003	-0.029
		(0.021)	(0.010)	(0.020)	(0.027)	(0.029)
A06 Retirement contributions	Baseline	-0.020***	0.003	-0.039***	-0.020***	-0.029***
		(0.007)	(0.005)	(0.006)	(0.007)	(0.008)
	Total	-0.021*	0.014**	-0.017	-0.005	-0.035**
		(0.012)	(0.005)	(0.011)	(0.015)	(0.016)
A07 Subsidies and discounts	Baseline	-0.020***	0.005	-0.038***	-0.022***	-0.032***
		(0.006)	(0.004)	(0.007)	(0.007)	(0.005)
	Total	-0.021	0.015*	-0.007	0.008	-0.029
		(0.016)	(0.008)	(0.015)	(0.020)	(0.022)
A08 Office space	Baseline	-0.021***	0.008*	-0.032***	-0.014*	-0.032***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	Total	0.003	-0.005	-0.016	-0.024	-0.011
		(0.011)	(0.015)	(0.012)	(0.023)	(0.012)
A09 Location and commuting	Baseline	-0.024***	0.006	-0.034***	-0.017*	-0.032***
		(0.007)	(0.005)	(0.007)	(0.009)	(0.009)
	Total	0.004	0.012	-0.020**	-0.001	-0.030***
		(0.009)	(0.008)	(0.010)	(0.014)	(0.010)
A10 Work equipment	Baseline	-0.016**	0.005	-0.034***	-0.023***	-0.028***
		(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
	Total	-0.053*	0.024**	-0.017	0.029	-0.047

Continued on next page

Table 9 – continued from previous page

		Consc.	Open.	Extra.	Agree.	Emot. Stab.
		(0.027)	(0.009)	(0.022)	(0.027)	(0.039)
A11 Work schedule flexibility	Baseline	-0.018***	0.011**	-0.033***	-0.014*	-0.032***
		(0.007)	(0.005)	(0.007)	(0.008)	(0.007)
	Total	-0.035*	-0.013	-0.025*	-0.016	-0.027*
		(0.018)	(0.013)	(0.015)	(0.012)	(0.014)
A12 Workplace safety	Baseline	-0.022***	0.007	-0.033***	-0.012	-0.031***
		(0.008)	(0.005)	(0.008)	(0.008)	(0.008)
	Total	-0.008	0.008	-0.021**	-0.035***	-0.036***
		(0.008)	(0.010)	(0.010)	(0.013)	(0.008)
A13 Job security	Baseline	-0.021***	0.007	-0.032***	-0.015*	-0.032***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	Total	0.010	0.016	-0.038*	0.030	-0.022
		(0.024)	(0.022)	(0.021)	(0.020)	(0.018)
A14 Remote work	Baseline	-0.020***	0.007	-0.031***	-0.013*	-0.030***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	Total	-0.033***	0.024	-0.052***	-0.045***	-0.057***
		(0.008)	(0.016)	(0.013)	(0.015)	(0.011)
A15 Work environment	Baseline	-0.030***	0.003	-0.045***	-0.039***	-0.046***
		(0.006)	(0.006)	(0.006)	(0.008)	(0.011)
	Total	-0.008	0.012*	-0.015*	0.002	-0.020*
		(0.011)	(0.006)	(0.009)	(0.012)	(0.011)
A16 Human capital dev.	Baseline	-0.028***	0.004	-0.060***	-0.038***	-0.054***
		(0.007)	(0.006)	(0.008)	(0.005)	(0.006)
	Total	-0.017**	0.009*	-0.019**	-0.007	-0.024**
		(0.008)	(0.005)	(0.007)	(0.010)	(0.010)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Baseline effects when amenity is absent vs total effects (baseline + interaction) when amenity is present. Regressions control for other traits, job requirements, location, time-month, occupation and firm fixed effects. $R^2 = 0.601$. Sample Size = 9,888,315

Traditional Benefits and Compensating Differentials

Traditional fringe benefits show patterns most consistent with compensating differentials theory. When health insurance is advertised, most trait penalties are eliminated: Conscientiousness penalties shift from 1.8 per cent to a non-significant 3.4 per cent, Extraversion penalties disappear entirely (from 3.8 per cent to a non-significant 0.5 per cent), and Agreeableness penalties are eliminated (from 1.7 per cent to a non-significant 0.3 per cent). Retirement contributions similarly moderate penalties for most traits.

These patterns suggest that when employers offer traditional benefits packages, personality trait requirements function as part of compensating differentials packages that allow lower monetary compensation while maintaining candidate attraction through enhanced total compensation.

Workplace Culture and Development Amenities

Workplace culture and development amenities show the strongest compensating differentials effects. Work environment mentions reduce penalties across all traits: Conscientiousness penalties decline from 3.0 per cent to a non-significant 0.8 per cent, Agreeableness shifts from a penalty (3.9 per cent) to no significant effect (0.2 per cent), and Emotional Stability penalties are reduced (4.6 to 2.0 per cent). Human capital development opportunities create similar moderating effects.

These findings provide strong support for compensating differentials mechanisms when personality trait language is combined with signals of positive organisational culture and employee development.

Amenities with Wage Premiums and Amplified Trait Penalties

However, several amenities show patterns inconsistent with simple compensating differentials theory. Remote work creates uniformly larger penalties for most traits: Conscientiousness penalties increase from 2.0 per cent to 3.3 per cent, Extraversion penalties amplify from 3.1 per cent to 5.2 per cent, and Emotional Stability penalties nearly double from 3.0 per cent to 5.7 per cent. This finding is unexpected: standard economic models suggest employers should value conscientious workers more when direct monitoring is limited, yet the data show the opposite pattern. The mechanisms behind this counterintuitive result remain unclear and warrant further investigation.

Variable compensation arrangements show more expected patterns for some traits. When bonuses are offered, the Conscientiousness penalty moderates (from 2.1 per cent to 1.8 per cent) and Openness shifts from a non-significant baseline effect to a positive premium (1.8 per cent). This pattern for Conscientiousness aligns with theoretical predictions: performance-based pay and conscientiousness may serve as substitutes in incentivizing effort, as Brenčič and McGee (2023b) suggest in their analysis. However, Agreeableness and Emotional Stability show larger penalties when bonuses are offered, indicating that variable compensation does not universally moderate trait penalties.

Physical Workplace Amenities

Physical workplace amenities generally reduce trait penalties in ways consistent with compensating differentials. Office space amenities eliminate penalties for most traits, with Conscientiousness shifting from a penalty (2.1 per cent) to a non-significant positive effect

(0.3 per cent). Location and commuting benefits similarly moderate penalties, particularly for Conscientiousness and Emotional Stability.

6.2.4 Assessment of the Compensating Differentials Hypothesis

The compensating differentials hypothesis receives substantial but selective support. Three key findings emerge.

First, trait-amenity clustering is selective rather than uniform. Extraversion shows the most consistent amenity associations, while other traits show contradictory or negative associations. This indicates that compensating differentials cannot operate through simple uniform bundling of traits with amenities.

Second, additive amenity controls provide limited explanatory power. The stability of trait coefficients when amenity controls are added suggests that baseline wage penalties do not primarily operate through employers offering different amenity packages. This indicates that compensating differentials must operate through interactive rather than additive mechanisms.

Third, the interaction analysis provides strong evidence for context-dependent compensating differentials mechanisms. When personality trait language is combined with traditional benefits (health insurance, retirement contributions), workplace culture signals (work environment, human capital development), and quality physical amenities, trait penalties are reduced or eliminated. However, some amenity combinations amplify rather than moderate trait penalties, particularly remote work and variable compensation.

To summarise, the mechanism operates effectively when personality trait language is combined with traditional benefits and positive workplace culture signals, allowing employers to offer lower monetary compensation while still attracting suitable candidates through enhanced person-organisation fit. However, compensating differentials cannot explain all instances of trait penalties, particularly when traits are combined with amenities that command market premiums.

7 Language Intensity

The baseline analysis established that advertisements mentioning personality traits offer lower wages on average. However, this binary measure treats all trait mentions equally, regardless of how much emphasis employers place on these characteristics. This section examines whether the intensity of personality trait language relates to wage offers.

The relationship between trait intensity and wages can be understood through information processing and signalling frameworks. Information architecture theory suggests that the frequency and prominence of content elements shape their interpretation (Rosenfeld

et al., 2015). When personality traits dominate job descriptions, they may trigger information saturation effects, where excessive repetition diminishes message value (Malhotra, 1982), potentially crowding out discussion of technical requirements. Conversely, Mahjoub and Kruyen (2021) note that detailed and specified job advertisements can positively influence perceptions about job appropriateness and truthfulness, helping potential applicants reduce search costs and better understand person-organisational fit (Wei et al., 2016).

I first present the methodology for measuring trait intensity, then examine how different levels of emphasis on personality traits and their placement in advertisements relate to wages.

7.1 Empirical Methodology

While the boolean approach identifies the presence of personality traits, it does not capture the emphasis or intensity with which traits are mentioned. Term frequency-inverse document frequency (TF-IDF) provides a methodological framework for capturing these linguistic patterns by measuring both local term prominence (how frequently traits appear within a single advertisement) and global distinctiveness (how unique these mentions are across all advertisements) (Salton and Buckley, 1988; Sparck Jones, 1972). High TF-IDF scores indicate advertisements where personality traits are both frequently mentioned and distinctive, suggesting deliberate emphasis beyond standard practice.

Technically, the TF-IDF method quantifies the importance of a word within a document by considering both its frequency in a specific document (Term Frequency) and its rarity across all documents (Inverse Document Frequency). For a term t in document d within a corpus of N documents:

$$\text{TF-IDF}(t, d) = \text{TF}(t, d) \times \text{IDF}(t)$$

where the Term Frequency is:

$$\text{TF}(t, d) = \text{frequency of term } t \text{ in document } d$$

and the Inverse Document Frequency is:

$$\text{IDF}(t) = \log\left(\frac{N}{\text{number of documents containing term } t}\right)$$

To measure trait intensity, I transform both job descriptions and trait-related keywords into TF-IDF vectors. I then calculate the cosine similarity between these vectors, which

measures the cosine of the angle between them:

$$\text{similarity}(d, \text{trait}) = \frac{\vec{d} \cdot \vec{\text{trait}}}{||\vec{d}|| \times ||\vec{\text{trait}}||}$$

where $\vec{d}$ is the TF-IDF vector of the job description and $\vec{\text{trait}}$ is the TF-IDF vector of the trait keywords.

This approach generates a continuous *trait intensity score* between -1 and 1 for each trait in each advertisement. A higher score indicates that the job posting places greater emphasis on the trait, accounting for both the frequency of trait mentions and their relative importance in the broader corpus. A detailed mathematical description of the TF-IDF calculation and cosine similarity measure is provided in Appendix C.

7.1.1 Empirical models

I investigate the intensity hypothesis through three complementary approaches. First, I examine the linear relationship between trait intensity and wages. For each trait t, I estimate:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \beta_t Trait\_Score_{ti} + \sum_{j \neq t}^{5} \beta_j Trait\_Score_{ji} \\ + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{10}$$

where $Trait\_Score_{ti}$ is the standardised intensity score for trait t in advertisement i, derived from the cosine similarity between the advertisement's TF-IDF vector and the trait's TF-IDF vector. Crucially, this specification includes all advertisements in the sample: those not mentioning the trait receive a score of zero, while those mentioning it receive positive intensity scores. This approach captures the combined effect of the decision to mention traits and the intensity chosen. The other traits ($Trait\_Score_{ji}$) are also included as standardised scores.

Second, to examine potential non-linearities and separate low-intensity from high-intensity mentions, I divide advertisements into quintiles based on their trait intensity scores. For each trait t, I estimate:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \sum_{q=1}^{5} \beta_{tq} (Trait_{ti} \times Q_{qi}) + \sum_{j \neq t}^{5} \beta_j Trait_{ji} \\ + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{11}$$

where Q_{qi} is a dummy variable indicating whether advertisement i belongs to the q-th quintile of trait t intensity. The baseline category is advertisements that do not mention the trait at all. In this specification, the other traits ($Trait_{ji}$) are included as binary

indicators.

The coefficients β_{tq} measure the wage effect for each quintile of trait intensity compared to not mentioning the trait, thus capturing the combined effect of mentioning and intensity level. However, the pattern across quintiles provides information about the intensive margin: if $\beta_{t5} > \beta_{t4} > ... > \beta_{t1}$, this indicates that higher intensity has additional wage effects beyond the initial decision to mention traits. This approach allows me to identify whether low-intensity mentions (Q1) have different wage effects compared to high-intensity mentions (Q5), and how these compare to not mentioning traits at all. I also provide a robustness exercise using simple frequency counts.

Third, I examine how the relationship between trait intensity and wages varies with firm characteristics by estimating, for each trait t and firm characteristic f:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \sum_{q=1}^{5} \beta_{tq}(Trait_{ti} \times Q_{qi}) + \sum_{j \neq t}^{5} \beta_j Trait_{ji} \\ + \gamma_f Firm\_Char_{jf} + \sum_{q=1}^{5} \nu_{tfq}(Trait_{ti} \times Q_{qi} \times Firm\_Char_{jf}) \\ + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \epsilon_i \end{aligned} \tag{12}$$

where $Firm\_Char_{jf}$ represents standardised firm characteristics such as size, age, profit, and a dummy for public listing status. The coefficients ν_{tfq} capture how the wage effect of trait intensity at each quintile q varies with firm characteristic f. This approach tests whether certain types of firms are more likely to use trait-related language for different purposes. For instance, smaller or younger firms might have stronger incentives to use trait-related language as a compensating differential, resulting in larger wage penalties at high intensity levels.

If the linguistic intensity hypothesis is correct, we would expect different wage effects across quintiles. Furthermore, these patterns might vary systematically across different types of firms.

7.2 Results

7.2.1 Intensity scores and linear effects

Table 10 shows the effects of a one standard deviation increase in trait scores on wages. The linear intensity analysis reveals that stronger emphasis on personality traits is associated with wage penalties for most traits. Extraversion and Conscientiousness show the largest intensity effects (-0.9 percentage points each), followed by Emotional Stability (-0.5 percentage points). Agreeableness and Openness show no significant linear relationship between intensity and wages. These results suggest that the wage penalty for personality

traits marginally increases with the emphasis placed on them in job advertisements.

Table 10: Effects of standardized personality trait intensity on wages

Agreeableness	Conscientiousness	Emotional Stability	Extraversion	Openness
-0.002	-0.009**	-0.005**	-0.009***	0.001
(0.009)	(0.003)	(0.002)	(0.003)	(0.003)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients show the effect of a one standard deviation increase in trait intensity on log wages. Regressions control for binary trait indicators, job requirements, location, time-month, occupation and firm fixed effects. Sample size: 2,921,390 observations.

7.2.2 Non-linear effects across intensity quintiles

To examine non-linear effects, I divide advertisements into quintiles based on trait intensity and estimate the wage effects for each quintile, controlling for location, month-year, occupation, and firm fixed effects. Figure 3 shows these results for each of the Big Five traits, with advertisements that do not mention the trait as the baseline[4], with advertisements that do not mention the trait as the baseline.

The results show clear non-linear patterns that provide important insights into the wage penalty paradox. For most traits, low-intensity mentions (first quintile) have small or negligible wage effects, while high-intensity mentions (fourth and fifth quintiles) show larger wage penalties. This pattern is most pronounced for Agreeableness, which shows positive coefficients in the first quintile that become increasingly negative through higher quintiles, reaching significant penalties in the fourth quintile. Conscientiousness follows a similar trajectory, with near-zero effects in the first quintile and progressively more negative effects in higher quintiles.

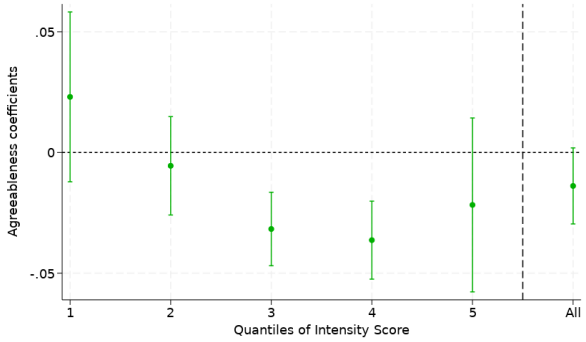
Extraversion has consistent penalties across all intensity levels, though these become more pronounced at higher intensities. Emotional Stability shows significant penalties beginning in the second quintile and maintaining these through higher intensity levels. Openness stands apart from other traits, maintaining relatively stable, small positive coefficients across most quintiles.

These intensity gradients suggest that employers use trait mentions strategically depending on emphasis level. When traits are mentioned with low intensity, they likely signal job requirements without triggering wage penalties. When traits receive high emphasis, they may serve different strategic purposes—potentially as compensating differential signals or demographic targeting mechanisms—that justify lower wage offers.

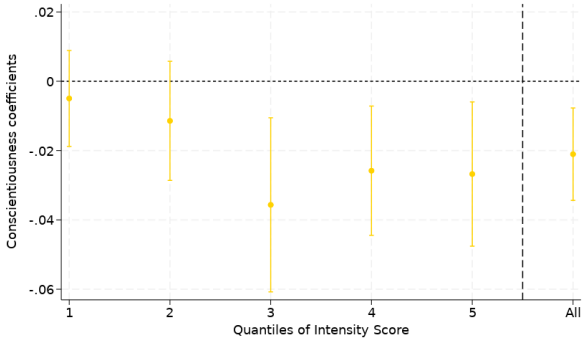
[4]Table A.9 in Appendix A presents the corresponding estimates

Figure 3: Binary Trait to Wage coefficients at different quantiles of the intensity score.

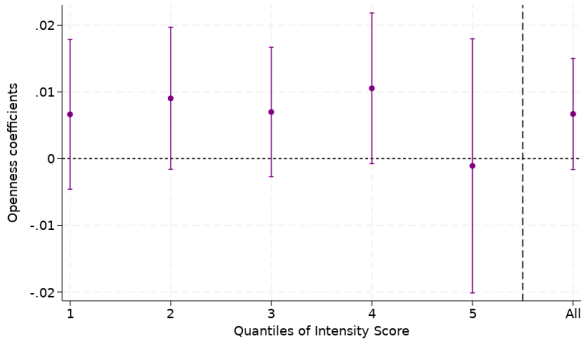
(a) Agreeableness



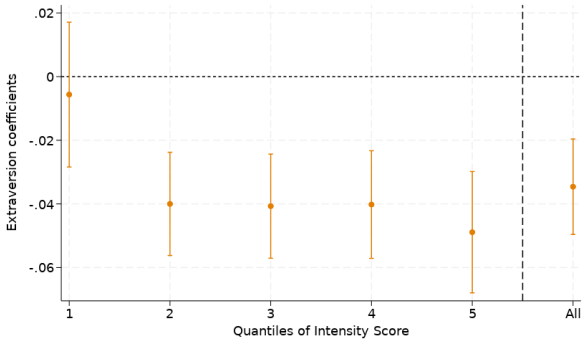
(b) Conscientiousness



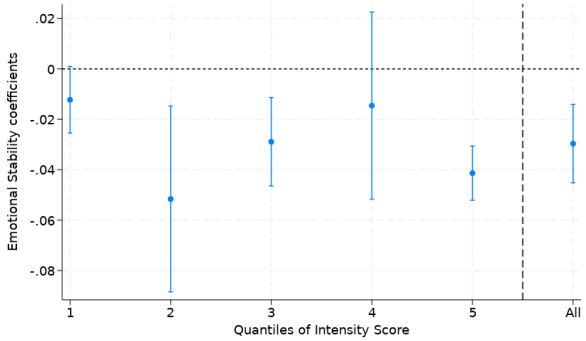
(c) Openness



(d) Extraversion



(e) Emotional Stability



Notes: Author's creation using Adzuna. Baseline is *no trait demand*.

7.2.3 Interaction with firm characteristics

The firm interaction analysis reveals that intensity effects vary systematically across different types of organisations, providing further evidence for strategic use of personality trait language. Tables A.5 through A.8 present these interaction effects across firm characteristics.

Larger firms demonstrate markedly different intensity patterns compared to average-sized firms. For Openness, large firms show significant premiums for low-intensity mentions (2.7 per cent in Q1) and high-intensity mentions (3.5 per cent in Q5), suggesting these firms may value this trait across intensity levels. Similarly, large firms show premiums for low-intensity Conscientiousness mentions (3.0 per cent in Q1), indicating that established firms with greater resources may use modest trait language to signal requirements.

Publicly quoted firms show dramatically different patterns, particularly for high-intensity trait mentions. For Conscientiousness, public firms impose severe penalties for high-intensity mentions, reaching -16.3 per cent in Q3 and -10.0 per cent in Q4, compared to much smaller penalties in non-public firms. Similarly, public firms penalize high-intensity Agreeableness mentions much more severely (-9.6 per cent in Q4) than private firms (-1.4 per cent). These patterns suggest that public firms may face different pressures that lead them to use high-intensity trait language for purposes other than identifying productive characteristics.

Firm age and profitability show more modest interaction effects, though older and more profitable firms generally show larger penalties for high-intensity trait mentions, particularly for Emotional Stability.

7.2.4 Frequency-Based Intensity Analysis

As a robustness check on the TF-IDF intensity analysis, I examine trait wage effects using simple frequency counts of trait mentions within advertisements. This approach provides a transparent measure of emphasis by counting raw mentions (1, 2, 3, 4, or 5+ times) without the contextual weighting inherent in TF-IDF measures. Table A.10 in Appendix A presents the wage effects for each frequency category, with advertisements not mentioning the trait as the baseline.

The frequency analysis confirms the non-linear intensity patterns observed in the TF-IDF analysis while providing additional insights into how raw repetition affects wages. Most traits show escalating penalties with increased mention frequency. Conscientiousness demonstrates a clear frequency gradient, with penalties increasing from -1.7 per cent for single mentions to -5.8 per cent for five or more mentions. Extraversion shows the most consistent escalation pattern, with penalties rising from -2.8 per cent (single mention) to

-9.2 per cent (five or more mentions).

Emotional Stability exhibits similar patterns, though with some variation at intermediate frequencies, ranging from -3.2 per cent for single mentions to -10.1 per cent at the highest frequency. Agreeableness shows more erratic patterns across frequency levels, with single mentions carrying -1.4 per cent penalties but higher frequencies showing mixed effects, including some positive but non-significant coefficients.

Openness maintains its exceptional status in the frequency analysis, showing positive premiums for moderate repetition (1.8 to 2.7 per cent for 2-4 mentions) before turning negative only at very high frequencies (-4.1 per cent for 5+ mentions). This pattern reinforces that Openness functions differently from other personality traits, with employers willing to pay premiums for moderate emphasis on this characteristic.

The consistency between TF-IDF quintile patterns and simple frequency counts provides strong evidence that the intensity effects are robust to different measurement approaches. Both methods demonstrate that low-emphasis trait mentions often carry minimal penalties while high-emphasis mentions show substantial penalties, supporting the hypothesis that employers use trait language strategically depending on the level of emphasis they wish to convey.

7.2.5 Trait Position Effects

I also examine whether the position of personality trait mentions within job advertisements affects their wage impact. Table 11 presents the wage effects across position quintiles, with advertisements not mentioning the trait as the baseline.

Table 11: Effects of trait position quintiles on wages

Trait	1-20% (Start)	21-40%	41-60%	61-80%	81-100% (End)
Agreeableness	-0.023***	-0.006	-0.013	-0.020**	-0.006
	(0.008)	(0.020)	(0.015)	(0.009)	(0.006)
Conscientiousness	-0.014**	-0.028***	-0.020	-0.015*	-0.024***
	(0.006)	(0.009)	(0.013)	(0.008)	(0.008)
Emotional Stability	-0.032**	-0.022	-0.040***	-0.031***	-0.031***
	(0.016)	(0.020)	(0.010)	(0.010)	(0.011)
Extraversion	-0.035***	-0.036***	-0.031***	-0.031***	-0.025***
	(0.010)	(0.010)	(0.008)	(0.007)	(0.008)
Openness	0.018*	0.003	0.005	0.001	0.008
	(0.009)	(0.006)	(0.005)	(0.011)	(0.006)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Baseline is no trait mention. Position quintiles based on trait's first appearance as a percentage of advertisement length. Regression control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 9,888,315 observations.

The position analysis provides complementary evidence that linguistic presentation

influences wage effects. Agreeableness shows the largest penalty (-2.3 per cent) when mentioned early in advertisements, with penalties diminishing when mentioned later in the text. Conscientiousness and Emotional Stability show penalties across most positions, ranging from -1.4 to -2.8 per cent for Conscientiousness and -2.2 to -4.0 per cent for Emotional Stability. Extraversion demonstrates consistent penalties of approximately -3.1 to -3.6 per cent across all positions, suggesting this trait carries wage penalties regardless of placement.

Only Openness shows a positive coefficient (1.8 per cent) when mentioned early, becoming non-significant when mentioned later. This pattern aligns with the intensity analysis, confirming that Openness differs fundamentally from other traits in its relationship with wages.

7.2.6 Assessment of the Linguistic Intensity Hypothesis

The linguistic intensity analysis provides evidence that employers use personality trait language strategically with different wage implications depending on the emphasis level. Three key findings emerge.

First, the non-linear relationship between trait intensity and wages shows that low-intensity mentions often carry minimal or no wage penalties, while high-intensity mentions consistently show larger penalties. This pattern suggests that moderate use of personality trait language likely reflects job requirements, while excessive emphasis serves different strategic purposes that justify lower wage offers.

Second, the firm interaction analysis reveals that established, larger firms show more favourable wage effects for low-intensity trait mentions, suggesting these organisations use trait language primarily to signal requirements. Conversely, publicly quoted firms impose severe penalties for high-intensity trait mentions, indicating that market pressures may lead these firms to use trait language for compensating differentials or signalling purposes rather than identifying productive characteristics.

Third, the position analysis confirms that linguistic presentation influences wage effects, although most traits carry penalties regardless of placement. The exception is Openness, which shows positive effects when mentioned prominently. The pattern of trait-related wage penalties appears consistently across all measurement approaches: binary indicators, continuous intensity scores, intensity quintiles, and placement analysis all show that personality trait language is associated with lower wages. This consistency across different measurement strategies indicates that the trait wage penalty reflects genuine employer behaviour rather than being driven by particular measurement choices.

These findings indicate that linguistic intensity provides important insights into the trait wage penalty paradox. The relationship between personality traits and wages is not

uniform but depends critically on how employers choose to emphasise these characteristics. The evidence suggests that employers use varying intensity levels to signal different messages: moderate emphasis signals job requirements, while high intensity may serve compensating differential functions or target specific demographic groups willing to accept lower wages for perceived better organisational fit.

8 Demographic Targeting

A fourth explanation draws on statistical discrimination models (Aigner and Cain, 1977), suggesting employers use personality-traits to strategically target demographic groups they believe will accept lower wages. For instance, employers might assume women or younger workers value job aspects related to certain personality traits (such as agreeableness or enthusiasm) more highly than monetary compensation. By tailoring advertisements with specific personality trait language, employers may induce self-selection among these groups (Delfgaauw and Dur, 2007; Salop and Salop, 1976).

Recent experimental evidence supports this demographic targeting explanation. Koçak et al. (2023) examined how younger and older job seekers respond to stereotyped traits in job advertisements. Older job seekers' application intentions decreased when job advertisements emphasised traits like "flexibility", which activated negative stereotypes about their age group. Similarly, Koçak and Derous (2025) found that women's attraction to top-executive positions was lower when job advertisements contained men-related stereotyped competencies (such as "leading" or "decision-making") compared to other competencies such as "coaching" or "customer service orientation".

Research on gendered language and wage offers in job advertisements further supports this explanation. Kuhn et al. (2020) found that jobs using feminine-coded language offered wages 7.8 per cent lower than gender-neutral postings. Similarly, Chaturvedi et al. (2021) showed that gendered language in advertisements influenced application patterns, with women more likely to apply for lower-paying positions described with feminine-coded language. This targeting mechanism suggests personality trait language might serve as an implicit demographic filter that allows employers to offer lower wages while still attracting their preferred candidates.

8.1 Empirical Methodology

To test the hypothesis that employers use personality trait language to target specific demographic groups who may accept lower wages, I employ a machine learning approach to identify implicit gender preferences in job advertisements. This builds on recent literature examining gendered language in recruitment. I implement a strategy inspired by Kuhn et al. (2020) and Chaturvedi et al. (2021). While the former implements a Bernoulli naive

Bayes classifier to infer the implicit gender preference in job's title, the latter constructs measures of *maleness* and *femaleness* using a multinomial logistic regression classifier on the advertisement explicit preferences. Chaturvedi et al. (2021) demonstrate that implicit hints within job text effectively direct where men and women send their applications, with jobs implicitly directed to women offering lower wages and receiving a higher fraction of female applications.

The UK Equality Act 2010 generally prohibits explicit gender requirements in job advertisements, with exceptions for occupational requirements such as same-gender care roles. Despite these restrictions, employers can use subtler linguistic patterns to signal gender preferences, such as female-related skills, number of qualifications or other gender-stereotyped words like *support, understanding, interpersonal* (Abraham et al., 2024; Gaucher et al., 2011). To detect these patterns without restricting the analysis to a dictionary, I use the BERT (Bidirectional Encoder Representations from Transformers) model, a state-of-the-art natural language processing model that understands words in context rather than in isolation.

Unlike traditional methods such as keyword counting or logistic regression, BERT captures complex contextual language patterns. Traditional approaches treat words as independent units, while BERT considers words in relation to their surrounding context. This makes BERT particularly effective at detecting subtle biases in language that simpler methods might miss. For example, phrases like "competitive individual" and "nurturing environment" may carry gendered connotations that are difficult to identify with keyword-based methods but can be captured by contextual models like BERT.

I create a training dataset by identifying advertisements with clearly gendered language. I use predefined gender-associated keywords from Chaturvedi et al. (2021) ("female", "woman", "women", "girl", "lady", "ladies" and "feminine" for female candidates, and "male", "males", "man", "men", "guy", "guys", "boy", "boys", "gent" or "gents" for male candidates) to identify explicitly gendered advertisements. For instance, an explicit women-targeted advertisement stated "*...work in a brand spanking new salon, all decorated to a glam queen standard with friendly girls who love beauty...*", while a men-targeted advertisement contained "*...a big personality who can lead from the front, with a wicked eye for detail, and is a highly organised and effective man manager...*".

I apply several preprocessing steps to the training data before estimating the BERT model. First, to ensure the algorithm learns the underlying patterns of gendered language rather than simply identifying explicit gender terms, once advertisements were categorised as explicit female-, male-, or neutral-oriented, I replaced the gender keywords with neutral tokens, and trained the model on the remaining language.

Then, to avoid false positives from care roles mentioning client gender, I exclude

advertisements where gendered terms appear near contextual phrases like "service for", "client is a", or "care for" unless they also contain explicit recruitment language like "female only" or "recruiting for". For instance, advertisements containing phrases like "care for elderly women" are not classified as gender-targeted unless they also contain phrases like "female applicants only", which would indicate a preference for the applicant's gender rather than describing clients.

I remove common stopwords (such as "the", "and", "a") and stem words to their root forms (e.g., converting "caring", "cares", and "cared" to "care"). Finally, I stratify the training data by occupation to prevent the model from simply learning occupation-specific language patterns, as some occupations (such as care roles) have a higher prevalence of explicit gender mentions.

From the 11.5 million vacancies analysed, 1.62 per cent were categorised as explicitly targeting women, 1.35 per cent targeting men, and 1 per cent mentioned both genders. This explicitly gendered subset forms the training data for the model, which then classifies the remaining advertisements according to their implicit gender orientation.

The validation metrics show that the model achieves 94.6 per cent accuracy on the test set, with an F1-score of 0.947 averaged across all classes. The model demonstrates consistent performance in classifying gender preferences, with precision rates of 95.4 per cent for female-targeted advertisements, 93.4 per cent for male-targeted, and 95.2 per cent for neutral advertisements. The confusion matrix analysis shows that the model correctly classifies the majority of advertisements, though there are some misclassifications between female and male categories (613 female ads classified as male, and 400 male ads classified as female). The model identifies distinctive linguistic patterns associated with gender targeting, with common terms like "support" appearing across categories but with different weights and contextual usage. A detailed analysis of model performance by occupation is presented in Appendix D.

8.1.1 Empirical models

Using this gender classification, I incorporate implicit gender targeting into the wage model by estimating,for each trait t:

$$\begin{aligned}\ln(wage_i) = \beta_0 + \beta_t Trait_{ti} + \sum_{j \neq t}^{5} \beta_j Trait_{ji} + \sum_{g=1}^{3} \gamma_g Gender\_Target_{gi} \\ + \beta_{gt}(Trait_{ti} \times Gender\_Target_{gi}) + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i\end{aligned} \tag{13}$$

where $Gender\_Target_{gi}$ is a set of categorical indicators for whether advertisement i implicitly targets women, men, or is gender-neutral (based on the BERT classification). The model includes both the main effect of gender targeting ($\gamma_g Gender\_Target_{gi}$) and its

interaction with each trait ($\beta_{gt}(Trait_{ti} \times Gender\_Target_{gi})$). The coefficients β_{gt} measure how the wage impact of personality trait t varies depending on the gender targeting category g.

If the demographic targeting hypothesis is correct, the wage penalty observed in the pooled sample could reflect a compositional effect: traits are mentioned more frequently in advertisements targeting demographic groups that receive lower wage offers overall. Under this hypothesis, when estimating wage effects separately by gender targeting category, trait penalties should be substantially reduced or eliminated in both categories. Alternatively, if trait language serves as a strategic targeting mechanism, we would expect systematically larger penalties in female-targeted advertisements where employers seek to attract workers willing to accept lower wages.

For testing both the intensity (Hypothesis 2) and gender targeting (Hypothesis 3) interactions together, I estimate for each trait t:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \beta_t Trait\_Score_{ti} + \sum_{j \neq t}^{5} \beta_j Trait_{ji} + \sum_{g=1}^{3} \gamma_g Gender\_Target_{gi} \\ + \beta_{gt}(Trait\_Score_{ti} \times Gender\_Target_{gi}) + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{14}$$

This combined approach allows me to test whether the intensity of trait language interacts with gender targeting to affect wages.

8.2 Results

8.2.1 Gender Targeting and Personality Trait Prevalence

Before analysing wage effects, I first examine whether employers strategically use different personality trait language when targeting specific demographic groups. Table 12 shows that female-targeted advertisements are significantly more likely to mention Openness (3.1 per cent), Extraversion (2 per cent), and Agreeableness (1.1 per cent). No significant differences were found for Conscientiousness or Emotional Stability. This pattern provides initial evidence that employers do use personality trait language differently when targeting female candidates, particularly traits associated with interpersonal skills and openness to experience.

8.2.2 Wage Effects by Gender Targeting

Using the BERT-based classification model, I analyse how personality trait wage effects differ between female-targeted and male-targeted advertisements. Table 13 presents these results.

Personality traits show significant wage penalties across both female-targeted and

Table 12: Effect of female-targeted ads on personality trait demand

	Agree.	Consc.	Emot.	Extra.	Open.
Effect of Female-targeting	0.011***	-0.002	-0.005	0.020***	0.031***
	(0.003)	(0.004)	(0.006)	(0.006)	(0.006)

Notes: Coefficients show the effect of female gender-targeting on trait demand. Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Regressions control for job requirements, location, time-month, occupation and firm fixed effects. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Sample size: 9,466,620 observations.

Table 13: Total effects of personality traits by gender targeting in job advertisements

	Consc.	Open.	Extra.	Agree.	Emot.
Female-intentioned ads	-0.018**	0.004	-0.032***	-0.022**	-0.046***
	(0.009)	(0.005)	(0.010)	(0.009)	(0.011)
Male-intentioned ads	-0.025***	0.003	-0.043***	-0.024***	-0.038***
	(0.007)	(0.005)	(0.006)	(0.008)	(0.008)
p-value: female = male	0.452	0.873	0.172	0.879	0.557

Notes: Total effects (base + interaction). Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. The p-value row tests H_0: total effect for female-intentioned ads = total effect for male-intentioned ads. This is equivalent to testing whether the gender interaction coefficient is zero. Regressions control for job requirements, location, time-month, occupation and firm fixed effects. N = 9,466,620 observations.

male-targeted advertisements. This indicates that the pooled wage penalty is not primarily driven by gender targeting. If demographic targeting explained the pooled penalty through compositional effects, penalties should diminish or disappear when estimated separately by gender category. Instead, penalties persist robustly in both categories. Most traits show larger penalties in male-targeted advertisements. Extraversion carries a stronger penalty in male-targeted advertisements (4.3 per cent versus 3.2 per cent), as does Conscientiousness (2.5 per cent versus 1.8 per cent) and Agreeableness (2.4 per cent versus 2.2 per cent). Only Emotional Stability shows a larger penalty in female-targeted advertisements (4.6 per cent versus 3.8 per cent). Openness remains the only trait without significant penalties in either gender category. A standrad t test fails to reject the null of no difference between the male and women coefficients.

These patterns contradict the demographic targeting hypothesis. The persistence of similar penalty structures across gender categories indicates that other mechanisms drive the trait wage penalty. These include compensating differentials, linguistic intensity effects, or organisational signalling rather than strategic demographic targeting.

8.2.3 Trait Intensity and Gender Targeting Interactions

To further explore this relationship, I analyse how trait intensity interacts with gender targeting. Table 14 presents the wage effects across different quintiles of trait intensity for both female-targeted and male-targeted advertisements.

Table 14: Effects of personality trait intensity by gender targeting (categorical)

Trait	Gender	Q1	Q2	Q3	Q4	Q5
Agreeableness	Male	-0.001	-0.016	-0.038***	-0.042***	-0.020
		(0.009)	(0.011)	(0.009)	(0.010)	(0.028)
	Female	0.002	-0.008	-0.033***	-0.041***	-0.038***
		(0.011)	(0.013)	(0.012)	(0.012)	(0.014)
Conscientiousness	Male	-0.011	-0.012	-0.038***	-0.024**	-0.031**
		(0.008)	(0.009)	(0.011)	(0.010)	(0.012)
	Female	0.001	-0.014	-0.008	-0.035*	-0.039**
		(0.008)	(0.010)	(0.012)	(0.018)	(0.016)
Emotional Stability	Male	-0.016**	-0.053***	-0.040***	-0.034**	-0.051***
		(0.007)	(0.018)	(0.011)	(0.017)	(0.008)
	Female	-0.025***	-0.075**	-0.035***	-0.039***	-0.054***
		(0.010)	(0.035)	(0.013)	(0.008)	(0.005)
Extraversion	Male	-0.018*	-0.048***	-0.046***	-0.043***	-0.049***
		(0.009)	(0.008)	(0.007)	(0.008)	(0.010)
	Female	-0.001	-0.027**	-0.037***	-0.039***	-0.060***
		(0.012)	(0.011)	(0.013)	(0.011)	(0.011)
Openness	Male	0.012*	0.004	0.001	0.008	-0.009
		(0.006)	(0.007)	(0.006)	(0.007)	(0.010)
	Female	0.001	0.004	0.007	0.005	-0.000
		(0.010)	(0.009)	(0.007)	(0.006)	(0.012)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Q1-Q5 represent quintiles of trait intensity (1=lowest, 5=highest). Regressions control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 9,466,620 observations.

The intensity analysis reveals that gender targeting modifies trait-wage relationships in ways that vary by trait and intensity level. For Conscientiousness, female-targeted advertisements show no significant penalty at low intensity (Q1: 0.1 per cent, non-significant) but increasingly negative effects at higher intensities, reaching 3.9 per cent in the highest quintile. Male-targeted advertisements show penalties beginning at moderate intensities, with effects ranging from 1.2 per cent (Q2, non-significant) to 3.1 per cent (Q5).

Emotional Stability has the most severe penalties across both gender categories, with particularly pronounced effects in the second quintile for female-targeted advertisements (7.5 per cent) and male-targeted advertisements (5.3 per cent). The penalties generally

persist across higher intensity levels, reaching 5.4 per cent for female-targeted and 5.1 per cent for male-targeted advertisements at the highest intensity.

Extraversion shows different patterns across gender targeting. In male-targeted advertisements, penalties are significant across all intensity levels, ranging from 1.8 per cent (Q1) to 4.9 per cent (Q5). Female-targeted advertisements show no significant penalty at the lowest intensity but larger penalties at higher intensities, reaching 6.0 per cent at the highest level.

Agreeableness exhibits similar patterns across gender categories, with both showing non-significant effects at low intensities transitioning to significant penalties of approximately 4.0 per cent at higher intensities (Q3-Q4). This similarity suggests that, regardless of gender targeting, employers use high-intensity Agreeableness language for similar strategic purposes.

Openness maintains its exceptional status, showing no significant wage effects across any intensity level for either gender category, with coefficients remaining small and predominantly positive.

8.2.4 Assessment of the Demographic Targeting Hypothesis

The demographic targeting hypothesis is not supported by the evidence. Three key findings lead to this conclusion.

First, employers do use personality trait language differently when targeting women versus men. Female-targeted advertisements are more likely to mention Openness, Extraversion, and Agreeableness. However, this differential usage does not translate into the wage patterns predicted by demographic targeting theory. If trait language served primarily to attract demographic groups willing to accept lower wages, we would expect larger penalties in female-targeted advertisements. Instead, most traits (Conscientiousness, Extraversion, Agreeableness) have larger penalties in male-targeted advertisements. Only Emotional Stability shows the opposite pattern.

Second, the persistence of significant penalties in both female-targeted and male-targeted advertisements indicates that the pooled wage penalty does not result from compositional effects driven by gender targeting. When estimated separately by gender category, trait penalties remain substantial and significant. This suggests that trait-wage relationships operate within gender categories rather than being explained by the differential targeting of demographic groups that receive different wage offers.

Third, both female-targeted and male-targeted advertisements show similar patterns of increasing penalties with trait intensity. Agreeableness shows identical penalty structures across gender categories. Both show non-significant effects at low intensities that transition

to 4.0 per cent penalties at higher intensities. This demonstrates that employers use trait intensity strategically regardless of gender targeting. This further undermines the demographic targeting explanation.

These findings rule out demographic targeting as the primary driver of the wage penalty paradox. The evidence instead supports the mechanisms documented in previous sections. Trait penalties reflect compensating differentials when combined with appropriate amenity bundles (Section 6). They vary systematically with linguistic intensity (Section 7). They correlate with occupation-level characteristics (Section 5). While employers do strategically differentiate trait language by gender targeting, this differentiation does not explain why trait mentions are associated with lower wages.

9 Conclusion

This research has examined the paradoxical relationship between personality trait mentions in job advertisements and posted wages in the UK labour market. While individual-level studies consistently find that traits like Conscientiousness and Emotional Stability increase personal wages, I documented a significant negative relationship between personality trait mentions in job advertisements and posted wages. Using data from 11.7 million online job advertisements posted between 2017 and 2022, I tested four potential explanations for this paradox: correlation with other low-pay characteristics, compensating differentials through amenity provision, linguistic intensity effects, and demographic targeting strategies.

The baseline analysis confirmed the wage penalty paradox. Advertisements mentioning personality traits offered 4.6 per cent lower wages on average, even after controlling for job, occupation, and firm characteristics. Most traits showed significant penalties: Extraversion (3.2 per cent), Emotional Stability (3.1 per cent), Conscientiousness (2.0 per cent), and Agreeableness (1.4 per cent). Only Openness showed a small positive effect (0.7 per cent), though this was only marginally significant.

Correlation with Low-Pay Characteristics

The analysis of correlation with low-pay characteristics provided partial but incomplete explanations for the wage penalty paradox. Occupation fixed effects produced the largest reduction in trait penalties, demonstrating that personality traits appear disproportionately in lower-paying occupations. However, significant penalties persisted even after controlling for occupation, job, and firm characteristics.

The occupation-level analysis revealed that trait penalties are smaller in occupations with higher proportions of women and migrants but larger in occupations with more degree holders and higher mean wages. This pattern—where traits face greater penalties in high-skill, high-wage contexts—contradicts simple productivity-based explanations and suggests strategic use of trait language in different occupational contexts.

Job-level interactions showed that personality traits receive more favourable treatment in entry-level positions and when explicit educational requirements are mentioned. However, traits face amplified penalties when combined with hard skills requirements and temporary contracts. Firm-level analysis provided limited support, with only public listing status showing meaningful effects that amplified rather than explained trait penalties.

Compensating Differentials Through Amenity Provision

The compensating differentials analysis revealed context-dependent support for this mechanism. Simple additive controls for amenities had minimal impact on trait penalties,

with the R^2 increasing only marginally from 0.598 to 0.601. This stability suggests that baseline trait penalties do not operate through systematic correlation with amenity provision.

However, the interaction analysis provided evidence for strategic compensating differentials mechanisms. When personality trait language was combined with traditional benefits (health insurance, retirement contributions), workplace culture signals (work environment, human capital development), and quality workplace amenities, trait penalties were reduced or eliminated. These patterns support compensating differentials theory: employers can offer lower wages when trait requirements are bundled with attractive non-monetary compensation and positive organisational culture signals.

Conversely, some amenity combinations amplified trait penalties. Remote work created uniformly larger penalties for most traits, while variable compensation arrangements showed mixed patterns. This heterogeneity suggests that employers strategically deploy compensating differential mechanisms for certain trait-amenity combinations while using personality trait language for different purposes in other contexts.

Linguistic Intensity Effects

The linguistic intensity analysis provided the strongest evidence for strategic employer behaviour and helped explain the wage penalty paradox. The relationship between trait emphasis and wages was highly non-linear: low-intensity mentions often carried minimal or no wage penalties, while high-intensity mentions consistently showed larger penalties.

This intensity gradient suggests that employers use trait mentions strategically depending on the emphasis level. Moderate use likely signals job requirements without triggering wage penalties, while excessive emphasis serves different strategic purposes—potentially compensating differential signals or demographic targeting mechanisms—that justify lower wage offers.

Firm interactions revealed that established, larger firms showed more favourable wage effects for low-intensity trait mentions, suggesting these organisations use trait language primarily to signal requirements. Conversely, publicly quoted firms imposed large penalties for high-intensity trait mentions, indicating that market pressures may lead these firms to use trait language for signalling purposes rather than identifying productive characteristics.

Demographic Targeting

The demographic targeting analysis provided limited support for this explanation. While employers do use personality trait language differently when targeting women versus men—with female-targeted advertisements more likely to mention Openness, Extraversion,

and Agreeableness—this differential usage did not translate into the wage patterns predicted by demographic targeting theory.

Most traits showed larger penalties in male-targeted advertisements, with only Emotional Stability showing a larger penalty in female-targeted contexts. The intensity analysis revealed that both female-targeted and male-targeted advertisements showed similar patterns of increasing penalties with trait intensity, suggesting that trait emphasis effects operate through mechanisms beyond demographic targeting.

Synthesis and Implications

These findings collectively show that the personality trait wage penalty paradox reflects multiple overlapping employer strategies rather than a single mechanism. Three key insights emerge about how employers use personality trait language in job advertisements.

First, the relationship between trait mentions and wages is context-dependent. The same trait can signal requirements, serve compensating differential purposes, or function as demographic targeting depending on how it is presented, what amenities are offered, and which type of organisation is hiring. This context-dependency explains why aggregate analyses show wage penalties despite productivity benefits in specific circumstances.

Second, linguistic presentation matters as much as content. The intensity of trait emphasis, its placement within advertisements, and its combination with other requirements influence wage effects. Low-intensity trait mentions often signal requirements without wage penalties, while high-intensity mentions serve strategic signalling purposes that justify lower compensation.

Third, organisational characteristics shape how personality trait language functions in wage determination. Established, larger firms use trait language differently than smaller, publicly traded, or newer organisations. This variation suggests that employer strategy, rather than uniform trait devaluation, drives the observed wage penalties.

Limitations and Future Research

Several limitations affect the interpretation of these results.

The analysis examines posted wages rather than realised compensation or wage growth trajectories. A competing explanation for the wage penalty paradox is that personality traits are non-verifiable characteristics where employers offer lower starting wages but provide steeper wage growth once these attributes become apparent through workplace performance (Altonji and Pierret, 2001; Farber and Gibbons, 1996; Lange, 2007). While cross-sectional vacancy data cannot directly test this mechanism, several patterns in the results are inconsistent with employer learning being the primary driver of the observed

penalties.

First, the employer learning literature predicts that non-verifiable characteristics should be particularly important in entry-level positions where workers lack demonstrated track records. Employers facing uncertainty about worker traits should offer lower starting wages in these contexts. However, the job-level interaction analysis (Section 5) shows the opposite pattern. Conscientiousness penalties are significantly smaller in low-seniority positions (interaction term +0.021, reducing the penalty from 3.2 per cent to 1.1 per cent) and disappear entirely when no experience is required (interaction term +0.085). If employer learning drove the wage penalty, we would observe larger penalties in entry-level positions where trait verification is most difficult. The data show smaller or non-existent penalties instead. Second, trait penalties are largest in high-skill, high-pay occupations. Third, the systematic interaction patterns with amenities (Section 6) and the non-linear intensity effects (Section 7) suggest employers use trait language strategically for multiple purposes beyond addressing verification problems. Low-intensity mentions carry minimal penalties while high-intensity mentions show substantial penalties, indicating deliberate emphasis choices rather than uniform responses to uncertainty.

While employer learning dynamics may contribute to some wage patterns, these inconsistencies indicate it is not the dominant explanation for why advertisements mentioning personality traits are associated with lower posted wages. Future research combining vacancy data with longitudinal employment records could directly examine how initial wage offers relate to subsequent earnings trajectories for workers in trait-emphasising positions. The identification approach treats advertisements not mentioning personality traits as not requiring them. Jobs might implicitly require traits without explicit mention. While extensive controls help address this concern, unmeasured trait requirements could bias results. However, this limitation affects all research examining job advertisement text and does not explain why jobs that explicitly mention traits offer lower wages compared to those that do not.

Data limitations prevented testing whether personality trait requirements correlate with other non-wage job characteristics not captured in advertisement text, or whether employers offering trait-demanding positions provide different benefits not advertised publicly. Future research could address these gaps through employer surveys, applicant-level data, or experimental approaches examining how job seekers respond to different intensities of personality-related language.

Contribution

This research contributes to understanding how personality traits are valued in contemporary labour markets by revealing the strategic complexity underlying employer

communication. Rather than simply reflecting productivity differences or uniform trait devaluation, personality trait language in job advertisements serves multiple strategic purposes that vary by context, intensity, and organisational characteristics.

This distinction carries practical implications for the use of job postings as a real-time labour market indicator (Hershbein and Kahn, 2018). If trait language reflects recruitment strategy rather than realised wage offers, then wage indices constructed from posted salaries, such as those used by the ONS, may be systematically biased for jobs requiring certain types of workers. More broadly, the gap between advertised requirements and actual hiring outcomes remains an open empirical question that individual-level matched employer-employee data could help resolve.

The findings have practical implications for both job seekers and employers. Job seekers should interpret personality trait language in advertisements as potentially signalling organisational culture, compensation structure, and strategic positioning rather than simply describing required characteristics. Employers should consider how trait-related language affects both the quantity and quality of applicants, recognising that high-intensity trait emphasis may signal compensation limitations that could deter qualified candidates.

More broadly, this research shows that job advertisements function as strategic communications rather than straightforward requirement descriptions. As labour markets evolve with technological advancement and changing work arrangements, understanding these strategic dimensions becomes increasingly important for explaining how employers attract workers and how job seekers evaluate opportunities.

By exposing the relationship between personality trait mentions and wages in job advertisements, this research provides new insights into employer communication strategies and their effects on labour market outcomes. The personality trait wage penalty paradox reflects the ways employers use language to signal organisational characteristics, manage compensation expectations, and attract their desired workforce rather than simple productivity-based valuations.

References

Abraham, L., Hallermeier, J., and Stein, A. (2024). Words matter: Experimental evidence from job applications. *Journal of Economic Behavior & Organization*, 225:348–391.

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, 40(S1):293–340.

Aigner, D. J. and Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *ILR Review*, 30(2):175–187.

Akerlof, G. A. and Kranton, R. E. (2005). Identity and the economics of organizations. *Journal of Economic Perspectives*, 19(1):9–32.

Alderotti, G., Rapallini, C., and Traverso, S. (2023). The Big Five personality traits and earnings: A meta-analysis. *Journal of Economic Psychology*, 94:102570.

Alekseeva, L., Azar, J., Gine, M., Samila, S., and Taska, B. (2021). The demand for ai skills in the labor market. *Labour Economics*, 71:102002.

Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Personality psychology and economics. In *Handbook of the Economics of Education*, volume 4, pages 1–181. Elsevier.

Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Arrow, K. (1971). The theory of discrimination. *Labor Economics*.

Banfi, S. and Villena-Roldan, B. (2019). Do high-wage jobs attract more applicants? directed search evidence from the online labor market. *Journal of Labor Economics*, 37(3):715–746.

Barrick, M. R. and Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, 44(1):1–26.

Batra, H., Michaud, A. M., and Mongey, S. (2023). Online job posts contain very little wage information. Working paper 31984, National Bureau of Economic Research.

Borghans, L., Duckworth, A. L., Heckman, J. J., and Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059.

Boschetti Adamczyk, W., Delaporte, I., and Escudero, V. (2025). Measuring quality of employment in emerging economies: A methodology for assessing job amenities using big data. Research brief, International Labour Organization.

Bowles, S., Gintis, H., and Osborne, M. (2001). Incentive-enhancing preferences: Personality, behavior, and earnings. *American Economic Review*, 91(2):155–158.

Braakmann, N. (2009). The role of psychological traits for the gender gap in full-time employment and wages: evidence from germany. (162). vailable at SSRN: https://ssrn.com/abstract=3169747 or http://dx.doi.org/10.2139/ssrn.3169747.

Breaugh, J. A. (2013). Employee recruitment. *Annual Review of Psychology*, 64:389–416.

Brenčič, V. (2012). Wage posting: evidence from job ads. *Canadian Journal of Eco-*

nomics/Revue canadienne d'économique, 45(4):1529–1559.

Brenčič, V. and McGee, A. (2023a). Demand for personality traits, tasks, and sorting. Discussion Paper 16576, Institute of Labor Economics (IZA).

Brenčič, V. and McGee, A. (2023b). Employers' demand for personality traits. Discussion Paper 16083, Institute of Labor Economics (IZA).

Burbano, V. C. (2016). Social responsibility messages and worker wage requirements: Field experimental evidence from online labor marketplaces. *Organization Science*, 27(4):1010–1028.

Cable, D. M. and Judge, T. A. (1996). Person–organization fit, job choice decisions, and organizational entry. *Organizational Behavior and Human Decision Processes*, 67(3):294–311.

Cassar, L. and Meier, S. (2018). Nonmonetary incentives and the implications of work as a source of meaning. *Journal of Economic Perspectives*, 32(3):215–238.

Chapman, D. S., Uggerslev, K. L., Carroll, S. A., Piasentin, K. A., and Jones, D. A. (2005). Applicant attraction to organizations and job choice: a meta-analytic review of the correlates of recruiting outcomes. *Journal of Applied Psychology*, 90(5):928.

Chaturvedi, S., Mahajan, K., and Siddique, Z. (2021). Words matter: Gender, jobs and applicant behavior. Discussion Paper 14497, Institute of Labor Economics (IZA).

Cobb-Clark, D. A. and Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, 115(1):11–15.

Cobb-Clark, D. A. and Tan, M. (2011). Noncognitive skills, occupational attainment, and relative wages. *Labour Economics*, 18(1):1–13.

Collischon, M. (2020). The returns to personality traits across the wage distribution. *Labour*, 34(1):48–79.

Costa Jr, P. T. and McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, 6(4):343–359.

Cubel, M., Nuevo-Chiquero, A., Sanchez-Pages, S., and Vidal-Fernandez, M. (2016). Do personality traits affect productivity? evidence from the laboratory. *The Economic Journal*, 126(592):654–681.

Delfgaauw, J. and Dur, R. (2007). Signaling and screening of workers' motivation. *Journal of Economic Behavior & Organization*, 62(4):605–624.

Deming (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, 132(4):1593–1640.

Deming and Kahn (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, 36(S1):S337–S369.

Denissen, J. J., Bleidorn, W., Hennecke, M., Luhmann, M., Orth, U., Specht, J., and Zimmermann, J. (2018). Uncovering the power of personality to shape income. *Psychological Science*, 29(1):3–13.

Draca, M., Duchini, E., Rathelot, R., Turrell, A., and Vattuone, G. (2022). Revolution

in progress? the rise of remote work in the uk. Working Paper 616, Centre for Competitive Advantage in the Global Economy (CAGE).

Escudero, V., Liepmann, H., and Vergara, D. (2024). Directed Search, Wages, and Non-Wage Amenities: evidence from an Online Job Board. Discussion Papers 17211, Institute of Labor Economics (IZA).

Farber, H. S. and Gibbons, R. (1996). Learning and wage dynamics. *The Quarterly Journal of Economics*, 111(4):1007–1047.

Flinn, C. J., Todd, P. E., and Zhang, W. (2025). Labor market returns to personality: A job search approach to understanding gender gaps. *Journal of Political Economy*, 133(4):1169–1234.

Flory, J. A., Leibbrandt, A., and List, J. A. (2015). Do competitive workplaces deter female workers? a large-scale natural field experiment on job entry decisions. *The Review of Economic Studies*, 82(1):122–155.

Gaucher, D., Friesen, J., and Kay, A. C. (2011). Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology*, 101(1):109–128.

Gensowski, M. (2018). Personality, IQ, and lifetime earnings. *Labour Economics*, 51:170–183.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. *Review of Personality and Social Psychology*, 2(1):141–165.

Goldberg, L. R. (1990). An alternative "description of personality": the big-five factor structure. *Journal of Personality and Social Psychology*, 59(6):1216—1229.

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, 4(1):26—-42.

Grugulis, I. and Vincent, S. (2009). Whose skill is it anyway?: "soft" skills and polarization. *Work, Employment and Society*, 23(4):597–615.

Handy, F. and Katz, E. (1998). The wage differential between nonprofit institutions and corporations: Getting more by paying less? *Journal of Comparative Economics*, 26(2):246–261.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4):451–464.

Heineck, G. (2011). Does it pay to be nice? personality and earnings in the United Kingdom. *ILR Review*, 64(5):1020–1038.

Hershbein, B. and Kahn, L. B. (2018). Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review*, 108(7):1737–1772.

John (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In *Handbook of personality: Theory and research*, pages 66—-100. Guilford Press.

John and Thomsen (2014). Heterogeneous returns to personality: the role of occupational choice. *Empirical Economics*, 47(2):553–592.

Judge, T. A. and Cable, D. M. (1997). Applicant personality, organizational culture, and organization attraction. *Personnel Psychology*, 50(2):359–394.

Khaouja, I., Kassou, I., and Ghogho, M. (2021). A Survey on Skill Identification From Online Job Ads. *IEEE Access*, 9:118134–118153.

Koçak, A. and Derous, E. (2025). Women's intention to apply to top-executive positions: The role of gender meta-stereotypes in job ads. *Sex Roles*, 91(2):4.

Koçak, A., Derous, E., Born, M. P., and Duyck, W. (2023). What (not) to add in your ad: When job ads discourage older or younger job seekers to apply. *International Journal of Selection and Assessment*, 31(1):92–104.

Koch, T. and Zerback, T. (2013). Helpful or harmful? How frequent repetition affects perceived statement credibility. *Journal of Communication*, 63(6):993–1010.

Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, 49(1):1–49.

Kristof-Brown, A. L., Zimmerman, R. D., and Johnson, E. C. (2005). Consequences of individuals' fit at work: A meta-analysis of person–job, person–organization, person–group, and person–supervisor fit. *Personnel Psychology*, 58(2):281–342.

Krueger, A. B. and Schkade, D. (2008). Sorting in the labor market do gregarious workers flock to interactive jobs? *Journal of Human Resources*, 43(4):859–883.

Kuhn, P. and Shen, K. (2023). What happens when employers can no longer discriminate in job ads? *American Economic Review*, 113(4):1013–1048.

Kuhn, P., Shen, K., and Zhang, S. (2020). Gender-targeted job ads in the recruitment process: Facts from a chinese job board. *Journal of Development Economics*, 147:102531.

Lambrecht, K. (1996). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*, volume 71. Cambridge University Press.

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics*, 25(1):1–35.

Lyu, W. and Liu, J. (2021). Soft skills, hard skills: What matters most? Evidence from job postings. *Applied Energy*, 300:117307.

Maestas, N., Mullen, K. J., Powell, D., Von Wachter, T., and Wenger, J. B. (2023). The value of working conditions in the United States and implications for the structure of wages. *American Economic Review*, 113(7):2007–2047.

Mahjoub, A. and Kruyen, P. M. (2021). Efficient recruitment with effective job advertisement: an exploratory literature review and research agenda. *International Journal of Organization Theory & Behavior*, 24(2):107–125.

Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of Consumer Research*, 8(4):419–430.

Marinescu, I. and Wolthoff, R. (2020). Opening the black box of the matching function: The power of words. *Journal of Labor Economics*, 38(2):535–568.

Matteson, M. L., Anderson, L., and Boyden, C. (2016). "soft skills": A phrase in search of meaning. *Libraries and the Academy*, 16(1):71–88.

McCrae, R. R. and John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):175–215.

McGee, A. and McGee, P. (2020). Whoever you want me to be: personality and incentives. Discussion Paper 13809, Institute of Labor Economics (IZA).

Mueller, G. and Plug, E. (2006). Estimating the effect of personality on male and female earnings. *ILR Review*, 60(1):3–22.

Nandi, A. and Nicoletti, C. (2014). Explaining personality pay gaps in the uk. *Applied Economics*, 46(26):3131–3150.

Nordman, C. J., Sarr, L. R., and Sharma, S. (2019). Skills, personality traits, and gender wage gaps: evidence from bangladesh. *Oxford Economic Papers*, 71(3):687–708.

Nyhus, E. K. and Pons, E. (2005). The effects of personality on earnings. *Journal of Economic Psychology*, 26(3):363–384.

Nyhus, E. K. and Pons, E. (2012). Personality and the gender wage gap. *Applied Economics*, 44(1):105–118.

Oldford, E. and Fiset, J. (2021). Decoding bias: Gendered language in finance internship job postings. *Journal of Behavioral and Experimental Finance*, 31:100544.

Ployhart, R. E., Weekley, J. A., and Baughman, K. (2006). The structure and function of human capital emergence: A multilevel examination of the attraction-selection-attrition model. *Academy of Management Journal*, 49(4):661–677.

Risse, L., Farrell, L., and Fry, T. R. L. (2018). Personality and pay: do gender gaps in confidence explain gender gaps in wages? *Oxford Economic Papers*, 70(4):919–949.

Roberson, Q. M., Collins, C. J., and Oreg, S. (2005). The effects of recruitment message specificity on applicant attraction to organizations. *Journal of Business and Psychology*, pages 319–339.

Rohrbach-Schmidt, D., Wehner, C., Krueger, S., and Ebner, C. (2023). Wage returns to job tasks and personality traits in germany. *International Journal of Manpower*, 44(9):55–71.

Rosen, S. (1986). The theory of equalizing differences. In *Handbook of Labour Economics*, volume 1, pages 641–692. Elsevier.

Rosenfeld, L., Morville, P., and Arango, J. (2015). *Information architecture: for the web and beyond*. O'Reilly Media, Inc.

Salop, J. and Salop, S. (1976). Self-selection and turnover in the labor market. *The Quarterly Journal of Economics*, 90(4):619–627.

Salton, G. and Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5):513–523.

Saucier, G. (1994). Mini-markers: A brief version of goldberg's unipolar big-five markers. *Journal of Personality Assessment*, 63(3):506–516.

Sockin, J. (2022). Show me the amenity: Are higher-paying firms better all around? Working Paper 9842, CESifo.

Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1):11–21.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Stevens, C. D. and Szmerekovsky, J. G. (2010). Attraction to employment advertisements: Advertisement wording and personality characteristics. *Journal of Managerial Issues*, pages 107–126.

Turrell, A., Speigner, B. J., Djumalieva, J., Copple, D., and Thurgood, J. (2019). Transforming naturally occurring text data into economic statistics: The case of online job vacancy postings. Working paper 25837, National Bureau of Economic Research.

Van Hoye, G. and Turban, D. B. (2015). Applicant–employee fit in personality: Testing predictions from similarity-attraction theory and trait activation theory. *International Journal of Selection and Assessment*, 23(3):210–223.

Wei, Y.-C., Chang, C.-C., Lin, L.-Y., and Liang, S.-C. (2016). A fit perspective approach in linking corporate image and intention-to-apply. *Journal of Business Research*, 69(6):2220–2225.

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, 96(5):849–861.

Wille, L. and Derous, E. (2018). When job ads turn you down: how requirements in job ads may stop instead of attract highly qualified women. *Sex Roles*, 79:464–475.

Woessmann, L. (2024). Skills and earnings: A multidimensional perspective on human capital. *Annual Review of Economics*, 17.

Woodcock, S. D. (2008). Wage differentials in the presence of unobserved worker, firm, and match heterogeneity. *Labour Economics*, 15(4):771–793.

10 Appendices

Appendix A - Figures and Tables

Table A.1: Personality Trait Coefficients on Log Posted Wage, by 1-digit SOC Occupation

SOC	Consc.	Open.	Extra.	Agree.	Emot.	N	R^2
1	-0.030***	0.017***	-0.034***	-0.010	-0.039***	600,960	0.584
	(0.011)	(0.006)	(0.007)	(0.012)	(0.011)		
2	-0.019***	-0.008	-0.032***	0.020	-0.035***	2,187,332	0.410
	(0.006)	(0.005)	(0.008)	(0.020)	(0.010)		
3	-0.030***	0.014***	-0.049***	-0.033***	-0.056***	1,875,955	0.464
	(0.005)	(0.005)	(0.009)	(0.010)	(0.007)		
4	-0.024***	0.016***	-0.057***	-0.047***	-0.011*	617,840	0.491
	(0.004)	(0.005)	(0.006)	(0.009)	(0.006)		
5	-0.019*	0.021**	-0.010	-0.005	0.008	732,619	0.543
	(0.012)	(0.009)	(0.008)	(0.013)	(0.025)		
6	-0.001	0.018***	-0.013***	-0.014	-0.035***	1,577,798	0.413
	(0.017)	(0.006)	(0.004)	(0.010)	(0.010)		
7	-0.009	0.022**	-0.036***	-0.043***	-0.020**	903,832	0.515
	(0.010)	(0.010)	(0.010)	(0.011)	(0.009)		
8	0.001	-0.031**	-0.039**	0.081	-0.018	485,450	0.573
	(0.012)	(0.015)	(0.018)	(0.076)	(0.030)		
9	-0.015	-0.001	-0.011	0.055*	0.018	855,124	0.537
	(0.013)	(0.008)	(0.014)	(0.033)	(0.030)		

Notes: Each row presents coefficients and standard errors (in parentheses) from separate regressions for each 1-digit SOC occupation. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. The dependent variable is the natural logarithm of posted salary. SOC codes: 1 = Managers, directors and senior officials; 2 = Professional occupations; 3 = Associate professional and technical occupations; 4 = Administrative and secretarial occupations; 5 = Skilled trades occupations; 6 = Caring, leisure and other service occupations; 7 = Sales and customer service occupations; 8 = Process, plant and machine operatives; 9 = Elementary occupations. Standard errors are clustered at the firm level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. All regressions include controls for soft skills, hard skills, remote work, advertisement length, contract type, work schedule, seniority level, education, experience, amenities availability, and negotiable salary. Fixed effects for location (NUTS2 region), month-year, and firm are included.

Table A.2: Personality Traits coefficients on Log posted wage, by 2-digit occupation

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
11	-0.037***	0.016***	-0.034***	0.004	-0.019*	390,200	0.614
	(0.013)	(0.005)	(0.008)	(0.015)	(0.010)		
12	-0.005	0.015	-0.025***	-0.027*	-0.055***	208,565	0.553
	(0.009)	(0.014)	(0.009)	(0.016)	(0.019)		
21	-0.023***	-0.011**	-0.049***	-0.002	-0.018**	864,886	0.389
	(0.005)	(0.005)	(0.005)	(0.011)	(0.007)		
22	0.007	-0.019	0.031	0.058*	-0.063***	458,612	0.392
	(0.015)	(0.015)	(0.022)	(0.035)	(0.022)		
23	-0.017**	0.024***	0.001	0.003	-0.041***	170,876	0.440
	(0.007)	(0.008)	(0.019)	(0.012)	(0.015)		
24	-0.027***	-0.005	-0.055***	-0.020***	-0.028***	687,599	0.448
	(0.004)	(0.005)	(0.005)	(0.007)	(0.006)		
31	-0.020***	0.013*	-0.061***	-0.049***	-0.016***	299,461	0.506
	(0.005)	(0.007)	(0.006)	(0.009)	(0.006)		
32	0.016	0.010	-0.051***	-0.037***	-0.029***	150,153	0.477
	(0.011)	(0.013)	(0.011)	(0.006)	(0.008)		
33	-0.042**	0.015	-0.032	0.015	-0.027	9,259	0.699
	(0.021)	(0.020)	(0.023)	(0.049)	(0.029)		
34	-0.022	0.009	-0.023	-0.010	-0.061*	103,482	0.644
	(0.029)	(0.031)	(0.044)	(0.070)	(0.031)		
35	-0.035***	0.010*	-0.051***	-0.019	-0.060***	1,305,530	0.463
	(0.006)	(0.006)	(0.011)	(0.013)	(0.009)		
41	-0.024***	0.015***	-0.056***	-0.043***	-0.013*	539,208	0.475
	(0.004)	(0.005)	(0.007)	(0.011)	(0.007)		
42	-0.018***	0.006	-0.045***	-0.039***	0.005	76,428	0.671
	(0.007)	(0.008)	(0.008)	(0.011)	(0.010)		
51	-0.011	0.048***	-0.043***	-0.016	-0.006	9,189	0.584
	(0.018)	(0.015)	(0.015)	(0.025)	(0.023)		
52	-0.018***	0.011	-0.018**	-0.015	0.011	351,376	0.495
	(0.006)	(0.008)	(0.009)	(0.016)	(0.015)		
53	0.024	0.080*	-0.008	0.013	-0.094	98,435	0.598
	(0.033)	(0.044)	(0.026)	(0.028)	(0.069)		
54	-0.035*	0.033**	0.012	0.001	0.012	271,633	0.378
	(0.019)	(0.015)	(0.013)	(0.019)	(0.036)		
61	0.017**	0.014**	-0.011***	-0.009	-0.028***	1,446,703	0.410
	(0.007)	(0.007)	(0.004)	(0.009)	(0.008)		
62	-0.170*	0.038**	-0.038*	-0.074**	-0.119*	130,177	0.538
	(0.102)	(0.018)	(0.022)	(0.035)	(0.061)		
71	-0.013	0.008	-0.036**	0.023	0.003	259,496	0.619
	(0.018)	(0.015)	(0.014)	(0.020)	(0.023)		
72	-0.009	0.024**	-0.035***	-0.067***	-0.024***	642,649	0.512
	(0.009)	(0.011)	(0.010)	(0.011)	(0.009)		

Continued on next page

Table A.2 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
81	0.002	-0.005	-0.017*	0.014	-0.005	72,579	0.652
	(0.008)	(0.012)	(0.009)	(0.020)	(0.016)		
82	0.004	-0.035**	-0.043**	0.085	-0.023	412,142	0.577
	(0.013)	(0.016)	(0.022)	(0.077)	(0.037)		
91	-0.021**	-0.012	0.000	0.017	-0.042**	47,292	0.483
	(0.010)	(0.015)	(0.013)	(0.039)	(0.018)		
92	-0.015	-0.000	-0.011	0.056*	0.019	807,273	0.543
	(0.013)	(0.008)	(0.014)	(0.033)	(0.030)		

Notes: Each row presents coefficients and standard errors (in parentheses) from separate regressions for each 2-digit SOC occupation. SOC codes follow the UK Standard Occupational Classification 2010. Full descriptions of occupation categories are available from the Office for National Statistics (www.ons.gov.uk/methodology/classificationsandstandards). "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. The dependent variable is the natural logarithm of posted salary. All regressions include controls for soft skills, hard skills, remote work, advertisement length, contract type, work schedule, seniority level, education, experience, amenities availability, and negotiable salary. Fixed effects for location (NUTS2 region), month-year, and firm are included. Standard errors are clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Personality Traits coefficients on Log posted wage, by 3-digit occupation

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
111	0.012	0.006	-0.040*	-0.017	0.038	6,293	0.837
	(0.025)	(0.026)	(0.021)	(0.039)	(0.045)		
112	-0.024***	-0.014	-0.033***	0.006	-0.037***	96,013	0.668
	(0.007)	(0.010)	(0.009)	(0.015)	(0.013)		
113	-0.019***	0.017***	-0.045***	0.001	-0.023**	117,821	0.472
	(0.006)	(0.006)	(0.006)	(0.011)	(0.011)		
115	0.003	0.049**	-0.021	0.039	-0.076**	6,636	0.697
	(0.018)	(0.023)	(0.018)	(0.039)	(0.031)		
116	-0.058**	0.010	0.018	-0.028	0.016	106,426	0.638
	(0.029)	(0.016)	(0.021)	(0.026)	(0.019)		
117	0.007	0.013	-0.059*	-0.190***	0.009	1,558	0.850
	(0.045)	(0.046)	(0.033)	(0.060)	(0.096)		
118	-0.022	0.031**	-0.027*	0.076*	-0.082***	24,316	0.481
	(0.013)	(0.015)	(0.014)	(0.043)	(0.027)		
119	-0.024*	-0.021	-0.002	0.013	0.028	26,336	0.729
	(0.015)	(0.017)	(0.031)	(0.027)	(0.031)		
121	0.005	-0.241***	0.126	-0.199***	-0.035	688	0.816
	(0.066)	(0.051)	(0.096)	(0.055)	(0.113)		
122	0.012	0.043**	0.002	-0.013	-0.083***	88,550	0.557
	(0.013)	(0.018)	(0.014)	(0.023)	(0.025)		

Continued on next page

Table A.3 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
124	-0.016	-0.049*	-0.040**	-0.029	-0.012	56,862	0.503
	(0.020)	(0.028)	(0.016)	(0.026)	(0.035)		
125	-0.016	0.020	-0.037***	-0.010	-0.060***	61,343	0.627
	(0.010)	(0.015)	(0.010)	(0.024)	(0.013)		
211	-0.025*	0.017*	-0.050***	0.012	-0.008	24,544	0.570
	(0.013)	(0.010)	(0.013)	(0.022)	(0.022)		
212	-0.006	0.009	-0.041***	0.005	-0.023	155,044	0.416
	(0.008)	(0.008)	(0.009)	(0.020)	(0.014)		
213	-0.027***	-0.018***	-0.049***	-0.006	-0.019**	657,904	0.379
	(0.006)	(0.006)	(0.005)	(0.011)	(0.008)		
214	-0.006	0.021	-0.057***	0.071**	0.115***	11,459	0.640
	(0.019)	(0.018)	(0.019)	(0.030)	(0.035)		
215	-0.046***	0.003	-0.066***	0.018	-0.072*	13,317	0.736
	(0.014)	(0.015)	(0.017)	(0.027)	(0.041)		
221	0.012	-0.016	-0.049	0.010	-0.153***	50,190	0.393
	(0.015)	(0.034)	(0.037)	(0.041)	(0.037)		
222	-0.012	-0.022	0.013	0.010	-0.024	33,197	0.420
	(0.017)	(0.014)	(0.012)	(0.018)	(0.019)		
223	0.013	0.000	0.049*	0.066*	-0.045**	374,840	0.470
	(0.017)	(0.013)	(0.025)	(0.038)	(0.018)		
231	-0.017**	0.024***	0.001	0.003	-0.041***	170,876	0.440
	(0.007)	(0.008)	(0.019)	(0.012)	(0.015)		
241	-0.025***	0.033**	-0.035***	-0.029***	0.007	39,213	0.524
	(0.007)	(0.014)	(0.007)	(0.009)	(0.014)		
242	-0.031***	0.003	-0.054***	-0.014*	-0.027***	422,880	0.430
	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)		
243	-0.022**	-0.007	-0.047***	-0.002	-0.018	74,139	0.479
	(0.009)	(0.011)	(0.009)	(0.019)	(0.015)		
244	0.007	0.017	-0.025**	-0.060***	-0.034*	34,789	0.716
	(0.012)	(0.012)	(0.012)	(0.023)	(0.018)		
245	-0.007	0.052	-0.074*	-0.060	0.076	871	0.830
	(0.037)	(0.038)	(0.039)	(0.058)	(0.054)		
246	-0.027***	-0.002	-0.043***	0.005	-0.031***	82,833	0.555
	(0.006)	(0.008)	(0.009)	(0.023)	(0.011)		
247	-0.015	0.001	-0.040***	0.012	-0.007	27,944	0.563
	(0.010)	(0.010)	(0.009)	(0.018)	(0.011)		
311	-0.003	-0.005	-0.052***	-0.010	0.015	112,425	0.544
	(0.008)	(0.010)	(0.010)	(0.015)	(0.010)		
312	-0.019	-0.002	-0.037**	-0.016	0.013	16,645	0.597
	(0.017)	(0.013)	(0.016)	(0.048)	(0.021)		
313	-0.031***	0.022***	-0.066***	-0.060***	-0.030***	168,785	0.531
	(0.005)	(0.008)	(0.007)	(0.011)	(0.006)		
321	0.011	0.043	-0.025	-0.049	-0.072	9,775	0.518

Continued on next page

Table A.3 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
	(0.034)	(0.062)	(0.034)	(0.039)	(0.050)		
323	0.018	0.008	-0.050***	-0.035***	-0.023***	140,227	0.470
	(0.011)	(0.013)	(0.011)	(0.006)	(0.008)		
331	-0.042**	0.015	-0.032	0.015	-0.027	9,259	0.699
	(0.021)	(0.020)	(0.023)	(0.049)	(0.029)		
341	-0.007	0.006	-0.020	0.023	-0.029	19,787	0.710
	(0.015)	(0.014)	(0.013)	(0.037)	(0.019)		
342	0.004	-0.089***	-0.055***	0.033	-0.019	30,049	0.673
	(0.012)	(0.015)	(0.013)	(0.024)	(0.017)		
344	-0.030	0.081	-0.008	0.016	-0.113	52,865	0.667
	(0.047)	(0.054)	(0.074)	(0.097)	(0.081)		
351	-0.015	0.161**	-0.188	0.136	0.404	849	0.897
	(0.110)	(0.080)	(0.128)	(0.255)	(0.288)		
352	-0.025	0.027	-0.045***	-0.049**	-0.020	20,679	0.541
	(0.016)	(0.017)	(0.012)	(0.022)	(0.019)		
353	-0.037***	0.026**	-0.058***	-0.054***	-0.054***	315,361	0.474
	(0.006)	(0.013)	(0.008)	(0.016)	(0.011)		
354	-0.033***	0.006	-0.048***	-0.006	-0.052***	642,188	0.502
	(0.008)	(0.007)	(0.012)	(0.014)	(0.013)		
355	0.025	0.044***	-0.097***	-0.019	0.018	2,914	0.718
	(0.015)	(0.016)	(0.035)	(0.015)	(0.031)		
356	-0.046***	-0.001	-0.030**	-0.014	-0.057***	316,191	0.498
	(0.008)	(0.011)	(0.013)	(0.023)	(0.013)		
411	0.045	0.004	-0.076***	-0.053	-0.010	6,719	0.742
	(0.030)	(0.026)	(0.028)	(0.047)	(0.033)		
412	-0.016***	0.025***	-0.058***	-0.040***	-0.039***	246,060	0.508
	(0.005)	(0.008)	(0.006)	(0.009)	(0.007)		
413	-0.032***	0.003	-0.036**	-0.025*	0.015	87,616	0.519
	(0.007)	(0.012)	(0.017)	(0.014)	(0.013)		
415	-0.014***	0.003	-0.028***	-0.026	-0.007	153,110	0.513
	(0.004)	(0.005)	(0.005)	(0.017)	(0.005)		
416	-0.029***	-0.005	-0.065***	-0.032**	-0.008	38,232	0.731
	(0.010)	(0.009)	(0.012)	(0.014)	(0.014)		
421	-0.018***	0.006	-0.045***	-0.039***	0.005	76,428	0.671
	(0.007)	(0.008)	(0.008)	(0.011)	(0.010)		
511	-0.011	0.048***	-0.043***	-0.016	-0.006	9,189	0.584
	(0.018)	(0.015)	(0.015)	(0.025)	(0.023)		
521	-0.008	-0.009	-0.026	-0.095	-0.036	15,616	0.583
	(0.014)	(0.020)	(0.018)	(0.068)	(0.031)		
522	-0.002	-0.007	0.011	-0.023	0.048*	99,264	0.536
	(0.006)	(0.010)	(0.017)	(0.026)	(0.025)		
523	-0.019	0.020	0.004	-0.026	0.010	57,531	0.543
	(0.013)	(0.021)	(0.024)	(0.050)	(0.028)		

Continued on next page

Table A.3 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
524	-0.018**	0.019**	-0.032***	-0.019	-0.022	173,056	0.474
	(0.008)	(0.008)	(0.007)	(0.015)	(0.022)		
525	0.007	0.008	-0.018	-0.030	-0.069***	3,912	0.680
	(0.018)	(0.030)	(0.027)	(0.030)	(0.025)		
531	0.030	0.106**	-0.006	0.016	-0.095	80,652	0.599
	(0.038)	(0.049)	(0.032)	(0.030)	(0.083)		
532	0.003	-0.072	-0.055**	-0.028	-0.004	8,870	0.664
	(0.020)	(0.054)	(0.024)	(0.046)	(0.057)		
533	-0.013	-0.021	-0.002	-0.066	-0.005	8,331	0.712
	(0.021)	(0.028)	(0.019)	(0.046)	(0.030)		
541	-0.013	0.151**	-0.044	0.029	-0.091	1,309	0.706
	(0.022)	(0.068)	(0.040)	(0.030)	(0.069)		
542	-0.028	0.011	-0.046**	-0.017	0.008	2,761	0.758
	(0.025)	(0.016)	(0.021)	(0.100)	(0.020)		
543	-0.034*	0.031**	0.011	0.002	0.010	262,994	0.373
	(0.020)	(0.015)	(0.013)	(0.019)	(0.037)		
544	-0.071***	0.049*	0.057**	-0.009	0.111	4,325	0.772
	(0.019)	(0.026)	(0.026)	(0.069)	(0.070)		
612	0.002	0.008	-0.019**	-0.010	-0.031***	147,936	0.503
	(0.011)	(0.008)	(0.009)	(0.013)	(0.010)		
613	-0.023	0.002	-0.101**	-0.021	-0.085*	5,886	0.636
	(0.035)	(0.021)	(0.046)	(0.038)	(0.043)		
614	0.018**	0.015**	-0.010**	-0.011	-0.027***	1,292,239	0.401
	(0.008)	(0.008)	(0.004)	(0.010)	(0.008)		
621	-0.029	0.033*	-0.061**	-0.024	-0.009	43,617	0.611
	(0.019)	(0.018)	(0.030)	(0.022)	(0.013)		
622	0.010	-0.010	0.004	-0.102**	0.001	13,135	0.574
	(0.021)	(0.036)	(0.022)	(0.049)	(0.041)		
623	-0.011	0.002	-0.013	0.004	-0.008	19,812	0.674
	(0.010)	(0.016)	(0.010)	(0.020)	(0.013)		
624	-0.312**	0.096**	-0.022	-0.131**	-0.227***	52,889	0.639
	(0.124)	(0.039)	(0.032)	(0.057)	(0.067)		
711	-0.020	-0.015	-0.005	0.015	0.006	194,211	0.579
	(0.017)	(0.012)	(0.014)	(0.022)	(0.025)		
712	-0.064***	-0.008	-0.039*	-0.022	-0.009	43,355	0.796
	(0.019)	(0.025)	(0.022)	(0.027)	(0.026)		
713	0.009	0.054*	-0.077***	-0.012	0.037	21,121	0.778
	(0.026)	(0.029)	(0.027)	(0.027)	(0.047)		
721	-0.012	0.001	-0.006	-0.040***	-0.010	438,368	0.518
	(0.009)	(0.009)	(0.010)	(0.009)	(0.007)		
722	-0.017	0.055***	-0.051***	-0.061***	-0.025	202,452	0.605
	(0.012)	(0.015)	(0.012)	(0.017)	(0.020)		
811	0.010	0.094**	0.065***	0.062	0.043	3,155	0.719

Continued on next page

Table A.3 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
	(0.022)	(0.041)	(0.024)	(0.138)	(0.042)		
812	0.001	-0.014	-0.017	0.010	0.013	8,122	0.621
	(0.018)	(0.020)	(0.015)	(0.052)	(0.027)		
813	-0.024**	-0.034**	-0.027**	-0.001	-0.013	21,095	0.737
	(0.011)	(0.015)	(0.013)	(0.034)	(0.020)		
814	0.019*	0.011	-0.010	0.005	-0.004	39,571	0.665
	(0.011)	(0.020)	(0.014)	(0.025)	(0.026)		
821	0.009	-0.042**	-0.031	0.123*	-0.053	363,254	0.599
	(0.014)	(0.017)	(0.023)	(0.072)	(0.045)		
822	-0.001	-0.006	-0.016	-0.043	0.022	39,867	0.624
	(0.009)	(0.021)	(0.012)	(0.041)	(0.025)		
823	0.027	-0.023	0.009	-0.020	0.016	8,332	0.663
	(0.017)	(0.020)	(0.014)	(0.034)	(0.020)		
911	-0.024	0.016	-0.049**	0.052	-0.026	3,114	0.710
	(0.022)	(0.054)	(0.023)	(0.063)	(0.042)		
912	-0.034**	-0.007	0.058***	0.053*	-0.047	27,677	0.371
	(0.015)	(0.036)	(0.022)	(0.030)	(0.070)		
913	-0.002	-0.007	0.004	-0.037	-0.036**	16,346	0.610
	(0.012)	(0.012)	(0.015)	(0.032)	(0.017)		
921	0.031	-0.027**	0.085***	0.121***	-0.039	55,434	0.728
	(0.023)	(0.013)	(0.031)	(0.032)	(0.036)		
923	0.006	-0.031***	-0.030**	0.009	-0.001	189,567	0.609
	(0.011)	(0.011)	(0.014)	(0.011)	(0.020)		
924	-0.018**	0.003	-0.019*	-0.006	-0.014	131,362	0.761
	(0.009)	(0.008)	(0.011)	(0.013)	(0.010)		
925	-0.042	0.189***	-0.044	-0.070	-0.151***	1,590	0.819
	(0.060)	(0.071)	(0.065)	(0.074)	(0.058)		
926	-0.016	0.012	-0.016	-0.068**	-0.020	151,947	0.403
	(0.013)	(0.026)	(0.022)	(0.034)	(0.012)		
927	0.028***	0.007	-0.038**	0.027	-0.060**	275,105	0.417
	(0.010)	(0.011)	(0.016)	(0.028)	(0.023)		

Notes: Each row presents coefficients and standard errors (in parentheses) from separate regressions for each 3-digit SOC occupation. SOC codes follow the UK Standard Occupational Classification 2010. Full descriptions of occupation categories are available from the Office for National Statistics (www.ons.gov.uk/methodology/classificationsandstandards). "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. The dependent variable is the natural logarithm of posted salary. All regressions include controls for soft skills, hard skills, remote work, advertisement length, contract type, work schedule, seniority level, education, experience, amenities availability, and negotiable salary. Fixed effects for location (NUTS2 region), month-year, and firm are included. Standard errors are clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Agreeableness intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	-0.005	-0.017	-0.003	-0.016	-0.010
		(0.012)	(0.010)	(0.011)	(0.021)	(0.018)
	High-value firms	0.005	-0.005	0.019	-0.010	-0.028
		(0.013)	(0.017)	(0.018)	(0.021)	(0.034)
Profit						
	Average firms	-0.004	-0.018	-0.006	-0.016	-0.012
		(0.013)	(0.011)	(0.012)	(0.020)	(0.020)
	High-value firms	0.009	-0.020	-0.003	-0.003	0.001
		(0.021)	(0.023)	(0.029)	(0.018)	(0.018)
Firm Size						
	Average firms	-0.002	-0.016	-0.005	-0.017	-0.011
		(0.011)	(0.011)	(0.011)	(0.020)	(0.021)
	High-value firms	0.035***	-0.005	0.002	-0.024	0.002
		(0.013)	(0.013)	(0.015)	(0.024)	(0.022)
Publicly Quoted						
	Non-quoted firms	-0.004	-0.015	-0.003	-0.014	-0.011
		(0.013)	(0.012)	(0.011)	(0.022)	(0.021)
	Quoted firms	-0.011	-0.057	-0.092***	-0.096*	-0.069**
		(0.031)	(0.050)	(0.033)	(0.051)	(0.030)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185,619 (SD=£612,396); Firm Size: mean=21,845 employees (SD=87,181).

Table A.5: Conscientiousness intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	-0.004	-0.007	-0.027**	-0.014	-0.014
		(0.010)	(0.014)	(0.011)	(0.011)	(0.014)
	High-value firms	0.003	-0.017	-0.017	0.006	-0.025*
		(0.012)	(0.017)	(0.015)	(0.014)	(0.013)
Profit						
	Average firms	-0.005	-0.008	-0.027**	-0.013	-0.014
		(0.011)	(0.014)	(0.011)	(0.011)	(0.014)
	High-value firms	0.001	-0.012	-0.029*	-0.011	-0.004
		(0.012)	(0.016)	(0.017)	(0.016)	(0.016)
Firm Size						
	Average firms	0.000	-0.007	-0.027**	-0.011	-0.015
		(0.011)	(0.015)	(0.011)	(0.011)	(0.013)
	High-value firms	0.030*	-0.006	-0.028**	0.001	-0.020
		(0.016)	(0.019)	(0.012)	(0.013)	(0.012)
Publicly Quoted						
	Non-quoted firms	-0.005	-0.004	-0.018*	-0.007	-0.013
		(0.011)	(0.014)	(0.010)	(0.010)	(0.014)
	Quoted firms	-0.021	-0.055*	-0.163***	-0.100**	-0.026
		(0.027)	(0.032)	(0.042)	(0.048)	(0.043)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185,619 (SD=£612,396); Firm Size: mean=21,845 employees (SD=87,181).

Table A.6: Openness intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	0.010	0.001	0.007	0.005	-0.006
		(0.011)	(0.007)	(0.010)	(0.011)	(0.013)
	High-value firms	0.001	0.001	0.010	0.008	0.004
		(0.016)	(0.009)	(0.012)	(0.014)	(0.015)
Profit						
	Average firms	0.008	0.001	0.007	0.005	-0.007
		(0.011)	(0.007)	(0.010)	(0.011)	(0.013)
	High-value firms	0.016	-0.002	0.009	0.006	0.008
		(0.014)	(0.012)	(0.012)	(0.011)	(0.011)
Firm Size						
	Average firms	0.010	0.001	0.007	0.004	-0.000
		(0.011)	(0.007)	(0.010)	(0.010)	(0.013)
	High-value firms	0.027**	-0.005	0.008	-0.004	0.035*
		(0.013)	(0.007)	(0.010)	(0.010)	(0.018)
Publicly Quoted						
	Non-quoted firms	0.011	0.003	0.008	0.007	-0.006
		(0.012)	(0.007)	(0.010)	(0.010)	(0.014)
	Quoted firms	-0.005	-0.027	-0.002	-0.012	-0.022
		(0.039)	(0.022)	(0.038)	(0.037)	(0.043)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185,619 (SD=£612,396); Firm Size: mean=21,845 employees (SD=87,181).

Table A.7: Extraversion intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	-0.011	-0.041***	-0.030**	-0.034**	-0.045***
		(0.011)	(0.012)	(0.013)	(0.015)	(0.013)
	High-value firms	-0.024	-0.068***	-0.034***	-0.038***	-0.023*
		(0.017)	(0.015)	(0.012)	(0.014)	(0.013)
Profit						
	Average firms	-0.011	-0.041***	-0.030**	-0.034**	-0.045***
		(0.011)	(0.013)	(0.013)	(0.015)	(0.013)
	High-value firms	-0.012	-0.062***	-0.027**	-0.033***	-0.045***
		(0.014)	(0.016)	(0.010)	(0.011)	(0.013)
Firm Size						
	Average firms	-0.010	-0.045***	-0.028**	-0.034**	-0.045***
		(0.011)	(0.014)	(0.012)	(0.015)	(0.013)
	High-value firms	-0.008	-0.061***	-0.017	-0.036***	-0.048***
		(0.012)	(0.017)	(0.010)	(0.013)	(0.012)
Publicly Quoted						
	Non-quoted firms	-0.013	-0.039***	-0.029**	-0.034**	-0.045***
		(0.012)	(0.014)	(0.014)	(0.016)	(0.014)
	Quoted firms	0.005	-0.101***	-0.047	-0.025	-0.042
		(0.038)	(0.037)	(0.029)	(0.055)	(0.033)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185,619 (SD=£612,396); Firm Size: mean=21,845 employees (SD=87,181).

Table A.8: Emotional Stability intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	-0.021	-0.024*	-0.032**	0.008	-0.032**
		(0.018)	(0.014)	(0.014)	(0.022)	(0.013)
	High-value firms	-0.025	-0.031**	-0.025	0.016	-0.052***
		(0.018)	(0.012)	(0.021)	(0.029)	(0.016)
Profit						
	Average firms	-0.020	-0.024*	-0.031**	0.008	-0.032**
		(0.018)	(0.014)	(0.014)	(0.022)	(0.013)
	High-value firms	-0.025	-0.020	-0.017	-0.000	-0.056***
		(0.018)	(0.017)	(0.027)	(0.021)	(0.015)
Firm Size						
	Average firms	-0.018	-0.020	-0.029**	0.008	-0.032**
		(0.017)	(0.014)	(0.013)	(0.023)	(0.015)
	High-value firms	0.008	0.010	-0.000	0.011	-0.038**
		(0.017)	(0.014)	(0.019)	(0.030)	(0.018)
Publicly Quoted						
	Non-quoted firms	-0.015	-0.022	-0.033**	0.011	-0.027*
		(0.019)	(0.015)	(0.013)	(0.023)	(0.014)
	Quoted firms	-0.079**	-0.066*	-0.047	-0.047**	-0.101**
		(0.031)	(0.035)	(0.049)	(0.022)	(0.041)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185,619 (SD=£612,396); Firm Size: mean=21,845 employees (SD=87,181).

Table A.9: Effects of personality trait intensity quintiles on wages

Intensity Measure	Any P.T.	Agree.	Consc.	Emot.	Extraversion	Open.
Quintile 1 (Lowest Intensity)	-0.024***	0.022	-0.005	-0.015**	-0.004	0.008
	(0.007)	(0.018)	(0.007)	(0.007)	(0.011)	(0.006)
Quintile 2	-0.039***	-0.005	-0.011	-0.053***	-0.038***	0.009*
	(0.008)	(0.010)	(0.009)	(0.018)	(0.008)	(0.005)
Quintile 3	-0.053***	-0.031***	-0.034***	-0.031***	-0.038***	0.007
	(0.007)	(0.008)	(0.013)	(0.009)	(0.008)	(0.005)
Quintile 4	-0.053***	-0.036***	-0.024**	-0.014	-0.036***	0.011*
	(0.009)	(0.008)	(0.010)	(0.019)	(0.008)	(0.006)
Quintile 5 (Highest Intensity)	-0.071***	-0.024	-0.026**	-0.043***	-0.045***	0.000
	(0.010)	(0.018)	(0.011)	(0.006)	(0.009)	(0.010)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Baseline is *no trait mentioned* (trait intensity = 0). Regressions control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 9,888,315 observations.

Table A.10: Effects of trait frequency categories on wages

Trait - Number of mentions	1	2	3	4	5+
Agreeableness	-0.014*	-0.018	0.014	-0.031	0.078
	(0.008)	(0.022)	(0.051)	(0.032)	(0.105)
Conscientiousness	-0.017**	-0.033**	-0.029**	-0.045***	-0.058***
	(0.007)	(0.013)	(0.011)	(0.016)	(0.017)
Emotional Stability	-0.032***	-0.020	-0.075***	-0.067**	-0.101***
	(0.006)	(0.025)	(0.017)	(0.026)	(0.020)
Extraversion	-0.028***	-0.037***	-0.059***	-0.072***	-0.092***
	(0.007)	(0.008)	(0.010)	(0.015)	(0.012)
Openness	0.004	0.018**	0.022*	0.027*	-0.041
	(0.004)	(0.007)	(0.013)	(0.015)	(0.073)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Baseline is no trait mention (trait frequency = 0). Regressions control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 9,888,315 observations.

Appendix B - Amenities' dictionaries

Category	Words searched
A00: General Benefits	
General Benefits	rewardscheme, flexiblebenefit, benefitsincluding, generousbenefits, benefitsplatform, comprehensivebenefits, staffbenefits, benefitsscheme, employeebenefits, competitivebenefits, fantasticbenefits, greatbenefits, excellentbenefits, companybenefits, rangebenefits, standardbenefits, corebenefits, variousbenefits
Benefits Packages	benefitspackage, greatpackage, rewardspackage, salarypackage, remunerationpackage
Recognition Schemes	recognitionschemes, employeemonthaward, monthlyawards, employeerecognition
Referral Schemes	referralscheme, referfriendincentive, referralprogramme, staffreferral, employeereferral, referfriendscheme
Tax Benefits	taxefficientscheme, taxefficientsavings, familytax, taxfreecycle
Loans Facilities	interestfreeloans, mortgagefacility, mileagebank
Relocation Support	relocationpackage, relocationallowance, visarelocation, sponsorshipvisa, visasponsorship, visacosts
Seasonal Benefits	christmas
A01: Bonuses and Commissions	
Basic Compensation	bonus, commission, incentive
Equity Compensation	stockoptions, sharesaving, sharescheme, sharesave, shareincentiveplan
Profit Sharing	profitsharing, profitshare
Performance Pay	performancerelated, payforperformance, performancerecognised, productivitypay
Sales Compensation	salesquota, salesgoal, salestarget
Premium Pay	premiumpay, salarypremium, extrapay, variablepay
Special Payments	paycash, averagepayout, tipsshared, annualpay, anniversaryreward
Savings Schemes	saveearnscheme, savingsscheme
Promotion Based	promotionbased
A02: Overtime Paid	
Overtime	extrahour, overtime, longhour, extrashift, doublepay, unsocialhour
A03: Paid Time Off	
Vacation Holidays	holiday, vacation, freeday
Sick Leave	sickday, sickleave, sicktime, sicknesspay, sickpay, sickdays, sickchild
Medical Leave	healthleave, medicalleave, medicalindemnity
Maternity Benefits	maternitypay, maternityleave, maternityday, maternitypaternity, maternitypolicies, maternitypackage, enhancedmaternity
Paternity Benefits	paternitypay, paternityleave, paternityday
Family Leave	parentalleave, parentaladoption, familypolicies, familydays, adoptionleave
Personal Time	personalday, paidtimeoff, paidleave, paidrest, paidbreak, dayspaid
Special Leave	bereavementleave, trainingleave, educationleave, studyleave, compensationleave, plannedabsence, movinghouseday, birthday, resetdays, earlyfriday

Continued on next page

Table B.1 – continued from previous page

Category	Words searched
Volunteer Time	communityvolunteering, annuallyvolunteer, volunteertime, volunteerday, volunteeringday, charitableday, charityprojects
Flexible Time	summerfridays, optionbuy
Wellbeing Time	wellbeingbenefits, wellbeingdays, cashhealthplan, healthcaredentalcare
A04: Health Insurance	
Health Insurance	healthinsurance, offerhealthinsurance, privatehealth, privatehealthinsurance, healthcarescheme, privatemedical, healthplan, medicalplan, healthcareplan, medicalbenefit, medicalcover, medicalinsurance
Dental Vision	dentalbenefit, dentalinsurance, offerdentalinsurance, dentalcoverage, dentalplan, dentalcare, dentaltreatment, dentalcover, includesdental, subsidiseddental, dentalvision, freedental, opticaldental, visioninsurance, visioncoverage
Life Insurance	lifeinsurance, lifecover, lifeaccidentinsurance
Accident Insurance	offeraccidentinsurance, includeaccidentinsurance, personalaccidentinsurance, travelaccidentinsurance
Health Coverage	healthcoverage, healthcontribution, familycoverage, familyhealthcare, healthwellbeing
Free Healthcare	freehealth, freemedical, freehealthcare, freelegaladvice
Health Savings	healthsavings, cashplan, medicash
Wellness Programs	wellnessprogramme, freefluvaccine, freeeyetest, onsiteflushots, annualflu, fluvouchers
Childcare Support	childcarescheme, childcarevouchers
Eye Care	eyecarecover, eyecarediscount, eyecarevoucher
Specialised Coverage	criticalillnesscover, criticalillnessscheme, disabilitycover, illnessinsurance, indemnitycover, mentalhealthfirst
Insurance Schemes	privateinsurance, discountedinsurance, welfarepackage, healthcover
A05: Retirement Contributions	
Pension Schemes	pension
Retirement Plans	retirementcontribution, retirementfund
Employer Contributions	employercontribution, employermatch, companymatch, twicecontribute, doublecontribution
Savings Accounts	savingsaccount, savingscontribution, savingsscheme
Provident Funds	providentfund
Education Support	tuitionreimbursement
Life Benefits	lifeassurance
A06: Food Subsidies and Discounts	
Breakfast Benefits	breakfastincluded, caterbreakfast, freebreakfast, freedailybreakfast, unlimitedbreakfast
Lunch Benefits	freelunch, lunchincluded, caterlunch, lunchallowance, lunchticket
Meal Benefits	freemeal, mealincluded, mealprovided, catermeal, mealsduty, mealsprovided, subsidisedmeal, onshiftmeals, workdaylunches

Continued on next page

Table B.1 – continued from previous page

Category	Words searched
Food General	freefood, foodincluded, foodprovided, foodcoupon, foodstamp, foodsubsidy, foodticket
Dinner Benefits	freedinner, dinnerincluded
Drinks Beverages	freedrinks, freecoffee, coffeetea, teacoffee, freebeer, freebeverages, freetea, drinksavailable, drinksfacilities, drinksfridge, freehotdrinks, fridaysdrinks, beerfridays, regulardrinks
Fruit Snacks	freefruit, freshfruit, fruiteveryday, fruitrefreshments, fruitsnacks, dailyfruit, healthysnacks, afternoonsnacks, officesnacks, freeicecream, freecake, freesoft, freeunhealthy, freehealthy
Gym Fitness	gymdiscount, gymmembership, gympass, gymsubsidies, gymnearby, fitnessclasses, freeyoga, swimmembership, yogasessions, weeklyyoga
Transport Benefits	freetransportation, transportationallowance, travelsubsidy, commutesubsidy, transitsubsidy, transportationsubsidy, transportvoucher, travelstipend
Discount Schemes	employeediscount, companydiscount, corporatediscount, staffdiscount, storediscount, retaildiscount, cardiscount, discountcar, discountcard, discountprogramme, discountschemes, discounttravel, discountvarious, discountedcinema, discountedfitness, discountedfood, discountedmembership, discountedstays, discountedsubscription, dealsdiscount, leisurediscount, hoteldiscount, traveldiscount, onlinediscount, instorediscounts, discountlocal
Clothing Uniform	clothesdiscount, companyuniform, comfortableuniform
Accommodation	liveposition, housingincluded, onsiteaccommodation, rentalsubsidy
Vouchers Rewards	cashback, seasontickets, shoppingvouchers, cinematickets, loyaltyrewards, rewardspoints, freetickets, voluntarybenefits
Wellness Perks	wellbeingoffers, wellnessportal, accessheadspace, headspacesubscription, meditationapp
Vehicle Benefits	carleasing, carloanscheme, fuelcard, leasingdeals
Platform Benefits	benefitsplatform, benefitsportal, advantagescheme, leasescheme, huboffers, onlinebenefits, rewardplatform, shoppingportal
Leisure Entertainment	pooltables, personalparcels, freeuse, leisuresavings
Coffee Facilities	excellentcoffeemachine, nespressocoffeemachine
Refreshments	refreshmentsprovided, dailymenu, subsidisedrate
Insurance Extras	bikeinsurance, travelinsurance, travelbenefits
Special Services	deathservice
A07: Office Space and Amenities	
Fitness Facilities	fitnessroom, officegym, onsitefitness, onsitegym, fitnessstudio, officeyoga
Recreational Spaces	gameroom, gamesarea, gamesroom, xboxroom, tabletennis, tennisroom, relaxationarea, breakarea, breakrooms, breakoutarea, chilloutarea
Dining Facilities	companycafeteria, onsitecafeteria, onsitecatering, onsiterestaurant, staffrestaurant, canteen, onsitebistro, stockedkitchen, subsidisedrestaurant

Continued on next page

Table B.1 – continued from previous page

Category	Words searched
Office Environment	comfortableoffice, comfortablespace, modernoffice, brandnewoffice, refurbishedoffice, greatoffices, openspace, openplanenvironment, openplanoffice, naturallight, greatviews, openroof
Specialised Rooms	prayerroom, wellnessroom, vapingarea, onsitecinema, onsitenursery
Work Facilities	conferenceroomfacilities, quietworkspaces, quietworkingspaces, coworkingspace
Facility Quality	amenities, facilitiesonsite, onsitefacilities, goodfacilities, greatfacilities, excellentfacilities, fantasticfacilities, modernfacilities, stateartfacilities, rangefacilities, comfortablefacilities, leisurefacilities, sportfacilities
Workplace Description	amazingworkplace, officeincludes, officemodern, developedoffice, stateartbuilding
Outdoor Spaces	roofterrace, rooftopterrace
Service Facilities	showerfacilities, showerfacility, onsiteshower, onsiterestroom, selfservicefacilities, usemicrowave, onsitesubsidised, onsiteswimming
Storage Amenities	stafflockers, staffrestarea
Location Features	locationalsoprovides, locationprovides
A08: Location and Commuting	
Location Quality	centrallocation, excellentlocation, goodlocation, greatlocation, finelocation, ideallocation, perfectlocation, peacefullocation, welllocated
City Center Access	citycenter, towncenter, basedheart, situatedheart, officeheart
Accessibility	easyaccess, easilyaccessible, convenientlylocated, quickaccess, easiestreach, closeaccess
Commuting	shortcommute, longcommute, easycommute, easilycommute, easilycommutable, perfectcommuting, easyreach
Parking Facilities	privateparking, payparking, companygarage, privategarage, carpark, onsitepark, carparking, freeparking, freecarparking, secureparking, siteparking, staffparking, streetparking, parkingavailable, parkingspace, localparking
Public Transport	accesspublictransport, publictransport, reliablepublic, viapublictransport, greattransport, transportlinks, transportationlinks, linkstransport, commuterlinks
Walking Distance	shortwalk, minutewalk, minuteswalk, walkingdistance, walkabledistance, minwalk, withinshortdistance
Cycling Facilities	bike2work, bikesheds, bikestorage, bikeworkscheme, cyclescheme, cycletowork, cyclework, cyclingscheme, taxfreebike, ride2work
Local Transport	buslinks, bustrain, localbus, localtransport, freebus, freelocalbus, discountedbus
Train Stations	nextcentralstation, mainlinestation, metrostationclose, metrostation
Location Context	basedcentre, centrelocated, centrelocation, situatedclose, minutescentral, outsidecitycentre, officenext
Nearby Amenities	localshops, nearlocal, childcareoptions
A09: Work Equipment and Allowances	

Continued on next page

Table B.1 – continued from previous page

Category	Words searched
Remote Work Support	internetcoupon, internetsubsidy, technologysubsidy, homeofficefund, teleworkfund, teleworkstipend, workhomefund, teleworkexpenses, homeofficeexpenses, workhomeexpenses, homeofficeallowance, teleworkallowance, workhomeallowance, teleworksubsidy, workhomesubsidy, homeofficeequipment, teleworkequipment, freebroadband, discountedbroadband, discountsenergybills
Mobile Devices	offermobilephone, includemobilephone, companyphone, freemobile, phoneandlaptop, phonelaptop, phoneprovided, phoneinsurance, phonework, offercompanycellphone, includecompanycellphone
Computing Equipment	offercompanycomputer, includecompanycomputer, offercompanylaptop, includecompanylaptop, companylaptop, twomonitors, twoscreens, dedicateddesk
Company Vehicles	companycar, companyvan, companyvehicle, vehicleallowance, vehiclescheme
Vehicle Expenses	fuelcard, fuelallowance, paidmileage, mileageexpense, carallowance
Company Cards	companycreditcard, companydebitcard
Clothing Equipment	companyuniform, freeuniform, uniformprovided, uniformprovision, provisionuniform, corporateclothing, protectivefootwear, protectiveequipment
Work Equipment	equipmentprovided, necessaryequipment
Perks Boxes	perkbox, perks, perbox, welcomebox
A10: Work Schedule Flexibility	
Flexible Scheduling	flexibleshift, flexibletime, flexiblehour, flexibleschedule, flxshift, flxtime, flxhour, flxschedule, flexibleofficehours, flxoption, flexibleoption
Work Life Balance	worklifebalance, lifebalance, lifework, healthywork, work.life, weeklyrest, familyfriendly
Flexible Work	flexiblework, flxwork, flexibleworking, modernworking, flexibilitywork, flexibilityaroundworking
A11: Workplace Safety and Ergonomics	
Workplace Safety	safeplace, worksafety, safework, safeworking, workplacesafety
Safety Culture	safetyculture, safetyfirst, employeesafety
Security	secureoffices, secureworkplace
Health Safety	healthsafety, awarehealth, focushealth
Safety Compliance	complysafe, complyhealth, compliancehealth, understandhealth, requirementshealth
Safety Regulations	safetylegislation, act1974, fireregulation, safetyact
Safety Policies	safetyrules, safetypolicies
Office Safety	safeoffice, hygienicworking
A12: Job Security	
Job Security	jobsecurity, jobstability
Company Stability	lowturnover, stableenvironment
Severance Benefits	severancepay, severancepackage, layoffcompensation
Unemployment Support	unemploymentinsurance, unemploymentbenefit

Continued on next page

Table B.1 – continued from previous page

Category	Words searched
Income Protection	incomeprotection, indemnityinsurance
A13: Work Environment and Impact on Society	
Workplace Culture	benefitculture, goodculture, companyculture, internalculture, collaborativeculture, positiveculture, workingculture, officeculture
Work Environment	bestworkplace, excellentworkplace, greatworkplace, amazingworkplace, dynamicworkplace, excellentenvironment, goodenvironment, warmenvironment, youngenvironment, multiculturalenvironment, relaxedenvironment, collaborativeenvironment, safeenvironment, socialenvironment, stimulatingenvironment, supportiveenvironment, uniqueenvironment, vibrantworkplace, workenvironment, workingenvironment
Workplace Atmosphere	goodatmosphere, relaxedatmosphere, lovelyatmosphere, casualdress, dresscode, vibrantoffice, sociableoffice, livelyoffice, pleasantoffice
Work Satisfaction	comfortwork, congenialwork, enjoyablework, meaningfulwork, positivework, relaxwork, healthywork, jobrecognition, jobsatisfaction, satisfactionjob, rewarding, fulfilling
Company Values	humanvalues, missionvision, companyvalues, values, promoteequality, valuediversity, responsiblecompany, culturematter
Social Activities	afterworkdrinks, companyevents, companyoutings, companyparties, companyretreats, companysocialevents, companytrips, funnysocialevents, gamesnights, regularsocialevents, regularsocials, socialactivities, socialgatherings, staffevents, teambonding, teambuilding, teamlunch, teamevents, outdooractivities
Company Reputation	bestplaceswork, bestreputation, localreputation, strongcompany
Employee Support	employeeassistance, employeesupport, freecounselling, wellbeingsupport, wellbeingworkshops, prioritisingwellbeing
Work Quality	enjoyableworking, excellentworkplace, fantasticworking, friendlyworking, greatworking, inclusiveworking, supportiveworking
Impact Purpose	hugeimpact, immediateimpact, impactsociety, differencelife, makedifference, makerealdifference, makingimpact, meaningfulproject, peoplelives, positiveimpact, withincommunity, impactworld, makeimpact, realimpact, significantimpact, strongimpact, positivechange, directimpact
Diversity Inclusion	multicultural, equalopportunities, valuemember, feelpart
Team Environment	interdisciplinary, multidisciplinary, dynamicteam, youngteam, greatteam, excitingteam, goodvibes
Work Environment Type	socialworkenvironment, workingethos
A14: Human Capital Development	
Career Growth	buildcareer, movecareerforward, career, progress
Skill Development	buildskills, developingskills, learningdevelopment, personaldevelopment, professionaldevelopment, developmentsupport

Continued on next page

Table B.1 – continued from previous page

Category	Words searched
Learning Opportunities	developmentopportunities, developmentopportunity, growthopportunities, opportunitieslearning, externallearning, continuousopportunities, opportunitiesgrowing, opportunitiesupskill
Company Growth	companygrowth, growingorganisation, growingcompany, growthcompany, growingbusiness, teamexpansion, growingteam
Training Education	training, elearningportal, learningallowance, learningzone, educationsubsidy, payeducation, course
Promotion Advancement	promoteinternally, promotionopportunities, promotionprospects, promotionprocess, promotiondevelopment, meritpromotion, salaryincrease
Merit Based	basemerit, meritincrease, meritbased
Personal Growth	personalgrowth, professionalgrowth, ongoinggrowth, growthopportunity, growthpossible, continuousimprovement, continuingprofessional
A15: Remote Work Opportunities	
Remote Work	conductedremote, fullyremote, homebased, homework, homeworking, jobisremote, jobisvirtual, locationisremote, locationremote, positionisremote, positionremote, remotework, rolewillberemote, telecommuting, telework, virtualjob, workfromhome, workhome, workremote, workingremote, flxwithoffice, remoteposition, roleisremote, remotebasis, hybridwork, hybridremote, workanywhere

Words were scrutinised as not to retrieve false positive words. Each word's context was analysed by sampling the dataset of 100 advertisements per word. When the word's context was not related to personality traits and a straightforward amendment of the word was not possible, this was removed from the dictionary. Words that did not retrieve any match were also not considered.

Before searching for words, modifications were made to some words to avoid false positives and account for different variations of words. The list of changes done to the text is the following:

01 General benefits and packages

"programe" for "programme", "rewards packages" for "rewardscheme", "rewards scheme" for "rewardscheme", "rewards" for "rewardscheme ", "reward" for "rewardscheme ", "flexible benefit" for "flexiblebenefit", "generous benefits" for "generousbenefits", "benefits platform" for "benefitsplatform", "comprehensive benefits" for "comprehensivebenefits", "staff benefits" for "staffbenefits", "recognition schemes" for "recognitionschemes", "benefits scheme" for "benefitsscheme", "employee benefits" for "employeebenefits", "competitive benefits" for "competitivebenefits", "fantastic benefits" for "fantasticbenefits", "great benefits" for "greatbenefits", "excellent benefits" for "excellentbenefits", "company benefits" for "companybenefits", "benefits package" for "benefitspackage", "great package" for "greatpackage", "rewards package" for "rewardspackage", "salary package" for "salarypackage", "remuneration package" for "remunerationpackage", "referral scheme" for "referralscheme", "range benefits" for "rangebenefits",

”standard benefits” for ”standardbenefits”, ”core benefits” for ”corebenefits”, ”benefits including” for ”benefitsincluding”, ”various benefits” for ”variousbenefits”, ”christmas” for ”christmas”, ”refer friend incentive” for ”referfriendincentive”, ”mileage bank” for ”mileagebank”, ”interest free loans” for ”interest-freeloans”, ”tax efficient scheme” for ”taxefficientscheme”, ”tax efficient savings” for ”taxefficientsavings”, ”family tax” for ”familytax”, ”mortgage facility” for ”mortgagefacility”, ”referral programme” for ”referralprogramme”, ”staff referral” for ”staffreferral”, ”tax free cycle” for ”taxfreecycle”, ”employee referral” for ”employeereferral”, ”refer friend scheme” for ”referfriendscheme”, ”employee month award” for ”employeemonthaward”, ”monthly awards” for ”monthlyawards”, ”visa costs” for ”visacosts”, ”relocation package” for ”relocationpackage”, ”relocation allowance” for ”relocationallowance”, ”visa relocation” for ”visarelocation”, ”sponsorship visa” for ”sponsorshipvisa”, ”visa sponsorship” for ”visasponsorship”, ”employee recognition” for ”employeerecognition”,

02 Payment and compensation variations

”bonuses” for ”bonus ”, ”commissions” for ”commission ”, ”incentives” for ”incentive ”, ”profit share” for ”profitsharing”, ”stock option” for ”stockoptions”, ”option stock” for ”stockoptions”, ”share profit” for ”profitsharing”, ”profit sharing” for ”profitsharing”, ”sharing profit” for ”profitsharing”, ”share profits” for ”profitsharing”, ”performance base” for ”performancerelated”, ”performance based” for ”performancerelated”, ”base performance” for ”performancerelated”, ”promotion base” for ”promotionbased”, ”base promotion” for ”promotionbased”, ”promotion based” for ”promotionbased”, ”extra pay” for ”extrapay”, ”pay extra” for ”extrapay”, ”sales goal” for ”salesgoal”, ”goal sales” for ”salesgoal”, ”sales target” for ”salestarget”, ”target sales” for ”salestarget”, ”premium pay” for ”premiumpay”, ”pay premium” for ”premiumpay”, ”salary premium” for ”salarypremium”, ”premium salary” for ”salarypremium”, ”productivity pay” for ”productivitypay”, ”pay productivity” for ”productivitypay”, ”sales quota” for ”salesquota”, ”quota sales” for ”salesquota”, ”variable pay” for ”variablepay”, ”pay for performance” for ”payforperformance”, ”performance pay” for ”payforperformance”, ”decommission” for ”decomm”, ”commissioners” for ”missioners”, ”share saving” for ”sharesaving ”, ”share schemes” for ”sharescheme ”, ”share scheme” for ”sharescheme ”, ”shares scheme” for ”sharescheme ”, ”tips shared” for ”tipsshared ”, ”share tips” for ”tipsshared ”, ”average payout” for ”averagepayout ”, ”annual pay” for ”annualpay”, ”share save” for ”sharesave”, ”anniversary reward” for ”anniversaryreward”, ”save earn scheme” for ”saveearnscheme”, ”savings scheme” for ”savingsscheme”, ”share incentive plan” for ”shareincentiveplan”, ”profit share” for ”profitshare”, ”performance recognised” for ”performancerecognised”

03 Work hours and overtime

”extra hour” for ”extrahour”, ”hour extra” for ”extrahour”, ”long hour” for ”longhour”, ”hour long” for ”longhour”, ”extra shift” for ”extrashift”, ”shift extra” for ”extrashift”, ”double pay” for ”doublepay”, ”pay double” for ”doublepay”, ”unsocial hour” for ”unsocialhour”

04 Time off

”breaks” for ”break”, ”sick day” for ”sickday”, ”day sick” for ”sickday”, ”sick leave” for ”sickleave”, ”leave sick” for ”sickleave”, ”sick time” for ”sicktime”, ”time sick” for ”sicktime”, ”health leave” for ”healthleave”, ”leave health” for ”healthleave”, ”medical leave” for ”medicalleave”, ”leave medical” for ”medicalleave”, ”maternity pay” for ”maternitypay”, ”pay maternity” for ”maternitypay”, ”maternity leave” for ”maternityleave”, ”leave maternity” for ”maternityleave”, ”maternity day” for ”maternityday”, ”day maternity” for ”maternityday”, ”paternity pay” for ”paternitypay”, ”pay paternity” for ”paternitypay”, ”paternity leave” for ”paternityleave”, ”leave paternity” for ”paternityleave”, ”paternity day” for ”paternityday”, ”day

paternity" for "paternityday", "bereavement leave" for "bereavementleave", "leave bereavement" for "bereavementleave", "personal day" for "personalday", "day personal" for "personalday", "paid time off" for "paidtimeoff", "time off paid" for "paidtimeoff", "paid leave" for "paidleave", "leave paid" for "paidleave", "free day" for "freeday", "day free" for "freeday", "training leave" for "trainingleave", "leave training" for "trainingleave", "education leave" for "educationleave", "leave education" for "educationleave", "sickness pay" for "sicknesspay","compensation leave" for "compensationleave","parental leave" for "parentalleave", "study leave" for "studyleave","medical indemnity" for "medicalindemnity","paid rest" for "paidrest", "parental adoption" for "parentaladoption","planned absence" for "plannedabsence","maternity paternity" for "maternitypaternity", "maternity . paternity" for "maternitypaternity","maternity , paternity" for "maternitypaternity","wellbeing benefits" for "wellbeingbenefits", "maternity policies" for "maternitypolicies","sick days" for "sickdays","moving house day" for "movinghouseday", "charity projects" for "charityprojects","sick child" for "sickchild","community volunteering" for "communityvolunteering", "annually volunteer" for "annuallyvolunteer","option buy" for "optionbuy","paid break" for "paidbreak", "birthday" for "birthday","days paid" for "dayspaid","sick pay" for "sickpay", "vacation" for "vacation","maternity package" for "maternitypackage","volunteer time" for "volunteertime", "summer fridays" for "summerfridays","enhanced maternity" for "enhancedmaternity","maternity pay" for "maternitypay", "reset days" for "resetdays","volunteer day" for "volunteerday","adoption leave" for "adoptionleave", "family policies" for "familypolicies","family days" for "familydays","volunteering day" for "volunteeringday", "cash health plan" for "cashhealthplan","healthcare dental care" for "healthcaredentalcare","early friday" for "earlyfriday", "charitable day" for "charitableday","wellbeing days" for "wellbeingdays"

05 Health insurance and benefits

"join family healthcare" for "join family", "family healthcare team" for "family team", "medical benefit" for "medicalbenefit", "benefit medical" for "medicalbenefit", "free health" for "freehealth", "health free" for "freehealth", "free medical" for "freemedical", "medical free" for "freemedical", "health contribution" for "healthcontribution", "contribution health" for "healthcontribution", "health coverage" for "healthcoverage", "coverage health" for "healthcoverage", "offer health insurance" for "offerhealthinsurance", "health insurance offer" for "offerhealthinsurance", "include health insurance" for "includehealthinsurance", "health insurance include" for "includehealthinsurance", "dental benefit" for "dentalbenefit", "benefit dental" for "dentalbenefit", "free dental" for "freedental", "dental free" for "freedental", "dental coverage" for "dentalcoverage", "coverage dental" for "dentalcoverage", "offer dental insurance" for "offerdentalinsurance", "dental insurance offer" for "offerdentalinsurance", "include dental insurance" for "includedentalinsurance", "dental insurance include" for "includedentalinsurance", "vision coverage" for "visioncoverage", "coverage vision" for "visioncoverage", "offer vision insurance" for "offervisioninsurance", "vision insurance offer" for "offervisioninsurance", "include vision insurance" for "includevisioninsurance", "vision insurance include" for "includevisioninsurance", "offer accident insurance" for "offeraccidentinsurance", "accident insurance offer" for "offeraccidentinsurance", "include accident insurance" for "includeaccidentinsurance", "accident insurance include" for "includeaccidentinsurance", "offer life insurance" for "offerlifeinsurance", "life insurance offer" for "offerlifeinsurance", "include life insurance" for "includelifeinsurance", "life insurance include" for "includelifeinsurance", "health savings" for "healthsavings", "savings health" for "healthsavings", "private insurance" for "privateinsurance", "insurance private" for "privateinsurance", "family coverage" for "familycoverage", "coverage family" for "familycoverage", "onsite flu shots" for "onsiteflushots", "child care scheme" for "childcarescheme", "childcare scheme" for "childcarescheme", "health care scheme" for "healthcarescheme", "childcare vouchers" for "childcarevouchers", "child care vouchers" for "childcarevouchers", "private medical" for "privatemedical", "dental vision insurance" for "dentalvisioninsurance", "medical plan" for "medicalplan", "private health" for "privatehealth", "health wellbeing plans" for "healthwellbeingplans", "private health insurance" for "privatehealthinsurance",

”life assurance” for ”lifeassurance”, ”private healthcare scheme” for ”healthcarescheme”, ”health plan” for ”healthplan”, ”cash plan” for ”cashplan”, ”optical dental” for ”opticaldental”, ”life insurance” for ”lifeinsurance”, ”eye care cover” for ”eyecarecover”, ”vision insurance” for ”visioninsurance”, ”flu vouchers” for ”fluvouchers”, ”dental insurance” for ”dentalinsurance”, ”health insurance” for ”healthinsurance”, ”free healthcare” for ”freehealthcare”, ”dental plan” for ”dentalplan”, ”discounted insurance” for ”discountedinsurance”, ”health care plan” for ”healthcareplan”, ”dental vision” for ”dentalvision”, ”dental cover” for ”dentalcover”, ”includes dental” for ”includesdental”, ”eyecare voucher” for ”eyecarevoucher ”, ”medicash” for ”medicash”, ”life cover” for ”lifecover”, ”health wellbeing” for ”healthwellbeing”, ”indemnity cover” for ”indemnitycover”, ”annual flu” for ”annualflu”, ”critical illness cover” for ”criticalillnesscover”, ”critical illness scheme” for ”criticalillnessscheme”, ”dental care” for ”dentalcare”, ”dental treatment” for ”dentaltreatment”, ”disability cover” for ”disabilitycover”, ”eye care discount” for ”eyecarediscount”, ”eye care voucher” for ”eyecarevoucher ”, ”welfare package” for ”welfarepackage”, ”health cover” for ”healthcover”, ”travel accident insurance” for ”travelaccidentinsurance”, ”free health” for ”freehealth”, ”free legal advice” for ”freelegaladvice”, ”life accident insurance” for ”lifeaccidentinsurance”, ”medical cover” for ”medicalcover”, ”mental health first” for ”mentalhealthfirst”, ”personal accident insurance” for ”personalaccidentinsurance”, ”subsidised dental” for ”subsidiseddental”, ”family healthcare” for ”familyhealthcare”, ”free medical” for ”freemedical”, ”illness insurance” for ”illnessinsurance”, ”free flu vaccine” for ”freefluvaccine”, ”free eye test” for ”freeeyetest”, ”wellness programme” for ”wellnessprogramme”, ”medical insurance” for ”medicalinsurance”

06 Retirement and savings

”company match values” for ”company values”, ”company match demand” for ”company demand”, ”team company match ” for ”team company ”, ”retirement contribution” for ”retirementcontribution”, ”contribution retirement” for ”retirementcontribution”, ”retirement fund” for ”retirementfund”, ”fund retirement” for ”retirementfund”, ”savings account” for ”savingsaccount”, ”account savings” for ”savingsaccount”, ”employer contribution” for ”employercontribution”, ”contribution employer” for ”employercontribution”, ”employer match” for ”employermatch”, ”match employer” for ”employermatch”, ”company match” for ”companymatch”, ”match company” for ”companymatch”, ”tuition reimbursement” for ”tuitionreimbursement”, ”reimbursement tuition” for ”tuitionreimbursement”, ”savings contribution” for ”savingscontribution”, ”contribution savings” for ”savingscontribution”, ”savings scheme” for ”savingsscheme”, ”scheme savings” for ”savingsscheme”, ”provident contribution” for ”providentcontribution”, ”contribution provident” for ”providentcontribution”, ”provident fund” for ”providentfund”, ”fund provident” for ”providentfund”, ”twice contribute” for ”twicecontribute”, ”double contribution” for ”doublecontribution”, ”life assurance” for ”lifeassurance”

07 Food, subsidies and discounts

”tea coffee” for ”teacoffee”, ”cinemas” for ”cinema”, ”discounts” for ”discount”, ”memberships” for ”membership”, ”subsidized” for ”subsidised”, ”breakfast included” for ”breakfastincluded”,”cater breakfast” for ”caterbreakfast”,”clothes discount” for ”clothesdiscount”, ”company uniform” for ”companyuniform”,”dinner included” for ”dinnerincluded”,”employee discount” for ”employeediscount”, ”food included” for ”foodincluded”,”free breakfast” for ”freebreakfast”,”free dinner” for ”freedinner”, ”free drinks” for ”freedrinks”,”free food” for ”freefood”,”free lunch” for ”freelunch”,”free meal” for ”freemeal”, ”free transportation” for ”freetransportation”,”gym discount” for ”gymdiscount”,”live position” for ”liveposition”, ”lunch included” for ”lunchincluded”,”meal included” for ”mealincluded”,”transportation allowance” for ”transportationallowance”, ”travel subsidy” for ”travelsubsidy”,”cater lunch” for ”caterlunch”,”cater meal” for ”catermeal”, ”commute subsidy” for ”commutesubsidy”,”discount coupon” for ”discountcoupon”,”food

coupon" for "foodcoupon", "food stamp" for "foodstamp","food subsidy" for "foodsubsidy","food ticket" for "foodticket","housing included" for "housingincluded", "lunch ticket" for "lunchticket","on site accommodation" for "onsiteaccommodation","rental subsidy" for "rentalsubsidy", "transit subsidy" for "transitsubsidy","transportation subsidy" for "transportationsubsidy","transport voucher" for "transportvoucher", "travel stipend" for "travelstipend","voluntary benefits" for "voluntarybenefits","cash back" for "cashback", "wellbeing offers" for "wellbeingoffers","lease scheme" for "leasescheme","advantage scheme" for "advantagescheme", "afternoon snacks" for "afternoonsnacks","benefits platform" for "benefitsplatform","benefits portal" for "benefitsportal", "bike insurance" for "bikeinsurance","car discount" for "cardiscount","car leasing" for "carleasing", "car loan scheme" for "carloanscheme","cinema tickets" for "cinematickets","comfortable uniform" for "comfortableuniform", "company discount" for "companydiscount","corporate discount" for "corporatediscount","daily fruit" for "dailyfruit", "daily menu" for "dailymenu","deals discount" for "dealsdiscount","death service" for "deathservice","discount car" for "discountcar","discount card" for "discountcard", "discount programme" for "discountprogramme","discount schemes" for "discountschemes","discount travel" for "discounttravel", "discount various" for "discountvarious","discounted cinema" for "discountedcinema","discounted fitness" for "discountedfitness", "discounted food" for "discountedfood","discounted membership" for "discountedmembership","discounted stays" for "discountedstays", "discounted subscription" for "discountedsubscription","drinks fridge" for "drinksfridge","food provided" for "foodprovided","unlimited breakfast" for "unlimitedbreakfast", "free coffee" for "freecoffee","free daily breakfast" for "freedailybreakfast","free fruit" for "freefruit", "free healthy" for "freehealthy","free hot drinks" for "freehotdrinks","free ice cream" for "freeicecream","free membership" for "freemembership", "free refreshments" for "freerefreshments","free soft" for "freesoft","free tea coffee" for "freeteacoffee", "free unhealthy" for "freeunhealthy","free yoga" for "freeyoga","fresh fruit" for "freshfruit","fruit every day" for "fruiteveryday", "fruit refreshments" for "fruitrefreshments","fruit snacks" for "fruitsnacks","fuel card" for "fuelcard","gym discount" for "gymdiscount", "gym membership" for "gymmembership","gym pass" for "gympass","gym subsidies" for "gymsubsidies","healthy snacks" for "healthysnacks", "hotel discount" for "hoteldiscount","hub offers" for "huboffers","leisure discount" for "leisurediscount","leisure savings" for "leisuresavings", "loyalty rewards" for "loyaltyrewards","lunch allowance" for "lunchallowance","meal provided" for "mealprovided", "meals duty" for "mealsduty","meals provided" for "mealsprovided","office snacks" for "officesnacks","online benefits" for "onlinebenefits","online discount" for "onlinediscount", "onshift meals" for "onshiftmeals","refreshments provided" for "refreshmentsprovided","retail discount" for "retaildiscount", "reward platform" for "rewardplatform","rewards points" for "rewardspoints","season tickets" for "seasontickets","shopping vouchers" for "shoppingvouchers","staff discount" for "staffdiscount", "store discount" for "storediscount","subsidised meal" for "subsidisedmeal","subsidised rate" for "subsidisedrate", "swim membership" for "swimmembership","tea coffee" for "teacoffee","travel benefits" for "travelbenefits", "travel discount" for "traveldiscount","travel insurance" for "travelinsurance","wellness portal" for "wellnessportal", "access headspace" for "accessheadspace","beer fridays" for "beerfridays","coffee tea" for "coffeetea", "drinks available" for "drinksavailable","drinks facilities" for "drinksfacilities","fitness classes" for "fitnessclasses", "free beer" for "freebeer","free beverages" for "freebeverages","free tea" for "freetea","fridays drinks" for "fridaysdrinks", "gym nearby" for "gymnearby","headspace subscription" for "headspacesubscription","instore discounts" for "instorediscounts", "meditation app" for "meditationapp","personal parcels" for "personalparcels","pool tables" for "pooltables", "excellent coffee machine" for "excellentcoffeemachine","nespresso coffee machine" for "nespressocoffeemachine", "free coffee machine" for "freecoffeemachine","shopping portal" for "shoppingportal","workday lunches" for "workdaylunches", "yoga sessions" for "yogasessions","discount local" for "discountlocal","free cake" for "freecake", "free tickets" for "freetickets","free use" for "freeuse","regular drinks" for "regulardrinks","weekly yoga" for "weeklyyoga"

08 Office space and amenities

”areas” for ”area”,”comfortable office environment” for ”good office environment”, ”fitness room” for ”fitnessroom”,”game room” for ”gameroom”,”office gym” for ”officegym”,”office yoga” for ”officeyoga”, ”comfortable office” for ”comfortableoffice”,”comfortable space” for ”comfortablespace”,”open space” for ”openspace”, ”natural light” for ”naturallight”,”company cafeteria” for ”companycafeteria”,”conference room facilities” for ”conferenceroomfacilities”, ”amenities” for ”amenities”,”amazing workplace” for ”amazingworkplace”,”brand new office” for ”brandnewoffice”, ”break area” for ”breakarea”,”break rooms” for ”breakrooms”,”breakout area” for ”breakoutarea”,”canteen” for ”canteen”, ”chillout area” for ”chilloutarea”,”comfortable facilities” for ”comfortablefacilities”,”coworking space” for ”coworkingspace”, ”developed office” for ”developedoffice”,”facilities onsite” for ”facilitiesonsite”,”fantastic facilities” for ”fantasticfacilities”, ”fitness studio” for ”fitnessstudio”,”games area” for ”gamesarea”,”games room” for ”gamesroom”,”good facilities” for ”goodfacilities”, ”great facilities” for ”greatfacilities”,”great facilities” for ”greatfacilities”,”great views” for ”greatviews”, ”leisure facilities” for ”leisurefacilities”,”modern facilities” for ”modernfacilities”,”modern office” for ”modernoffice”, ”office includes” for ”officeincludes”,”office modern” for ”officemodern”,”onsite cafeteria” for ”onsitecafeteria”, ”onsite catering” for ”onsitecatering”,”onsite cinema” for ”onsitecinema”,”onsite facilities” for ”onsitefacilities”, ”onsite fitness” for ”onsitefitness”,”onsite gym” for ”onsitegym”,”onsite kitchen” for ”onsitekitchen”, ”onsite nursery” for ”onsitenursery”,”onsite recreation” for ”onsiterecreation”,”onsite restaurant” for ”onsiterestaurant”, ”onsite restroom” for ”onsiterestroom”,”onsite shower” for ”onsiteshower”,”onsite subsidised” for ”onsitesubsidised”, ”onsite swimming” for ”onsiteswimming”,”open plan environment” for ”openplanenvironment”,”open plan office” for ”openplanoffice”, ”prayer room” for ”prayerroom”,”quiet work spaces” for ”quietworkspaces”,”range facilities” for ”rangefacilities”, ”refurbished office” for ”refurbishedoffice”,”relaxation area” for ”relaxationarea”,”roof terrace” for ”roofterrace”, ”rooftop terrace” for ”rooftopterrace”,”self service facilities” for ”selfservicefacilities”,”shower facilities” for ”showerfacilities”, ”shower facility” for ”showerfacility”,”sport facilities” for ”sportfacilities”,”staff rest area” for ”staffrestarea”, ”staff restaurant” for ”staffrestaurant”,”stocked kitchen” for ”stockedkitchen”,”subsidised restaurant” for ”subsidisedrestaurant”, ”table tennis” for ”tabletennis”,”tennis room” for ”tennisroom”,”use microwave” for ”usemicrowave”, ”vaping area” for ”vapingarea”,”wellness room” for ”wellnessroom”,”xbox room” for ”xboxroom”,”staff lockers” for ”stafflockers”, ”excellent facilities” for ”excellentfacilities”,”great offices” for ”greatoffices”,”location also provides” for ”locationalsoprovides”, ”location provides” for ”locationprovides”,”onsite bistro” for ”onsitebistro”,”open roof” for ”openroof”,”quiet working spaces” for ”quietworkingspaces”, ”state art facilities” for ”stateartfacilities”

09 Location and commuting

”spaces” for ”space”, ”city center” for ”citycenter”,”central location” for ”centrallocation”,”excellent location” for ”excellentlocation”, ”good location” for ”goodlocation”,”easy access” for ”easyaccess”,”private parking” for ”privateparking”,”pay parking” for ”payparking”, ”company garage” for ”companygarage”,”private garage” for ”privategarage”,”short commute” for ”shortcommute”,”long commute” for ”longcommute”, ”local shops” for ”localshops”,”access public transport” for ”accesspublictransport”,”based centre” for ”basedcentre”, ”based heart” for ”basedheart”,”bike 2 work” for ”bike2work”,”bike sheds” for ”bikesheds”,”bike storage” for ”bikestorage”, ”bike work scheme” for ”bikeworkscheme”,”bike2 work” for ”bike2work”,”bus links” for ”buslinks”,”bus train” for ”bustrain”, ”car park” for ”carpark”,”car parking” for ”carparking”,”central location” for ”centrallocation”,”centre located” for ”centrelocated”, ”centre location” for ”centrelocation”,”childcare options” for ”childcareoptions”,”commuter links” for ”commuterlinks”, ”conveniently located” for ”convenientlylocated”,”coworking space” for ”coworkingspace”,”cycle scheme” for ”cyclescheme”,”cycle to work” for ”cycletowork”, ”cycle work” for ”cyclework”,”cycling scheme” for ”cyclingscheme”,”discounted bus” for ”discountedbus”,”easily accessible” for ”easilyaccessible”, ”easily

commutable” for ”easilycommutable”,”easy commute” for ”easycommute”,”easy reach” for ”easyreach”, ”excellent location” for ”excellentlocation”,”fine location” for ”finelocation”,”free bus” for ”freebus”, ”free carparking” for ”freecarparking”,”free local bus” for ”freelocalbus”,”free parking” for ”freeparking”, ”good location” for ”goodlocation”,”great location” for ”greatlocation”,”great transport” for ”greattransport”, ”leasing deals” for ”leasingdeals”,”links transport” for ”linkstransport”,”local bus” for ”localbus”, ”local transport” for ”localtransport”,”min walk” for ”minwalk”,”minute walk” for ”minutewalk”, ”minutes central” for ”minutescentral”,”near local” for ”nearlocal”,”next central station” for ”nextcentralstation”, ”office heart” for ”officeheart”,”office next” for ”officenext”,”on site park” for ”onsitepark”,”onsite park” for ”onsitepark”, ”outside city centre” for ”outsidecitycentre”,”public transport” for ”publictransport”,”reliable public” for ”reliablepublic”, ”ride2 work” for ”ride2work”,”secure parking” for ”secureparking”,”short walk” for ”shortwalk”, ”site parking” for ”siteparking”,”situated close” for ”situatedclose”,”situated heart” for ”situatedheart”, ”staff parking” for ”staffparking”,”street parking” for ”streetparking”,”tax free bike” for ”taxfreebike”,”town center” for ”towncenter”, ”transport links” for ”transportlinks”,”transportation links” for ”transportationlinks”,”via public transport” for ”viapublictransport”, ”walking distance” for ”walkingdistance”,”well located” for ”welllocated”,”within short distance” for ”withinshortdistance”, ”easiest reach” for ”easiestreach”,”easily commute” for ”easilycommute”,”ideal location” for ”ideallocation”, ”local parking” for ”localparking”,”mainline station” for ”mainlinestation”,”metro station close” for ”metrostationclose”, ”minutes walk” for ”minuteswalk”,”parking available” for ”parkingavailable”,”parking space” for ”parkingspace”, ”peaceful location” for ”peacefullocation”,”perfect commuting” for ”perfectcommuting”,”perfect location” for ”perfectlocation”, ”quick access” for ”quickaccess”,”walkable distance” for ”walkabledistance”,”state art building” for ”stateartbuilding”,”metro station” for ”metrostation”

10 Work equipment and technology

”maintain company” for ”maintain comp ”,”maintain equipment” for ”maintain equ ”,”uniforms” for ”uniform”, ”internet coupon” for ”internetcoupon”, ”coupon internet” for ”internetcoupon”, ”internet subsidy” for ”internetsubsidy”, ”subsidy internet” for ”internetsubsidy”, ”technology subsidy” for ”technologysubsidy”, ”subsidy technology” for ”technologysubsidy”, ”home office fund” for ”homeofficefund”, ”fund home office” for ”homeofficefund”, ”telework fund” for ”teleworkfund”, ”fund telework” for ”teleworkfund”, ”telework stipend” for ”teleworkstipend”, ”stipend telework” for ”teleworkstipend”, ”work home fund” for ”workhomefund”, ”fund work home” for ”workhomefund”, ”telework expenses” for ”teleworkexpenses”, ”expenses telework” for ”teleworkexpenses”, ”home office expenses” for ”homeofficeexpenses”, ”expenses home office” for ”homeofficeexpenses”, ”work home expenses” for ”workhomeexpenses”, ”expenses work home” for ”workhomeexpenses”, ”home office allowance” for ”homeofficeallowance”, ”allowance home office” for ”homeofficeallowance”, ”telework allowance” for ”teleworkallowance”, ”allowance telework” for ”teleworkallowance”, ”work home allowance” for ”workhomeallowance”, ”allowance work home” for ”workhomeallowance”, ”telework subsidy” for ”teleworksubsidy”, ”subsidy telework” for ”teleworksubsidy”, ”work home subsidy” for ”workhomesubsidy”, ”subsidy work home” for ”workhomesubsidy”, ”home office equipment” for ”homeofficeequipment”, ”equipment home office” for ”homeofficeequipment”, ”telework equipment” for ”teleworkequipment”, ”equipment telework” for ”teleworkequipment”, ”offer mobile phone” for ”offermobilephone”, ”mobile phone offer” for ”offermobilephone”, ”include mobile phone” for ”includemobilephone”, ”mobile phone include” for ”includemobilephone”, ”offer company cellphone” for ”offercompanycellphone”, ”company cellphone offer” for ”offercompanycellphone”, ”include company cellphone” for ”includecompanycellphone”, ”company cellphone include” for ”includecompanycellphone”, ”offer company computer” for ”offercompanycomputer”, ”company computer offer” for ”offercompanycomputer”, ”include company computer” for ”includecompanycomputer”, ”company computer include” for ”includecompanycomputer”, ”offer company laptop” for ”offercompanylaptop”, ”company laptop offer” for ”offercompanylaptop”, ”include company laptop” for ”includecompanylaptop”, ”company laptop include” for ”includecompa-

nylaptop", "close access" for "closeaccess","company car" for "companycar","company credit card" for "companycreditcard", "company debit card" for "companydebitcard","company laptop" for "companylaptop","company phone" for "companyphone", "company uniform" for "companyuniform","company van" for "companyvan","company vehicle" for "companyvehicle", "corporate clothing" for "corporateclothing","dedicated desk" for "dedicateddesk","discounts energy bills" for "discountsenergybills", "equipment provided" for "equipmentprovided","free broadband" for "freebroadband","free mobile" for "freemobile", "free uniform" for "freeuniform","free uniform" for "freeuniform","fuel card" for "fuelcard", "necessary equipment" for "necessaryequipment","paid mileage" for "paidmileage","perkbox" for "perkbox", "perks" for "perks","phone and laptop" for "phoneandlaptop","phone insurance" for "phoneinsurance", "phone laptop" for "phonelaptop","phone provided" for "phoneprovided","protective footwear" for "protectivefootwear", "provision uniform" for "provisionuniform","two monitors" for "twomonitors","two screens" for "twoscreens", "uniform provided" for "uniformprovided","uniform provision" for "uniformprovision","vehicle allowance" for "vehicleallowance", "vehicle scheme" for "vehiclescheme","welcome box" for "welcomebox","protective equipment" for "protectiveequipment", "phone work" for "phonework","car allowance" for "carallowance","discounted broadband" for "discountedbroadband", "fuel allowance" for "fuelallowance","mileage expense" for "mileageexpense","perbox" for "perbox"

11 Work flexibility

"flexible shift" for "flexibleshift", "shift flexible" for "flexibleshift", "flexible time" for "flexibletime", "time flexible" for "flexibletime", "flexible work" for "flexiblework", "work flexible" for "flexiblework", "flexible hour" for "flexiblehour", "hour flexible" for "flexiblehour", "flexible schedule" for "flexibleschedule", "schedule flexible" for "flexibleschedule", "work life balance" for "worklifebalance", "life work balance" for "worklifebalance", "family friendly" for "familyfriendly", "friendly family" for "familyfriendly", "weekly rest" for "weeklyrest", "rest weekly" for "weeklyrest", "flexibility around working" for "flexibilityaroundworking", "flexibility work" for "flexibilitywork", "flexible office hours" for "flexibleofficehours", "flexible working" for "flexibleworking", "flx hour" for "flxhour", "flx schedule" for "flxschedule","flx shift" for "flxshift", "flx time" for "flxtime","flx work" for "flxwork","life balance" for "lifebalance","life work" for "lifework", "flx schedule" for "flxschedule","flx work" for "flxwork","healthy work" for "healthywork", "modern working" for "modernworking","family friendly" for "familyfriendly","work . life" for "work.life", "worklife balance" for "worklifebalance","work life balance" for "worklifebalance","flx option" for "flxoption", "flexible option" for "flexibleoption", "flx shift" for "flexiblework", "flx with" for "flexiblework", "flx basis" for "flexiblework", "flx about" for "flexiblework", "hours are flx" for "flexiblework", "flx annual" for "flexiblework", "is flx" for "flexiblework", "flx schedule" for "flexiblework", "fully flx" for "flexiblework", "flx pay" for "flexiblework", "flx opportunities" for "flexiblework", "flx subcontract" for "flexiblework", "flx between" for "flexiblework", "flx start" for "flexiblework", "flx dayst" for "flexiblework", "flx within" for "flexiblework", "choose flx" for "flexiblework", "times flx" for "flexiblework", "flx schedules" for "flexiblework", "flx vacancies" for "flexiblework", "scheme flx" for "flexiblework", "time flx" for "flexiblework", "flx and optional" for "flexiblework", "flx self employed" for "flexiblework", "flx in relation to" for "flexiblework", "flx job" for "flexiblework", "options flx" for "flexiblework", "flx rotas" for "flexiblework", "flx employment" for "flexiblework", "days are flx" for "flexiblework", "freelance and flx" for "flexiblework"

12 Workplace safety and ergonomics

"safe place" for "safeplace", "place safe" for "safeplace", "safety culture" for "safetyculture", "culture safety" for "safetyculture","employee safety" for "employeesafety", "safety employee" for "employeesafety", "work safety" for "worksafety", "safety work" for "worksafety","safety first" for "safetyfirst", "first safety" for "safetyfirst", "secure offices" for "secureoffices", "safe work" for "safework","health safety" for

”healthsafety”, ”safety rules” for ”safetyrules”, ”secure workplace” for ”secureworkplace”, ”focus health” for ”focushealth”,”safe working” for ”safeworking”, ”workplace safety” for ”workplacesafety”, ”aware health” for ”awarehealth”, ”comply safe” for ”complysafe”, ”safety legislation” for ”safetylegislation”, ”act 1974” for ”act1974”,”fire regulation” for ”fireregulation”, ”safety act” for ”safetyact”,”comply health” for ”complyhealth”, ”compliance health” for ”compliancehealth”, ”understand health” for ”understandhealth”,”requirements health” for ”requirementshealth”, ”safe working” for ”safeworking”,”safety policies” for ”safetypolicies”,”safe office” for ”safeoffice”,”hygienic working” for ”hygienicworking”

13 Job security

”severance pay” for ”severancepay”, ”pay severance” for ”severancepay”, ”severance package” for ”severancepackage”, ”package severance” for ”severancepackage”, ”layoff compensation” for ”layoffcompensation”, ”compensation layoff” for ”layoffcompensation”, ”unemployment insurance” for ”unemploymentinsurance”, ”insurance unemployment” for ”unemploymentinsurance”, ”unemployment benefit” for ”unemploymentbenefit”, ”benefit unemployment” for ”unemploymentbenefit”, ”low turnover” for ”lowturnover”, ”turnover low” for ”lowturnover”, ”job security” for ”jobsecurity”, ”security job” for ”jobsecurity”, ”job stability” for ”jobstability”, ”stability job” for ”jobstability”, ”low turnover” for ”lowturnover”,”job security” for ”jobsecurity”,”job stability” for ”jobstability”, ”indemnity insurance” for ”indemnityinsurance”,”income protection” for ”incomeprotection”,”stable environment” for ”stableenvironment”, ”severance pay” for ”severancepay”,”severance package” for ”severancepackage”,”layoff compensation” for ”layoffcompensation”, ”unemployment insurance” for ”unemploymentinsurance”,”unemployment benefit” for ”unemploymentbenefit”

14 Work environment and culture

”working conditions” for ”work conditions”,”working hour” for ”work hour”,”benefit culture” for ”benefitculture”, ”best workplace” for ”bestworkplace”,”comfort work” for ”comfortwork”,”congenial work” for ”congenialwork”, ”dynamic workplace” for ”dynamicworkplace”,”enjoyable work” for ”enjoyablework”,”excellent environment” for ”excellentenvironment”, ”good atmosphere” for ”goodatmosphere”,”good culture” for ”goodculture”,”good environment” for ”goodenvironment”, ”great place work” for ”greatplacework”,”healthy work” for ”healthywork”,”human values” for ”humanvalues”, ”job recognition” for ”jobrecognition”,”meaningful work” for ”meaningfulwork”,”mission vision” for ”missionvision”, ”positive work” for ”positivework”,”promote equality” for ”promoteequality”,”relaxed environment” for ”relaxedenvironment”, ”relax work” for ”relaxwork”,”strong company” for ”strongcompany”,”warm environment” for ”warmenvironment”, ”young environment” for ”youngenvironment”,”multicultural environment” for ”multiculturalenvironment”,”afterwork drinks” for ”afterworkdrinks”, ”best places work” for ”bestplaceswork”,”best reputation” for ”bestreputation”,”casual dress” for ”casualdress”, ”collaborative culture” for ”collaborativeculture”,”collaborative environment” for ”collaborativeenvironment”, ”company culture” for ”companyculture”,”company events” for ”companyevents”,”company outings” for ”companyoutings”, ”company parties” for ”companyparties”,”company retreats” for ”companyretreats”,”company social events” for ”companysocialevents”, ”company trips” for ”companytrips”,”company values” for ”companyvalues”,”culture matter” for ”culturematter”,”difference life” for ”differencelife”, ”dress code” for ”dresscode”,”employee assistance” for ”employeeassistance”,”employee support” for ”employeesupport”, ”enjoyable working” for ”enjoyableworking”,”excellent workplace” for ”excellentworkplace”,”fantastic working” for ”fantasticworking”, ”free counselling” for ”freecounselling”,”friendly working” for ”friendlyworking”,”funny social events” for ”funnysocialevents”, ”games nights” for ”gamesnights”,”good vibes” for ”goodvibes”,”great working” for ”greatworking”,”great workplace” for ”greatworkplace”, ”huge impact” for ”hugeimpact”,”immediate impact” for ”immediateimpact”,”impact society”

for ”impactsociety”, ”inclusive working” for ”inclusiveworking”,”inclusive working” for ”inclusiveworking”,”internal culture” for ”internalculture”, ”job satisfaction” for ”jobsatisfaction”,”lively office” for ”livelyoffice”,”local reputation” for ”localreputation”, ”lovely atmosphere” for ”lovelyatmosphere”,”make difference” for ”makedifference”,”make real difference” for ”makerealdifference”, ”making impact” for ”makingimpact”,”meaningful project” for ”meaningfulproject”,”multi cultural” for ”multicultural”, ”outdoor activities” for ”outdooractivities”,”people lives” for ”peoplelives”,”pleasant office” for ”pleasantoffice”, ”positive culture” for ”positiveculture”,”positive impact” for ”positiveimpact”,”prioritising wellbeing” for ”prioritisingwellbeing”, ”regular social events” for ”regularsocialevents”,”regular socials” for ”regularsocials”,”relaxed atmosphere” for ”relaxedatmosphere”, ”responsible company” for ”responsiblecompany”,”rewarding” for ”rewarding”,”safe environment” for ”safeenvironment”, ”satisfaction job” for ”satisfactionjob”,”sociable office” for ”sociableoffice”,”social activities” for ”socialactivities”, ”social environment” for ”socialenvironment”,”social gatherings” for ”socialgatherings”,”staff events” for ”staffevents”, ”stimulating environment” for ”stimulatingenvironment”,”supportive environment” for ”supportiveenvironment”, ”supportive working” for ”supportiveworking”,”team bonding” for ”teambonding”,”team building” for ”teambuilding”, ”unique environment” for ”uniqueenvironment”,”value diversity” for ”valuediversity”,”values” for ”values”, ”vibrant office” for ”vibrantoffice”,”vibrant workplace” for ”vibrantworkplace”,”wellbeing support” for ”wellbeingsupport”, ”wellbeing workshops” for ”wellbeingworkshops”,”within community” for ”withincommunity”,”work environment” for ”workenvironment”, ”working culture” for ”workingculture”,”working environment” for ”workingenvironment”,”direct impact” for ”directimpact”, ”equal opportunities” for ”equalopportunities”,”fulfilling” for ”fulfilling”,”impact world” for ”impactworld”, ”make impact” for ”makeimpact”,”working culture” for ”workingculture”,”working ethos” for ”workingethos”, ”office culture” for ”officeculture”,”real impact” for ”realimpact”,”significant impact” for ”significantimpact”, ”social work environment” for ”socialworkenvironment”,”strong impact” for ”strongimpact”,”positive change” for ”positivechange”, ”valuemember” for ”valuemember”, ”feelpart” for ”feelpart”,”interdisciplinary” for ”interdisciplinary”,”multidisciplinary” for ”multidisciplinary”, ”dynamicteam” for ”dynamicteam”,”youngteam” for ”youngteam”,”great team” for ”greatteam”,”team lunch” for ”teamlunch”, ”team events” for ”teamevents”,”exciting team” for ”excitingteam”

15 Learning and development

”learning disabilities” for ”learnd disabilities”, ”base merit” for ”basemerit”, ”merit base” for ”basemerit”,”build career” for ”buildcareer”, ”career build” for ”buildcareer”, ”company growth” for ”companygrowth”,”growth company” for ”companygrowth”,”education subsidy” for ”educationsubsidy”,”growth opportunity” for ”growthopportunity”, ”growth possible” for ”growthpossible”,”merit increase” for ”meritincrease”,”merit promotion” for ”meritpromotion”, ”pay education” for ”payeducation”,”personal growth” for ”personalgrowth”,”professional development” for ”professionaldevelopment”, ”professional growth” for ”professionalgrowth”,”promotion process” for ”promotionprocess”,”salary increase” for ”salaryincrease”, ”build skills” for ”buildskills”,”career” for ”career”,”development opportunities” for ”developmentopportunities”, ”development opportunity” for ”developmentopportunity”,”elearning portal” for ”elearningportal”, ”external learning” for ”externallearning”,”growth opportunities” for ”growthopportunities”, ”learning allowance” for ”learningallowance”,”learning development” for ”learningdevelopment”, ”move career forward” for ”movecareerforward”,”personal development” for ”personaldevelopment”, ”professional development” for ”professionaldevelopment”,”promote internally” for ”promoteinternally”, ”promotion opportunities” for ”promotionopportunities”,”promotion prospects” for ”promotionprospects”, ”training” for ”training”,”developing skills” for ”developingskills”,”development support” for ”developmentsupport”, ”ongoing growth” for ”ongoinggrowth”,”continuous opportunities” for ”continuousopportunities”, ”promotion prospects” for ”promotionprospects”,”growing organisation” for ”growingorganisation”, ”growing company” for ”growingcompany”,”opportunities learning” for ”opportunitieslearning”,

”learning development” for ”learningdevelopment”,”learning zone” for ”learningzone”, ”growth company” for ”growthcompany”,”continuous improvement” for ”continuousimprovement”, ”growing business” for ”growingbusiness”,”continuing professional” for ”continuingprofessional”, ”merit based” for ”meritbased”,”opportunities growing” for ”opportunitiesgrowing”, ”opportunities upskill” for ”opportunitiesupskill”,”development opportunities” for ”developmentopportunities”, ”promotion opportunities” for ”promotionopportunities”,”promotion development” for ”promotiondevelopment”, ”teamexpansion” for ”teamexpansion”,”growingteam” for ”growingteam

16 Remote Work availability

”conducted remote” for ”conductedremote”, ”fully remote” for ”fullyremote”, ”home based” for ”homebased”, ”home work” for ”homework”, ”job is virtual” for ”jobisvirtual”, ”location is remote” for ”locationisremote”, ”location remote” for ”locationremote”, ”position is remote” for ”positionisremote”, ”position remote” for ”positionremote”, ”remote work” for ”remotework”, ”role will be remote” for ”rolewillberemote”, ”virtual job” for ”virtualjob”, ”work from home” for ”workfromhome”, ”work home” for ”workhome”, ”work remote” for ”workremote”, ”working remote” for ”workingremote”, ”flx with office” for ”flxwithoffice”, ”remote position” for ”remoteposition”, ”role is remote” for ”roleisremote”, ”remote basis” for ”remotebasis”, ”hybrid work” for ”hybridwork”, ”hybrid remote” for ”hybridremote”, ”tele commuting” for ”telecommuting”, ”tele work” for ”telework ”, ”no telecommuting” for ”no”, ”does not offer teleworker” for ”no”, ”no taking work home” for ”no”, ”take work home” for ”no”

Appendix C - Term Frequency - Inverse Document Frequency (TF-IDF) and Cosine Similarity

TF-IDF

The Term Frequency - Inverse Document Frequency (TF-IDF) is a statistical measure used to evaluate how important a word is. It assigns a weight to each term (word) in a document (vacancy text) in a corpus (set of all vacancies). It comprises two components: Term Frequency (TF) and Inverse Document Frequency (IDF).

Term frequency $tf(t, d)$ is the simplest approach where each term t is assigned a weight equal to the number of times it appears in the document d. Formally, it is defined as:

$$tf(t, d) = \text{frequency of term } t \text{ in document } d \tag{15}$$

Inverse Document Frequency $idf(t)$, on the other hand, is a measure of the amount of information the term provides, i.e., if it's common or rare across all documents. It is the logarithmically scaled inverse fraction of the documents that contain the term. Formally, it is defined as:

$$idf(t) = \log\left(\frac{N}{df(t)}\right) \tag{16}$$

where N is the total number of documents in the corpus and $df(t)$ is the number of documents in the corpus that contain the term t.

The TF-IDF weight of a term is the product of its TF and IDF:

$$tfidf(t, d) = tf(t, d) \times idf(t) \tag{17}$$

Cosine Similarity

In the context of Information Retrieval, each document is represented as a vector, and the similarity between the two documents is quantified using the cosine of the angle between their vectors. Text vectorisation involves transforming textual data into a numerical format so that it can be analyzed mathematically. One common approach is to use a bag-of-words model. In this model, I create a vocabulary of unique terms (words) present in the corpus and represent each document as a vector in this high-dimensional space. Each dimension corresponds to a unique term from the vocabulary.

For instance, consider a corpus with two documents: Document 1: "Cat sits on the mat." and Document 2: "Dog sits on the log.". The vocabulary for this corpus would be cat, dog, sits, on, the, mat, log. Each document is then represented as a 7-dimensional

vector based on the frequency of these words. For example, the vector for Document 1 would look like [1, 0, 1, 1, 1, 1, 0], representing the presence of 'cat', 'sits', 'on', 'the', and 'mat', and the absence of 'dog' and 'log'. Once the documents are represented as vectors, the cosine similarity can be calculated. The cosine similarity measures the cosine of the angle between these two vectors. Mathematically, it is the dot product of the vectors divided by the product of their magnitudes (Euclidean lengths). This measure ranges from -1 to 1, where 1 indicates identical vectors, 0 indicates orthogonality (no similarity), and -1 indicates completely opposite vectors.

Given two documents d_1 and d_2, the cosine similarity is defined as follows:

$$sim(d_1, d_2) = \frac{V(d_1) \cdot V(d_2)}{||V(d_1)|| \times ||V(d_2)||} \tag{18}$$

where $V(d_1)$ and $V(d_2)$ are the vector representations of documents d_1 and d_2 respectively, $V(d_1) \cdot V(d_2)$ is the dot product of the vectors, and $||V(d_1)||$ and $||V(d_2)||$ are their Euclidean lengths.

Appendix D - BERT Model for Gender Classification

This appendix provides technical details on the BERT model used to identify implicit gender targeting in job advertisements.

Model Architecture and Implementation

I implemented a fine-tuned BERT model using the pre-trained bert-base-uncased model from the Hugging Face Transformers library. The model architecture consists of the BERT encoder with 12 transformer layers and a classification head for the three categories: female-targeted, male-targeted, and neutral.

During implementation, I added special tokens `[GENDER]` and `[CONTEXT]` to the tokenizer vocabulary to handle gender-specific terms and contextual phrases during preprocessing.

Data Preprocessing and Feature Engineering

The data preprocessing pipeline involved several steps to prepare the text data for model training:

Identifying Explicitly Gendered Advertisements

To create the training dataset, I identified advertisements with explicit gender preferences using two sets of criteria:

Gender Keywords: I used the following sets of keywords to identify potential gender preferences:

- Female indicators: “female”, “woman”, “women”, “girl”, “girls”, “lady”, “ladies”, “feminine”
- Male indicators: “male”, “males”, “man”, “men”, “guy”, “guys”, “boy”, “boys”, “gent”, “gents”

Contextual Filtering: To distinguish between references to worker gender versus client gender, I applied contextual rules based on two comprehensive lists:

- Contextual clues that indicate client references (for exclusion):

  > “service for”, “patients”, “assist a”, “men client”, “support to”, “supports women”, “supporting”, “care for”, “care to”, “caring for”, “client is a”, “therapy for”, “treatment for”, “work for”, “work with”, “services

for", "support for", "assistance to", "assistance for", "programme for", "counselling for", "advice for", "guidance for", "empower", "service is for", "dedicated to", "abuse", "in need", "working with", "pregnant", "isle of man"

- Gender-specific phrases that indicate worker requirements (for inclusion):

  "applicants only", "applicant only", "female only", "men only", "male only", "male support", "recruit", "recruiting for", "be female", "be male", "available for a female", "encourage"

If a gender keyword appeared within the same sentence as a contextual clue (within 10 words before or after), it was assumed to be referring to a client or service user rather than a job applicant, unless a gender-specific phrase was also present indicating an explicit worker requirement.

Text Cleaning and Normalisation

For model training, I applied the following text processing steps:

1. Basic Text Cleaning:
   - Convert to lowercase
   - Remove punctuation and special characters
   - Remove numbers and digit-only tokens
2. Linguistic Processing:
   - Remove common stopwords (e.g., "the", "and", "a")
   - Apply stemming to reduce words to their root forms
   - Remove very short words (fewer than 4 characters)
3. Gender Word Handling:
   - Replace explicit gender keywords with the `[GENDER]` token
   - Replace contextual phrases with the `[CONTEXT]` token

This preprocessing ensures the model learns underlying linguistic patterns rather than relying on explicit gender terms. For example, an original text like:

"We are looking for female care assistants to provide care for elderly women clients"

Would be transformed to:

"looking [GENDER] care assist provide [CONTEXT] elderly [GENDER] client"

This preprocessing removes stopwords, applies stemming (e.g., "assistants" becomes "assist"), and replaces gender terms and contextual phrases with special tokens.

Training Data Stratification

To prevent occupation-specific biases, I stratified the training data by occupation codes. For each occupation-gender combination, I sampled a minimum of 10 and a maximum of 5,000 advertisements. This balanced approach ensures that occupations with high concentrations of explicit gender preferences (such as care roles) do not dominate the training data.

Model Training

The model was trained using the following procedure:

1. Split the explicitly gendered advertisements into training (70 per cent), validation (15 per cent), and test (15 per cent) sets with stratification
2. Apply class weighting to handle imbalanced classes
3. Train for 3 epochs with early stopping based on validation loss

The training process took a total of 87,614 seconds (approximately 24.3 hours) across three epochs. As shown in Table D.1, the model demonstrated significant improvement during training, with training loss decreasing from 0.23 in the first epoch to 0.08 by the third epoch. Validation accuracy improved from 90.3 to 94.7 per cent, while validation loss decreased from 0.26 to 0.16.

Table D.1: Model Training Metrics by Epoch

Epoch	Training Loss	Validation Loss	Validation Accuracy	Training Time (s)
1	0.230	0.260	0.903	8,440
2	0.164	0.180	0.939	48,905
3	0.084	0.165	0.947	30,269

Model Performance and Validation

The final model achieved an overall accuracy of 94.7 per cent on the test set, with comparable performance across all three classification categories. As shown in Table D.2, the model demonstrated high precision and recall for all target classes, with the neutral class showing the highest F1-score (0.950).

Table D.2: Final Classification Metrics

Class	Precision	Recall	F1-score	Support
Female	0.954	0.938	0.946	13,307
Male	0.934	0.953	0.943	11,787
Neutral	0.952	0.948	0.950	7,481
Macro average	0.947	0.947	0.947	32,575
Weighted average	0.946	0.946	0.946	32,575

The confusion matrix in Table D.3 shows that the model correctly classified the majority of advertisements in each category. The most common misclassifications were between female and male categories, with 613 female-targeted advertisements incorrectly classified as male-targeted, and 400 male-targeted advertisements incorrectly classified as female-targeted.

Table D.3: Confusion Matrix

	Predicted Female	Predicted Male	Predicted Neutral
True Female	12,488	613	206
True Male	400	11,235	152
True Neutral	199	187	7,095

Application to the Analysis Dataset

While the model was trained on explicitly gendered advertisements, for the wage analysis in the main paper, I excluded all advertisements with explicit gender preferences (2.8 per cent of the total dataset). This methodological choice focuses the analysis on implicit rather than explicit gender targeting, which better captures subtle linguistic patterns that might influence wage offers without overt discrimination. After applying the trained model to the remaining advertisements without explicit gender mentions, 28 per cent were classified as female-targeted, 70 per cent as male-targeted, and 2 per cent as neutral. This substantial gender imbalance in the predicted classifications provides important context for interpreting the wage effects across gender categories and reflects implicit gender associations in occupational language.

Performance Across Occupations

To validate the model's consistency across different occupational contexts, I evaluated its performance on subsets of advertisements grouped by Standard Occupational Classification (SOC) codes. Table D.4 presents a selection of occupations with their corresponding accuracy and gender distribution ratios.

The model maintained high accuracy across most occupational categories, with accuracies ranging from 81.8 to 98.6 per cent. Notably, the model performed well in occupations with both balanced and imbalanced gender distributions. For instance, in

Table D.4: Model Performance Across Selected Occupations

SOC	Occupation Type	N	Acc.	F. Ratio	M. Ratio
213	IT Prof.	1,324	0.968	0.554	0.415
223	Healthcare Prof.	2,129	0.964	0.309	0.356
354	Sales Prof.	1,580	0.965	0.478	0.291
356	Public Services	1,630	0.971	0.483	0.270
614	Caring Services	2,324	0.818	0.331	0.338
721	Customer Service	930	0.984	0.859	0.114
821	Transport Drivers	917	0.977	0.110	0.796
927	Elementary Services	1,483	0.986	0.365	0.183

Notes: SOC = Standard Occupational Classification; N = Sample size; Acc. = Accuracy; F. Ratio = Female Ratio; M. Ratio = Male Ratio

Transport Drivers (SOC 821), which has a strong male bias (79.6 per cent), the model achieved 97.7 per cent accuracy. Similarly, in Customer Service Occupations (SOC 721), which has a strong female bias (85.9 per cent), the model achieved 98.4 per cent accuracy.

The lowest accuracy (81.8 per cent) was observed in Caring Personal Services (SOC 614), which has a more balanced gender distribution. This likely reflects the challenging nature of gender classification in care-related contexts, where advertisements frequently mention both male and female clients, potentially creating ambiguous patterns for the model to learn.

Word Importance Analysis

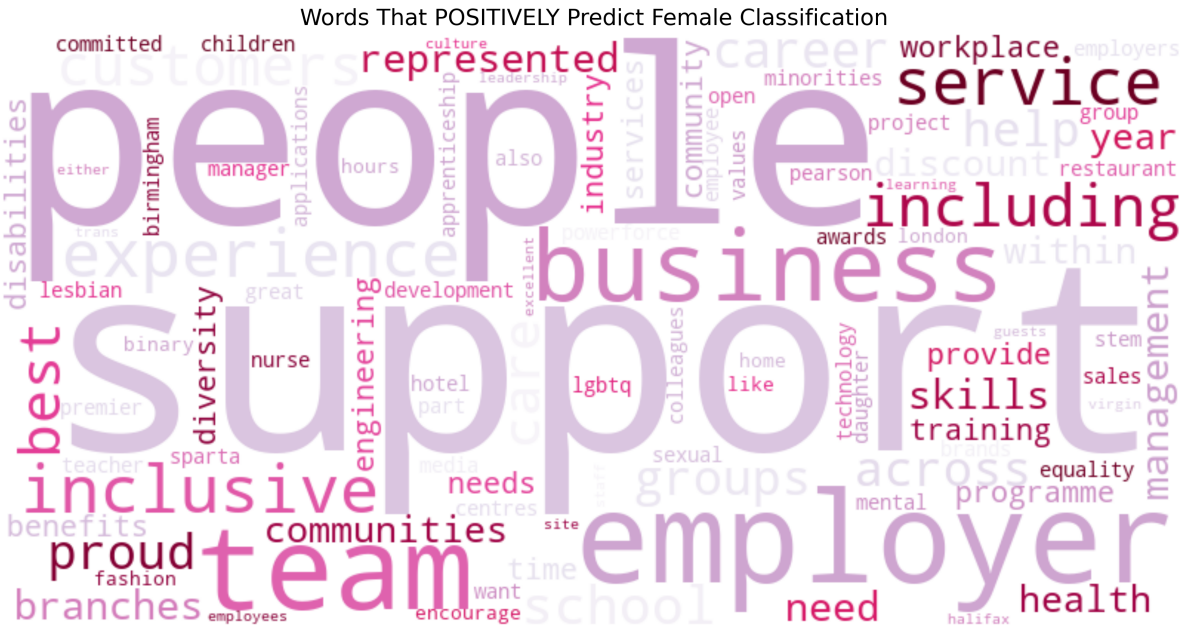
To understand the linguistic features associated with gender targeting, I extracted words with the highest predictive power from the model. Table D.5 presents the top 10 words most predictive of female-targeted, male-targeted, and neutral advertisements.

Table D.5: Top 10 Most Predictive Words by Gender Category

Female Words	Score	Male Words	Score	Neutral Words	Score
support	0.0054	support	0.0057	support	0.0085
people	0.0044	manager	0.0050	requirement	0.0078
employer	0.0034	service	0.0047	people	0.0072
team	0.0029	mail	0.0042	roles	0.0059
business	0.0027	royal	0.0023	employees	0.0049
experience	0.0025	team	0.0021	candidates	0.0045
inclusive	0.0025	delivery	0.0020	avon	0.0041
service	0.0024	people	0.0018	school	0.0036
including	0.0023	skills	0.0017	adult	0.0035
customers	0.0021	driving	0.0016	hotel	0.0035

Notably, some terms appear as predictive across multiple gender categories, though with different weights. "Support" is the top predictor across all three categories, but with a higher weight for neutral advertisements (0.0085) compared to male (0.0057) or female (0.0054) advertisements. This does not indicate a lack of differentiation between

Figure D.1: Word cloud of words predictive of female targeting



categories, but rather illustrates that gender targeting is often expressed through subtle variations in how common terms are used and combined, rather than through entirely separate vocabularies.

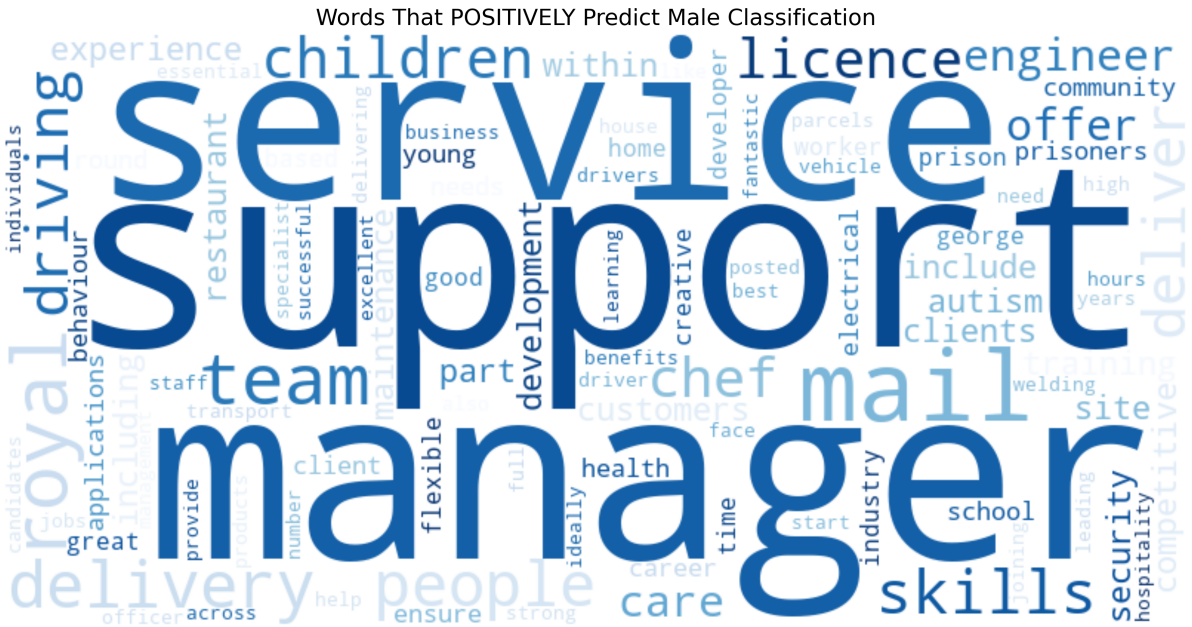
The model is likely capturing context-dependent usage patterns—the same word may appear in different linguistic environments across gender categories. For instance, "support" might co-occur with care-related terms in female-targeted advertisements, with technical terms in male-targeted advertisements, and with more formal recruitment language in neutral advertisements. The variation in predictive weights across categories (particularly the higher weight in neutral advertisements) suggests that these contextual differences are meaningful for classification.

"People" also appears in all three categories, while "team" appears in both female and male categories, again indicating that the model relies on contextual patterns rather than isolated keywords. This aligns with BERT's design as a contextual language model that considers words in relation to their surrounding text.

Despite these overlaps, distinctive patterns emerge. Female-targeted advertisements are more associated with terms like "inclusive", "employer", and "business", while male-targeted advertisements show stronger associations with "manager", "mail", "royal", and "driving". Neutral advertisements are distinctively associated with more formal recruitment terminology such as "requirement", "roles", "employees", and "candidates".

Word clouds visualizing these predictive words are shown in Figures D.1, D.2 and D.3.

Figure D.2: Word cloud of words predictive of male targeting



Limitations

The approach has several limitations:

1. Training Data Limitations: The model is trained on explicitly gendered advertisements, which may not be representative of the full advertisement population.

2. Occupation-Specific Language: Despite stratification, occupation-specific language patterns may still influence gender classification. The variability in performance across occupations, as shown in Table D.4, indicates that context-specific factors affect model accuracy.

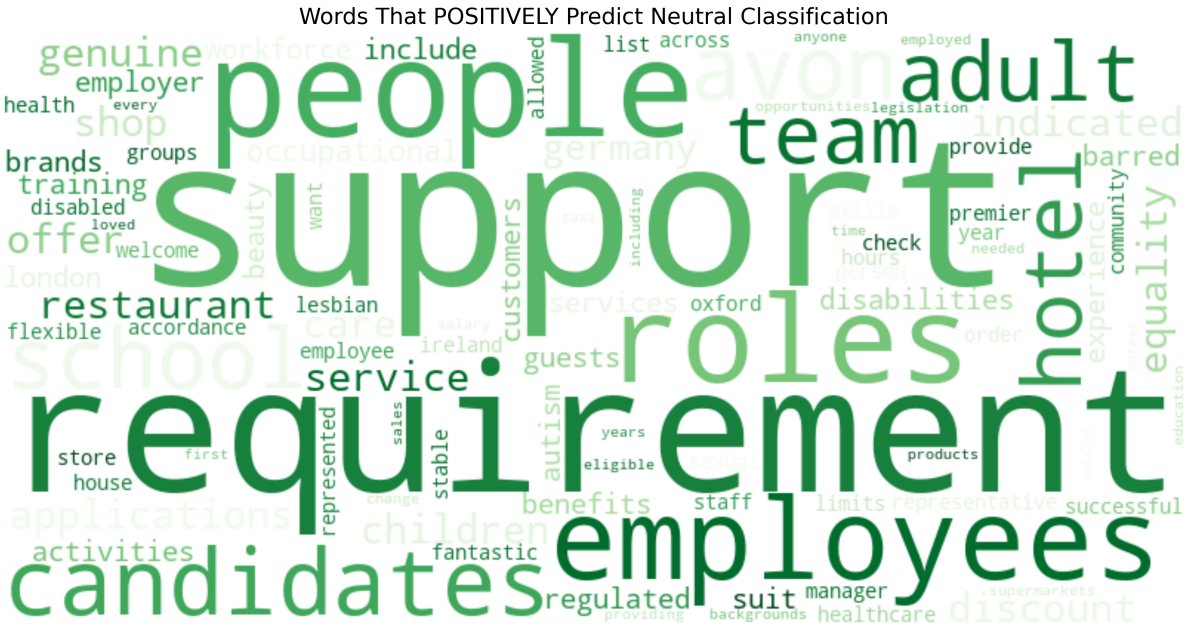
3. Reinforcement of Stereotypes: The model may reinforce existing gender stereotypes in occupational language, as it learns patterns from human-written advertisements that may contain implicit biases.

4. Binary Framework: The approach implements a primarily binary gender framework that does not account for non-binary gender identities.

5. Temporal Stability: Language patterns associated with gender may evolve over time, affecting the model's performance on advertisements from different time periods.

Despite these limitations, the model provides a valuable tool for examining the relationship between gender targeting, personality trait language, and wage offers in job advertisements.

Figure D.3: Word cloud of words predictive of neutral targeting



The effectiveness of formalisation policies in Latin America

Matias Golman[*] Verónica Escudero[†] Elva López Mourelo[‡]

Abstract

This paper presents the first comprehensive meta-analysis of formalisation policies in Latin America, analysing 79 impact evaluations with 527 estimates from interventions implemented between 1990 and 2020. We calculate standardised mean differences for 444 estimates and conduct thorough publication bias analysis using FAT-PET-PEESE methods. Unlike previous meta-analyses, we examine interventions targeting both worker and firm formalisation, include Spanish-language studies and unpublished government reports, and provide the first systematic assessment of effect magnitudes for the region. Our findings show that formalisation policies predominantly produce positive effects, with adverse impacts being rare (5.1% of estimates). However, systematic publication bias, when corrected, substantially reduces effect sizes. For formality outcomes, active labour market policies show consistent positive effects (0.238 standard deviations, $p < 0.05$), as do information campaigns (0.226 standard deviations, $p < 0.10$) and labour inspections (0.320 standard deviations, $p < 0.10$). For earnings outcomes, social protection programmes with activation components demonstrate the largest effects (1.258 standard deviations, $p < 0.01$), though negligible formality impact (0.033 standard deviations, not significant). This divergence suggests these programmes primarily affect income rather than formal employment status. Programme benefits strengthen over time, with long-term effects exceeding short-term impacts by approximately 0.10 standard deviations for formality and 0.32 standard deviations for earnings. Multi-component interventions outperform single-component policies by 0.07 standard deviations. These findings underscore the need for realistic expectations while supporting evidence-based, multifaceted approaches to formalisation.

Keywords: formalisation policies, Latin America, meta-analysis, publication bias, active labour market programmes, social protection

JEL Codes: J08, J21, J46, O17, C83

[*]University of Nottingham. Email: matias.golman@gmail.com.
[†]International Labour Organization. Email: escudero@ilo.org.
[‡]International Labour Organization. Email: lopezmourelo@ilo.org.

Contents

1 Introduction **236**

2 Theoretical framework and evidence from previous meta-analyses **238**

2.1 Theoretical perspectives on informality 239

2.2 Formalisation policies 240

2.2.1 Active labour market policies 240

2.2.2 Social protection with activation components 241

2.2.3 Information campaigns 241

2.2.4 Labour inspections 242

2.2.5 Payroll tax reductions 242

2.2.6 Simplification of formalisation processes 243

2.2.7 Combined approaches 244

3 Methodology **244**

3.1 Search and Selection Strategy 244

3.2 Data Extraction and Sample Description 245

3.3 Analytical Approaches 247

3.4 Standardised Mean Differences 247

3.5 Meta-Analytical Models 250

3.6 Publication Bias Testing and Correction 250

3.7 Meta-Regression with Covariates 252

4 Analysis and Results **253**

4.1 Sign and Significance Patterns 253

4.2 Sample Coverage and Overall Effects 253

4.3 Publication Bias 255

4.4 Conditional Effects from Meta-Regression 258

4.4.1 Specification 259

4.4.2 Programme Type Effects 259

4.4.3 Effect Measurement Timing and Programme Duration 261

4.4.4 Demographic Patterns 262

4.4.5 Methodological and Contextual Factors 262

4.4.6 Synthesis 263

5 Discussion and conclusions **263**

Appendices **279**

1 Introduction

Informality remains a defining characteristic of labour markets across Latin America, with approximately 140 of the region's 263 million workers operating in informal employment as of 2016 (Salazar-Xirinachs and Chacaltana, 2018). Following the standards of the International Labour Organisation, informal employment refers to work that is not covered by formal systems such as labour law, social security, and tax regulations (ILO, 2024). This represents a significant economic and social challenge. Informal workers lack access to social protection, face precarious working conditions, and typically earn lower and more irregular wages than their formal counterparts. Informal firms face substantial barriers to growth, including limited access to credit, restricted participation in government procurement, and exclusion from formal supply chains. At the aggregate level, informality undermines equity, efficiency, tax collection, social security coverage, productivity, and economic growth (Bertranou, 2025).

The persistence of high informality rates has prompted governments in the region to implement diverse policy interventions aimed at promoting transitions to formality. These range from active labour market policies (training courses, internships, and employment subsidies or intermediation) to institutional reforms, including simplified business registration, reduced payroll taxes, enhanced labour inspection, and information campaigns on formalisation benefits (Escudero et al., 2019; Jessen and Kluve, 2021). This diversity reflects both the multifaceted nature of informality and the competing theoretical frameworks that guide policy design. Despite substantial investment in these interventions, systematic evidence on their effectiveness remains limited. Individual impact evaluations exist, but no comprehensive effort has been made to synthesise these findings to identify which approaches work, for whom, and under what conditions.

Previous meta-analyses have examined related questions but with limitations. Escudero et al. (2019) analysed active labour market policies in Latin America, finding positive effects on formal employment, particularly for women and youth, but focused primarily on employment outcomes rather than formalisation itself. Jessen and Kluve (2021) examined formalisation interventions globally, identifying 170 impact estimates from 38 studies, but their analysis was not Latin America-specific and omitted Spanish-language studies and unpublished reports. Floridi et al. (2020) conducted a meta-analysis of formalisation interventions, but restricted their scope to firm outcomes and did not focus specifically on Latin America. In addition, none of these studies calculated standardised mean differences for the impacts of formalisation policy in the region, limiting the comparability of effect sizes in different outcomes and studies.

This paper addresses these gaps by conducting the first comprehensive meta-analysis of formalisation policies specifically for Latin America. We built an extensive database of 79 impact evaluations covering interventions implemented between 1990 and 2020, with

527 impact estimates. Our analysis makes several advances over previous work. First, we calculate standardised mean differences (Hedges' g) for 444 estimates, providing the first systematic assessment of effect magnitudes for formalisation policies in the region. Second, we examine a comprehensive range of policy interventions targeting both worker and firm formalisation, going beyond the focus on either active labour market policies or firm-level interventions in isolation. Third, we include evaluations published in Spanish alongside English-language studies, and incorporate unpublished reports often excluded from global meta-analyses. We address potential quality concerns by applying strict inclusion criteria requiring experimental or quasi-experimental methods. Fourth, we conduct a thorough publication bias analysis using FAT-PET-PEESE correction methods. We examine effectiveness across three key outcomes: probability of being formal (for employees or firms), monthly earnings, and hours worked. We employ both random-effects and fixed-effects meta-analytical models to account for heterogeneity, and conduct extensive subgroup analyses to identify conditions under which different intervention types prove most effective.

Our findings indicate that formalisation policies in the region predominantly produce positive effects, with adverse outcomes being rare (5.1% of estimates). However, uncorrected pooled estimates overstate true effects substantially. For formality outcomes, bias correction reduces estimates by approximately one-third; for earnings outcomes, correction reduces estimates by half or more, depending on the method applied. Among intervention types, active labour market policies show consistent positive impacts on formality outcomes (0.238 standard deviations, $p < 0.05$). Social protection programmes that combine income support with activation components produce the largest effects on earnings (1.258 standard deviations, $p < 0.01$) but show negligible impacts on formality itself. Information campaigns and labour inspections also show positive effects on formality, although evidence bases remain more limited.

Three additional patterns emerge from our analysis. First, programme benefits strengthen over time. Long-term effects (measured 25+ months after programme completion) exceed short-term impacts by approximately 0.10 standard deviations for formality outcomes and 0.32 standard deviations for earnings outcomes. This pattern has important implications for evaluation design, as studies with follow-up periods shorter than two years may substantially underestimate programme impacts. Second, multi-component programmes outperform single-component interventions. Programmes combining multiple policy tools (such as training with apprenticeships or labour intermediation) show effects 0.07 standard deviations larger than single-component approaches. This suggests potential complementarities between interventions that address multiple barriers to formalisation. Third, methodological choices influence reported effect sizes. Experimental evaluations report effects 0.06 to 0.09 standard deviations smaller than comparable quasi-experimental studies, likely reflecting stricter methodological standards rather than

a lower programme performance.

These results have important implications for policy design and evaluation in Latin America's formalisation agenda. Active labour market policies emerge as the most consistently effective tool for achieving formality outcomes, with substantial supporting evidence across diverse contexts. For policymakers seeking to improve earnings among vulnerable populations, social protection programmes with activation components show large effects, though these do not translate into formal employment transitions. Information campaigns offer promise as low-cost complements, particularly where knowledge gaps constitute important barriers. The strengthening of programme benefits over time underscores the importance of sustained policy implementation and longer evaluation periods. The consistent pattern of smaller effects after publication bias correction emphasises the need for realistic expectations about programme impacts.

Several evidence gaps require attention. Rigorous evaluations of payroll tax reductions, labour inspections, and simplification measures remain limited, with the evidence base heavily concentrated in active labour market policies (79% of estimates). Longer-term follow-up studies are essential, as programme benefits appear to strengthen over time. Research on combined policy approaches could identify complementarities between interventions, given that multi-component programmes show larger effects. Understanding mechanisms would enable better adaptation across contexts. Finally, firm-level formalisation requires more attention, as evidence remains concentrated on worker formalisation.

The paper proceeds as follows. Section 2 reviews conceptual foundations and existing empirical evidence on formalisation policies. Section 3 describes our methodology for data collection, coding, and meta-analytical techniques. Section 4 presents our results, including overall effect sizes and heterogeneity analysis by policy type, programme characteristics, and target population. Section 5 discusses policy implications and concludes with recommendations for future research and policy design.

2 Theoretical framework and evidence from previous meta-analyses

We define formalisation policies as interventions aimed at increasing the rate of formality among individuals and firms in Latin America. We include six types of policy interventions: (1) Active labour market policies (ALMPs), which include training, apprenticeships and internships, hiring subsidies, and labour intermediation; (2) Social protection policies with activation components; (3) Information campaigns; (4) Labour inspections; (5)

Payroll tax reductions; (6) Simplification of formalisation processes.[1] These categories reflect distinct theoretical mechanisms through which governments attempt to encourage transitions from informal to formal employment and enterprise registration.

2.1 Theoretical perspectives on informality

Understanding why informality emerges and persists is central to the theoretical and policy debates on formalisation. The conceptualisation of informality has evolved substantially since its introduction in the early 1970s. Initially associated with subsistence self-employment, the concept now encompasses a diverse range of occupations in both formal and informal economies (Bertranou, 2025). Three main theoretical perspectives have emerged to explain informality's origins and persistence, each with distinct implications for policy design.

The dualist perspective views informality as a residual sector that absorbs workers excluded from formal employment (Layard et al., 2005). Rigidities in formal labour markets, such as minimum wages, union agreements, or efficiency wage considerations, create a limited number of formal jobs. Workers who are unable to access formal employment turn to informal activities for survival (Fields, 2009). This perspective suggests that expanding formal job creation through economic growth and reducing barriers to formal employment should reduce informality. Formalisation policies under this view should focus on enhancing workers' productivity and employability to enable transitions into available formal positions.

The structuralist perspective emphasises the links between formal and informal economies (Portes et al., 1989). Informality arises from productive decentralisation processes, as formal firms subcontract production to reduce costs and increase flexibility. Informal units supply inputs and services to formal enterprises, creating functional interdependencies (Moser, 1978). Unlike the dualist view, this perspective suggests that economic growth alone may not reduce informality if formal sector expansion relies on informal production chains. Formalisation policies must therefore address both demand-side factors (formal firms' incentives to hire formally) and supply-side factors (informal workers' and firms' capacity to meet formal requirements).

The exit or legalist perspective posits that workers and firms choose informality after evaluating the costs and benefits of formalisation (Maloney, 2004; Levy, 2008). High compliance costs, complex regulations, or weak links between contributions and benefits may make informality attractive. Some workers may access social protection through non-contributory programmes or family members, reducing formal employment benefits (Perry et al., 2007). Firms may avoid registration costs and regulatory burdens by re-

[1] In our analysis, we consider a separate seventh category for policies that combine information campaigns and payroll tax reductions.

maining informal (De Soto, 1989). This perspective suggests that reducing formalisation costs and strengthening benefits should encourage transitions to formality. However, it recognises that even workers who choose informality can face poverty and precarious conditions, having selected this option as the least-bad alternative given their circumstances (Bertranou, 2025).

More recent integrated approaches recognise that the informal sector itself is heterogeneous (Fields, 1990, 2023; Ranis and Stewart, 1999). Some informal workers engage in subsistence activities with minimal formal sector connections (consistent with dualism). Others are microentrepreneurs who strategically choose partial compliance with regulations (consistent with exit). A third segment comprises workers and small firms subordinated to larger formal enterprises through subcontracting chains (consistent with structuralism). This heterogeneity implies that different policy tools may be effective for different segments of the informal sector.

2.2 Formalisation policies

Individual studies measure formality using varied operational definitions reflecting the ILO conceptual framework but adapting to data availability. For workers, common definitions include social security registration, pension system affiliation, possession of formal employment contracts, or access to employer-provided health insurance. For firms, definitions typically capture business registration with tax authorities, legal entity status, or participation in formal social security systems. We exclude studies using weak proxies such as contract type, earnings thresholds, or bookkeeping practices. Our meta-analysis addresses this heterogeneity by focussing on whether studies measure transitions towards formality, regardless of the specific operational definition employed.

2.2.1 Active labour market policies

Active labour market policies (ALMPs) address skill deficits, search frictions, and employment costs that may keep workers in informal employment (Layard et al., 2005; Kluve, 2010). Training programmes aim to improve human capital and productivity, with effects materialising as skills accumulate (Card et al., 2010). Apprenticeships and internships combine classroom learning with workplace experience, addressing skill gaps and information asymmetries about worker quality (Lerman, 2019). Labour intermediation services improve job matching efficiency and reduce search costs (Schmid et al., 2001). Employment subsidies alter relative hiring costs between formal and informal workers (Katz, 1996), although sustained impacts depend on whether temporary support enables permanent productivity increases (Abel et al., 2022).

Training programmes aimed at disadvantaged youth represent the most frequently evaluated intervention, which typically combines three months of classroom instruction

with similar-length private sector internships. Major programmes include Colombia's *Jóvenes en Acción*, Peru's *Projoven*, and the Dominican Republic's *Juventud y Empleo* (Attanasio et al., 2011; Díaz and Jaramillo, 2006; Card et al., 2011). More intensive approaches include Brazil's *PRONATEC* offering 1,200 hours of technical instruction (Camargo et al., 2021) and Chile's *FOTRAB* combining 500 classroom hours with 360 hours of on-the-job training (Doerr and Novella, 2024). Labour intermediation reforms in Colombia and Mexico transformed basic job-matching into integrated systems offering counselling, job search assistance, and placement services (Pignatti, 2016; Morales-González et al., 2019). Hiring subsidies include direct worker bonuses, wage supplements, and combined subsidy-intermediation models (Abel et al., 2022; Berniell and de la Mata, 2017; Novella and Valencia, 2022).

2.2.2 Social protection with activation components

Social protection programmes with activation components combine income support with requirements or incentives for labour market participation. Income support allows for a more effective job search by reducing the pressure to accept immediately available opportunities (Marimon and Zilibotti, 1999), while activation requirements address skills and information barriers that prevent access to formal employment.

Temporary public employment initiatives constitute the primary type of intervention, providing work on community service projects or infrastructure development alongside income support. The duration of the programme typically ranges from three to eight months. Argentina's *Jefes de Hogar* Programme, implemented during the 2002 crisis, required twenty hours of community work, training, or private company employment for household heads who received 150 pesos monthly (Gasparini et al., 2009). Peru's *Construyendo Perú* provided four-month employment on small-scale urban infrastructure projects with job training (Escudero, 2018; Macroconsult, 2012). *Uruguay Trabaja* combined six to eight months of community service with competency workshops, vocational training, health services, and documentation assistance (Blanchard et al., 2025; Nogueira Puentes, 2018).

Impact pathways involve immediate effects (participation creates temporary formal employment and income) and longer-term effects through improved employability. Skills training, work experience, and formal work history signals should increase participants' attractiveness to private formal sector employers.

2.2.3 Information campaigns

Information campaigns address knowledge gaps and misperceptions about formalisation processes, costs, and benefits. Workers and firms may remain informal due to lack of information about registration procedures, overestimation of associated costs, or unawareness

of benefits such as social protection access, credit availability, and government procurement opportunities (Bruhn and McKenzie, 2014). These interventions aim to correct knowledge deficits without altering the underlying economic incentives.

Several studies examined campaigns promoting Brazil's *Microempreendedor Individual* programme through in-person consultancy, virtual assistance through mobile messaging, and information sessions with text message reminders (Zucco et al., 2020, 2023; Lenz, 2017). Other interventions targeted domestic worker employers in Argentina through email and letter campaigns (Ohaco, 2023; Feld, 2024). Some approaches combined information with complementary elements: Brazil's *Descomplicar* tested brochures about *SIMPLES* registration procedures alongside fee waivers and free accounting services (Bruhn, 2013), while Colombia's Act No. 1429 workshops combined information provision with payroll tax reductions (Galiani et al., 2017).

These low-cost interventions are effective when information gaps constitute the primary barrier to formalisation. Where economic factors such as high taxes or procedural complexity dominate, information campaigns alone may generate limited responses (De Mel et al., 2013).

2.2.4 Labour inspections

Labour inspections aim to increase formalisation by raising the expected cost of remaining informal through monitoring and enforcement. Firms compare compliance costs (taxes, social security contributions, administrative requirements) against expected costs of non-compliance (Allingham and Sandmo, 1972; Becker, 1968). Increased inspection intensity raises detection probability, shifting calculations towards compliance. Inspections may also generate spillover effects if firms perceive generally higher enforcement (Ronconi, 2010).

Studies examine substantial variation in institutional design throughout the region, including decentralised systems with provincial variation in enforcement intensity in Argentina (Viollaz, 2018b), graduated enforcement regimes where requirements vary by firm size in Peru (Viollaz, 2018a), expansion of inspector numbers in Colombia (Pignatti, 2020), and interactions between trade liberalisation and local inspection intensity in Brazil (Ponczek and Ulyssea, 2022). While intensified enforcement raises detection probability and pushes firms towards compliance, higher labour costs may simultaneously reduce employment or shift production towards capital-intensive methods (Ronconi, 2010).

2.2.5 Payroll tax reductions

Payroll tax reductions lower non-wage labour costs associated with formal employment. In Latin American countries, formal employment typically involves employer contributions to social security, health insurance, and mandatory levies that substantially increase

total labour costs beyond gross wages. Reducing these costs narrows the price differences between formal and informal labour, making formal hiring more attractive to employers and potentially increasing workers' take-home pay in competitive markets (Gruber, 1997). Effect magnitudes depend on labour supply and demand elasticities and on whether tax savings are captured by firms, passed to workers, or shared.

Reforms show considerable heterogeneity in design. Colombia's 2012 reform reduced employer payroll taxes and health contributions by 13.5 percentage points for workers earning up to ten times the minimum wage, phased in over 2013–2014 (Bernal et al., 2017; Morales and Medina, 2017; Kugler et al., 2017). Brazil's *SIMPLES* programme consolidated multiple employer taxes into a single payment while substantially reducing overall burdens for eligible micro- and small firms (Fajnzylber et al., 2011; Monteiro and Assunção, 2012; Piza, 2018). The Individual Microentrepreneur Programme extended this to own-account workers, combining web-based registration with reduced social security contributions (Rocha et al., 2018; Pereda et al., 2022). Brazil's 2006 reform for domestic workers allowed employers to deduct social security contributions from personal income taxes, effectively reducing the employer cost of formal registration (Madalozzo and Bortoluzzo, 2011).

Lower tax rates increase formal labour demand by reducing relative costs, with magnitudes shaped by substitution possibilities between formal and informal workers (Ulyssea, 2018). Responses differ substantially between extensive margins (whether firms formalise) and intensive margins (the share of formal workers within firms), particularly when fixed registration costs remain binding for smaller enterprises.

2.2.6 Simplification of formalisation processes

Simplification interventions reduce time and administrative burden associated with formal registration, targeting barriers including number of required procedures, completion time, and compliance complexity (Djankov, 2009). Rather than changing price incentives, simplification reduces fixed formalisation costs. Complex procedures also create corruption opportunities and disadvantage less educated entrepreneurs.

Mexico's Rapid Business Opening System created centralised facilities enabling simultaneous completion of federal, state, and municipal procedures, reducing the average registration time from 30 to 1.4 days (Aparicio, 2014; Bruhn, 2013). Colombia's 2006 reform unified payment systems for health insurance and pensions, simplifying compliance (Calderón-Mejía and Marinescu, 2012). Other approaches combined simplification with tax reductions through web-based platforms, consolidating registration requirements (Rocha et al., 2018; Fajnzylber et al., 2011; Monteiro and Assunção, 2012).

Simplification generates heterogeneous responses across firm types. Registration barriers bind more tightly for retail and service firms with low capital requirements, while

manufacturing firms face additional constraints that administrative reforms alone cannot address (Bruhn, 2013; Monteiro and Assunção, 2012). Entrepreneur education levels and geographic proximity to registration offices further shape reform effectiveness.

2.2.7 Combined approaches

Many evaluated policies combine multiple intervention types, recognising that informality arises from intersecting barriers. Combined approaches may produce larger effects than individual interventions if constraints interact. Primary combinations include tax reduction with administrative simplification (Fajnzylber et al., 2011; Monteiro and Assunção, 2012; Piza, 2018; Rocha et al., 2018), regulatory change with enforcement and information campaigns (Feld, 2024), and ALMP packages that incorporate training, internships, job search assistance, and financial support (Alzuá and Brassiolo, 2006; Attanasio et al., 2011; Van Gameren, 2010).

Combined approaches present evaluation challenges: they may generate larger effects by addressing multiple constraints, but hinder identification of which components drive impacts. Most evaluations assess overall packages rather than disentangling individual element contributions, reflecting both practical constraints and the reality that many programmes were designed as integrated packages rather than separable components.

3 Methodology

3.1 Search and Selection Strategy

We selected impact evaluations through a systematic process. First, we established the following inclusion criteria: studies must assess the effects of labour formalisation policies implemented in Latin America, evaluate impacts at the individual or firm level, employ experimental or quasi-experimental methods controlling for selection into treatment, and provide sufficient statistical information (coefficient, standard error, sample size) to calculate effect sizes.

We searched multiple sources to identify relevant studies: international repositories (3ie Impact Evaluation Studies Registry and Inter-American Development Bank), academic databases (Scopus, Web of Science, SSRN, IDEAS, Google Scholar, and social-protection.org), studies from previous meta-analyses by Card et al. (2010), Kluve (2010), Escudero et al. (2019), Jessen and Kluve (2021), and Floridi et al. (2021), and evaluations conducted by government agencies within the region.

Search terms included "formalisation", "registration", "(in)formality", "labour intermediation", "public employment service", "labour inspection", "hiring subsidies", "job training", "social protection", "conditional transfer programmes", "wage subsidies", "payroll tax reduction" and "active policies" in Spanish or English. We conducted searches

without restrictions on publication date, capturing studies from the earliest available to the most recent. For each study identified, we conducted forward and backward citation searches using snowballing throughout the entire process. We completed the search at the end of February 2024, identifying 152 studies eligible for inclusion.

Applying the selection criteria reduced the sample from 152 studies to 79 studies with 527 estimates. We excluded 45 studies at the paper level (removing 150 estimates) because they did not meet fundamental inclusion criteria: social protection programmes without activation components (67 estimates), studies not presenting estimations for probability of being formal (38 estimates), studies not evaluating formalisation policy effectiveness as a primary objective (30 estimates), and studies employing non-experimental identification methods (15 estimates).

We then applied estimate-level refinement to the remaining 112 studies, excluding 1,302 estimates for methodological and definitional reasons. We then applied estimate-level refinement to the remaining 112 studies, excluding 1,302 estimates for methodological and definitional reasons. Major exclusions included estimates from individual programme cohorts or intake rounds when aggregate or initial cohort treatment effects were available (247 estimates); for instance, when a study reported separate impacts for the 2005, 2006, and 2007 intakes of a training programme, we retained only the pooled estimate across all cohorts or, when no aggregate was available, the first cohort reported. Other major exclusions included outcome variables outside our scope (427 estimates) and cases where only combined treatments were available when disaggregated effects were required (166 estimates). Methodological exclusions included estimates lacking standard errors (110 estimates) and multiple formality definitions (95 estimates). We created 19 additional pooled gender estimates for studies reporting separate male and female effects but lacking overall estimates. Table A.1 in Appendix A provides a detailed breakdown of all exclusions.

3.2 Data Extraction and Sample Description

We extracted information systematically from studies meeting our inclusion criteria: (i) analysed intervention; (ii) outcome measuring intervention effectiveness; (iii) study characteristics (authors, title, type, and year of publication); (iv) intervention characteristics (country, target population, implementation period, duration, and number of components); (v) empirical analysis (methodological approach, data source, type of estimate and time horizon); and (vi) estimated impact (coefficient, standard errors, treatment, and control group sizes).

When studies reported separate estimates for different interventions, outcome variables, subgroups, or time horizons, we recorded each estimate separately. However, we applied consistency rules to avoid over-representation of individual studies. For studies

presenting multiple treatment variations or cohorts, we extracted the aggregate treatment effect rather than the disaggregated results. When multiple identification strategies were reported, we selected the estimate identified as preferred or constituting the main focus of the findings. When multiple definitions of formality were used, we selected the estimate based on the most comprehensive definition.

Two authors selected studies and constructed the database independently, working on the same sample separately and then comparing the results to minimise bias. Additional technical details on data construction procedures, including the handling of interaction terms, the creation of pooled estimates, and sample size adjustments for standardised mean difference calculations, are provided in Appendix A.

Our final sample includes 79 impact evaluation studies with 527 impact estimates from formalisation policies implemented in Latin America between 1990 and 2020. Table 1 presents key characteristics of the sample. On average, each study contributes 6.7 estimates distributed among three outcome categories (formality, hours worked, earnings) and disaggregated by gender, age, and time horizon. This represents the most comprehensive collection of formalisation policy evaluations in the LAC region to date.

The temporal distribution shows growing policy activity, with 44% of estimates evaluating interventions from the 2010s and 34% from the 2000s. Unlike previous global meta-analyses that excluded Spanish-language studies or focused solely on published papers, our systematic search captured a comprehensive regional coverage: 19% of the studies are in Spanish and 44% are published in peer-reviewed journals. This inclusive approach addresses the limitations noted by Jessen and Kluve (2021) regarding Spanish-language studies and unpublished reports.

At the estimate level, 31% employ experimental designs while 69% use quasi-experimental methods. Among quasi-experimental approaches, difference-in-differences is more prevalent (35%), followed by instrumental variables (15%), propensity score matching (11%), and regression discontinuity (8%). Geographic coverage spans major LAC sub-regions: Southern Cone (46%), Andean countries (44%), and Central America and Caribbean (10%). Notably, 63% of estimates evaluate nationally implemented policies, while 37% focus on sub-national interventions.

Active labour market policies represent the largest category (79% of estimates). Our broader scope includes social protection programmes with activation components (10%), payroll tax policies (3%), information campaigns (3%), labour inspections (2%), and simplification measures (1%). Within ALMPs, 88% involve training programmes alone or in combination with other policies. The most common configuration combines training with apprenticeships or internships (49% of ALMP estimates). At the estimate level, 63% evaluate multi-component programmes while 37% assess single-component interventions.

The target population analysis reveals clear policy priorities: 53% focus on youth interventions, 34% target unemployed and job seekers, 8% focus on firms and self-employed

workers, and 5% examine the general population. The outcome coverage shows that 59% examine formality, 10% examine hours worked, and 30% examine monthly earnings. Gender analysis is well-represented: 23% provide male-specific results, 24% provide female-specific estimates, and 53% report pooled results. Regarding age, 51% focus on youth populations (aged 15 to 24 or 15 to 30) while 49% examine pooled adult populations (18+). Temporal scope varies: 52% capture short-term effects (0 to 12 months), 22% examine medium-term impacts (13 to 24 months), and 26% assess long-term effects (25+ months).

3.3 Analytical Approaches

We employ two complementary analytical methods. First, we conduct a preliminary sign and significance analysis that examines patterns in the direction and statistical significance of the effects of treatment. This approach requires only the sign and significance reported for each estimate, making it applicable to the full sample of 527 estimates regardless of whether the studies report complete sample size information. Inspired by Escudero et al. (2019) and Kluve et al. (2019), we create a binary outcome variable coded as 1 if the estimate shows a statistically significant positive effect and 0 otherwise, then estimate linear probability models examining how the likelihood of obtaining positive significant results varies with study characteristics.

Second, we calculate standardised mean differences for all estimates with available treatment and control group sample sizes. This forms our primary analytical approach, preserving effect size information and enabling explicit publication bias correction.

3.4 Standardised Mean Differences

Studies in our sample assess the impacts of formalisation policy using three main outcomes: probability of being formal, monthly earnings, and hours worked. Each outcome is measured differently across studies. Table 2 presents the distribution of the outcome variables across the entire sample and by policy type.

The probability of being formal presents the greatest definitional variation. Our sample includes measures from both the employee and the firm perspectives. The definitions of formality for employees include the registration of social security (*RAIS* in Brazil, *SIPA* in Argentina, *PILA* in Colombia), the contribution to the pension and health insurance systems, the possession of formal contracts, and the access to social protection benefits. Firm formality measures include business registration with tax authorities, possession of municipal or state licences, registration in programmes such as *MEI* in Brazil, and compliance with regulatory obligations.

Similarly, earnings measures vary between levels and logarithmic transformations, capturing individual monthly wages, firm-level average wages, or regional wage indicators.

Table 1: Characteristics of included estimates

Characteristic	Freq.	%	Characteristic	Freq.	%
Policy type			Outcome category		
ALMPs	417	79.1%	Formality	313	59.4%
Labour inspections	11	2.1%	Hours worked	55	10.4%
Payroll tax reduction	13	2.5%	Monthly earnings	159	30.2%
PTR + Simplification	10	1.9%			
Sensitization	16	3.0%	Target population		
Simplification	5	0.9%	Youth	279	52.9%
Social protection	55	10.4%	Unemployed and job seekers	181	34.3%
			Firms and self-employed	43	8.2%
Sub-region			General workers	24	4.6%
South Cone	240	45.5%			
Andean countries	233	44.2%	Programme components		
Central America	36	6.8%	Single component	194	36.8%
Caribbean	18	3.4%	Multiple components	333	63.2%
Geographical coverage			Gender		
National	333	63.2%	All	279	52.9%
Sub-national	194	36.8%	Men	119	22.6%
			Women	129	24.5%
Evaluation period					
1990s	96	18.2%	Age		
2000s	180	34.2%	Pooled (18+)	256	48.6%
2010s	230	43.6%	Youth (15–24/30)	271	51.4%
2020s	21	4.0%			
			Timing of follow-up		
Method used			Short-term (0–12 months)	275	52.2%
DID	186	35.3%	Medium-term (13–24 months)	117	22.2%
IV	80	15.2%	Long-term (25+ months)	135	25.6%
PSM	59	11.2%			
RD	41	7.8%			
Randomised	161	30.6%	Total estimates	527	

Notes: ALMPs refers to Active Labour Market Policies, including hiring subsidies, apprenticeships, training programmes, and employment services. PTR refers to Payroll Tax Reduction. Social protection refers to policies with activation components (e.g., cash transfers conditional on training or job search). Formality includes both worker and firm formalisation outcomes. Youth (15–24/30) aggregates estimates targeting individuals aged 15–24 or 15–30, while Pooled (18+) includes estimates for the general adult population. Timing of follow-up refers to the period between programme participation and outcome measurement. Studies can contribute multiple estimates through disaggregation by outcome, time horizon, or population subgroup.

Table 2: Estimates availability by policy type and outcome

Policy Type	Formality	Hours	Income	Total
ALMPs	240	40	137	417
Labour inspections	6	2	3	11
Payroll tax	10	0	3	13
Payroll+Simplification	10	0	0	10
Sensitization	16	0	0	16
Simplification	5	0	0	5
Social protection	26	13	16	55
Total	313	55	159	527

Notes: Rows represent policy categories: ALMPs = Active Labour Market Programmes (training, apprenticeships, labour intermediation, hiring subsidies); Labour inspections = enforcement measures to increase formalization compliance; Payroll tax = reductions in employer social security contributions; Payroll + Simplification = combined interventions reducing taxes and streamlining registration procedures; Sensitization = information campaigns; Simplification = business registration and administrative procedure reforms; Social protection = programmes with activation components linking benefits to formalization. Columns represent outcome measures: Formality = probability of formal employment or firm registration; Hours = weekly hours worked (total or formal sector); Income = monthly earnings or wages (total or formal).

Hours worked estimates span weekly hours in levels or logarithms, with some studies focussing on formal sector hours specifically. Table B.1 in Appendix B provides a complete list of the definitions of formality, earnings, and hours worked used in all studies.

To enable systematic comparison across these different measurement approaches, we calculate standardised mean differences (SMD) for all estimates with available treatment and control group sample sizes. The SMD represents the difference between treatment and control group means expressed in standard deviation units, providing a scale-free metric that facilitates comparison across studies with different outcome measures (Higgings et al., 2023).

We use Hedges' g, which includes a small-sample bias correction, making it preferable to Cohen's d for studies with limited sample sizes. Hedges' g is calculated as:

$$g = \frac{Y_t - Y_c}{S_p} \times \left[1 - \frac{3}{4(df) - 1}\right] \tag{1}$$

where $Y_t - Y_c$ represents the estimated treatment effect, S_p denotes the pooled standard deviation, and the bracketed term provides the small-sample correction factor. Since pooled standard deviations are rarely reported in primary studies, we calculate S_p using:

$$S_p = SE \times \sqrt{\frac{n_c + n_t}{n_c \times n_t}} \times \sqrt{n_c + n_t - 2} \tag{2}$$

where n_t and n_c represent treatment and control group sample sizes, and SE is the standard error of the estimated treatment effect. A standardised effect of 0.2 is generally

considered small, 0.5 medium, and 0.8 large, regardless of the original metric or baseline rate.

3.5 Meta-Analytical Models

Our analysis follows the comprehensive approach recommended by Irsova et al. (2024) for modern meta-analysis. We begin by presenting summary measures of overall effects across formalisation policy categories and outcomes. This is followed by a heterogeneity analysis and a publication bias assessment. We then estimate conditional effects through meta-regression models that correct for publication bias while controlling for study characteristics.

Meta-analysis employs two primary statistical models: fixed-effects and random-effects. In meta-analysis terminology, these terms differ from panel data econometrics. The distinction concerns whether studies represent a fixed set (the entire population of interest) or a random sample from a larger population. Fixed-effects models assume that all studies share the same true effect size, with observed variation reflecting only sampling error. These models use weights $w_j = 1/\sigma_j^2$, where σ_j^2 is the within-study variance. Random-effects models allow true effects to vary across studies due to genuine differences in implementation, context, or populations. These models incorporate between-study heterogeneity through weights $w_j = 1/(\sigma_j^2 + \tau^2)$, where τ^2 captures variation across studies beyond sampling error.

We present results from both specifications. Random-effects models are appropriate when true effects genuinely vary across studies due to differences in design, implementation, and context, as is common in policy evaluation research. The model explicitly incorporates between-study variance (τ^2), producing correct standard errors and confidence intervals when heterogeneity is present (Borenstein et al., 2021). However, random-effects models can produce biased estimates when publication selection is present (Stanley and Doucouliagos, 2011; Picchio, 2020). Our publication bias correction methods address this limitation.

Following Irsova et al. (2024), we also report unrestricted weighted least squares (UWLS) estimates. UWLS provides a simple summary statistic that uses inverse variance weighting without imposing additional structure on between-study heterogeneity. UWLS avoids the efficiency-bias tradeoff inherent in FE and RE models, providing a weighted average that does not require assumptions about the structure of heterogeneity between studies (Stanley, 2008).

3.6 Publication Bias Testing and Correction

Publication bias arises when researchers systematically under-report results that contradict theoretical expectations or policy objectives. In formalisation policy evaluation,

researchers and policymakers may be reluctant to report small or statistically insignificant effects, particularly when substantial resources have been invested in programme implementation. When imprecise but large positive estimates are treated more favourably than imprecise but small estimates, a systematic bias emerges that overstates the impact of the programme in the literature (Stanley and Doucouliagos, 2011).

Our sample includes both published peer-reviewed studies and unpublished working papers, reducing but not eliminating publication bias. Academic journals may favour studies based on effect size magnitude and statistical significance rather than methodological quality alone (Stanley and Doucouliagos, 2011; Brodeur et al., 2016).

We assess publication bias through visual inspection, formal statistical tests, and multiple bias correction approaches. Visual inspection relies on funnel plots that display effect sizes against precision measures (inverse standard errors). In the absence of selection bias, studies should be distributed symmetrically around the mean effect, with greater dispersion among less precise estimates, creating an inverted funnel shape. Asymmetry, particularly sparse representation of small effects when precision is low, indicates systematic under-reporting.

For formal testing, we implement the Funnel Asymmetry Test-Precision Effect Test (FAT-PET) methodology (Egger et al., 1997; Stanley, 2008). This regresses the effect sizes on their standard errors:

$$g_i = \alpha + \beta \times SE(g_i) + \varepsilon_i \tag{3}$$

where g_i denotes the standardised effect size for the estimate i. The intercept α provides the precision-effect test (PET): the estimated effect when the precision is perfect ($SE = 0$), representing the average effect adjusted for publication bias. The slope coefficient β provides the funnel asymmetry test (FAT): a test of whether the effect sizes systematically relate to their standard errors. A statistically significant β coefficient indicates funnel plot asymmetry, suggesting publication bias. The model uses weighted least squares with inverse variance weights ($1/SE(g_i)^2$) to address known heteroskedasticity.

Following Irsova et al. (2024), we employ multiple approaches to estimate Equation 3. We estimate the FAT-PET model using weighted least squares following Stanley (2008), ordinary least squares with standard errors clustered at the study level to account for within-study dependencies, fixed effects specifications, instrumental variables using sample size as an instrument for standard errors to address potential endogeneity, and precision-weighted regression using inverse standard errors as weights. This multimethod approach reveals the sensitivity of conclusions to different assumptions.

When FAT-PET indicates nonzero effects, we use the Precision Effect Estimate with Standard Error (PEESE) model, which employs quadratic approximation to minimise bias (Stanley and Doucouliagos, 2014). PEESE replaces the standard error term with

the variance term:

$$g_i = \alpha + \beta \times SE(g_i)^2 + \varepsilon_i \tag{4}$$

The quadratic specification corrects for the non-linear relationship between effect sizes and their variances when publication bias is present. Linear FAT-PET (Equation 3) underestimates the true effect when the genuine effect is non-zero because it assumes a constant relationship between effect size and standard error. PEESE allows this relationship to vary with the level of imprecision, which better approximates how publication selection operates in practice (Stanley and Doucouliagos, 2014). The PEESE constant term provides the bias-corrected-effects estimate. We implement PEESE with clustered standard errors at the study level.

3.7 Meta-Regression with Covariates

Beyond publication bias correction, we address additional sources of effect heterogeneity by including comprehensive control variables. We account for within-study dependencies through clustered standard errors at the study level. Study-specific factors include programme type, duration, and multi-component design indicators, along with participant characteristics such as age and gender distributions. Methodological controls account for differences in evaluation design between studies. We include GDP growth rates in the year of programme implementation to control for macroeconomic conditions that may influence labour market transitions and programme effectiveness.

To incorporate these factors while maintaining publication bias correction, we extend the FAT-PET specification to include covariates:

$$g_i = \alpha + \beta \times SE(g_i) + \gamma X_i + \varepsilon_i \tag{5}$$

where i indexes individual estimates, X_i represents the covariate vector, and γ contains the estimates of the corresponding parameters. For PEESE specifications with covariates:

$$g_i = \alpha + \beta \times SE(g_i)^2 + \gamma X_i + \varepsilon_i \tag{6}$$

Our preferred specification uses the PEESE methodology with random-effects meta-regression, incorporating clustered standard errors at the study level to account for dependencies within the study estimate, following Irsova et al. (2024) recommendations for robust inference in meta-analysis.

4 Analysis and Results

We present our meta-analytical findings in four parts. First, we conduct a preliminary sign and significance analysis examining the distribution of positive, negative, and insignificant results across the entire sample. Second, we examine the overall effect sizes and heterogeneity patterns using standardised mean differences. Third, we assess publication bias using multiple diagnostic approaches. Fourth, we estimate conditional effects through a meta-regression analysis that simultaneously corrects for publication bias and controls for study characteristics.

4.1 Sign and Significance Patterns

We conducted a preliminary sign and significance analysis on the full sample of 527 estimates. This approach examines whether estimates are statistically significant and positive without requiring complete sample size information. Of all estimates, 287 (54.5%) show statistically significant positive effects, 213 (40.4%) are statistically insignificant, and 27 (5.1%) show statistically significant negative effects. Linear probability models regressing a binary indicator for positive significant results on study characteristics reveal three notable patterns. First, experimental evaluation designs consistently show large negative associations with the probability of reporting significant positive results, with coefficients ranging from -0.301 to -0.349 ($p < 0.01$), likely reflecting stricter methodological standards rather than worse programme performance. Second, information campaigns show positive associations of 0.505 to 0.543 ($p < 0.10$). Third, country fixed effects substantially increase explanatory power from $R^2 = 0.130$ to $R^2 = 0.223$, indicating country-level factors explain considerable variation beyond observable programme characteristics.

Most other factors show no consistent relationship with the likelihood of obtaining positive significant results. Programme type differences between ALMPs, social protection, and other interventions are not statistically significant. Similarly, neither the timing of effect measurement (short, medium, or long-term), programme duration, nor demographic targeting (gender or age focus) systematically predicts whether a study reports positive significant findings. This preliminary analysis provides descriptive context but discards information about effect magnitudes and does not correct for publication bias, motivating our focus on standardised mean differences in the remainder of this section. The complete results appear in Appendix B.

4.2 Sample Coverage and Overall Effects

From the 527 estimates available in our full sample, 444 (84.3%) disclosed sample sizes for treatment and control groups, allowing calculation of standardised mean differences using Hedges' g. Table 3 presents the distribution of estimates with SMD data across pro-

gramme types and outcome categories. Active labour market policies dominate with 366 estimates (82% of the SMD sample), while social protection programmes with an activation component contribute 55 estimates. Several categories show limited representation: labour inspections (1 estimate), payroll tax interventions (5 estimates), and simplification measures (0 estimates). This concentration limits our ability to draw conclusions about these types of intervention.

Table 3: Estimates with Hedge'g availability by policy type and outcome

Policy Type	Formality	Hours	Income	Total
ALMPs	222	31	113	366
Labour inspections	1	0	0	1
Payroll tax	3	0	2	5
Payroll+Simplification	2	0	0	2
Sensitization	15	0	0	15
Simplification	0	0	0	0
Social protection	26	13	16	55
Total	269	44	131	444

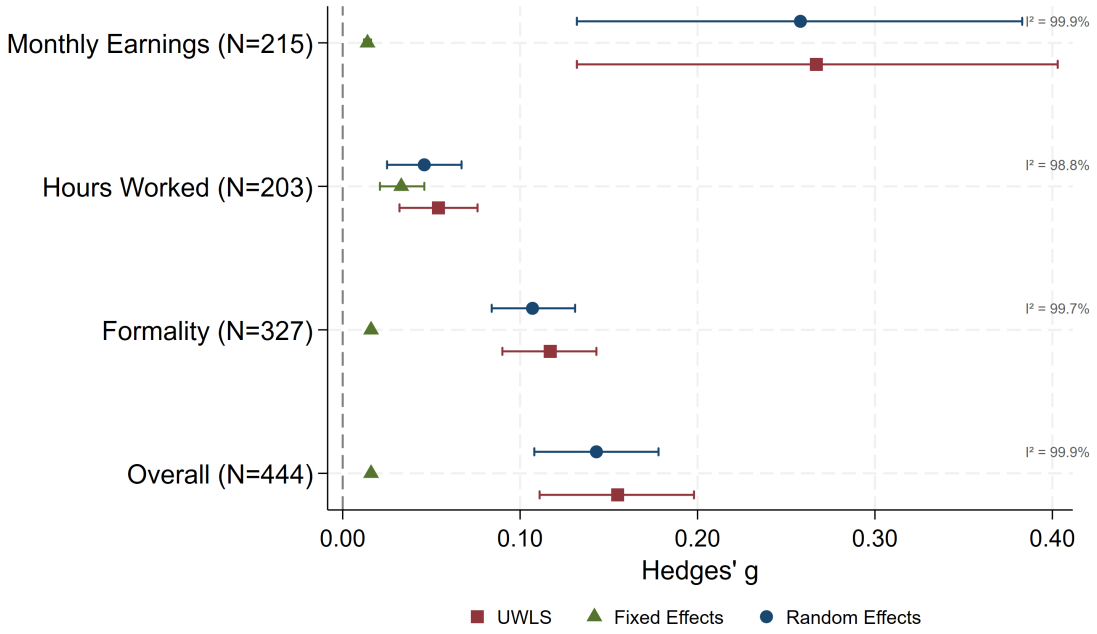
Notes: Rows represent policy categories: ALMPs = Active Labour Market Programmes (training, apprenticeships, labour intermediation, hiring subsidies); Labour inspections = enforcement measures to increase formalization compliance; Payroll tax = reductions in employer social security contributions; Payroll+Simplification = combined interventions reducing taxes and streamlining registration procedures; Sensitization = information campaigns and awareness-raising activities; Simplification = business registration and administrative procedure reforms; Social protection = programmes with activation components linking benefits to formalization. Columns represent outcome measures: Formality = probability of formal employment or firm registration; Hours = weekly hours worked (total or formal sector); Income = monthly earnings or wages.

Figure 1 presents pooled effect sizes across all studies using the three estimation approaches described in Section 3.5. The random effects estimate for the pooled sample shows a positive and statistically significant effect of 0.143 standard deviations, representing approximately one-seventh of a standard deviation improvement in outcomes. The substantial difference between fixed effects (0.016) and random effects estimates indicates considerable between-study heterogeneity. The I^2 statistic of 99.9% confirms that variation across studies reflects genuine differences in treatment effects rather than sampling error alone.

The effects vary by outcome type. The formalisation outcomes show consistent and statistically significant effects in all estimation methods: UWLS of 0.117, FE of 0.017, and RE of 0.107. Hours worked shows smaller effects with RE of 0.046, suggesting that formalisation policies primarily affect employment status and earnings rather than work intensity. Monthly earnings shows the largest point estimates with RE of 0.258, but with wide confidence intervals reflecting high uncertainty.

Programme-specific analyses appear in Appendix B Table B.2. ALMPs produce

Figure 1: Meta-Analytic Effect Sizes by Outcome Category



Notes: Meta-analytic effect sizes with 95% confidence intervals comparing three estimation methods: unrestricted weighted least squares (UWLS), fixed effects (FE), and random effects (RE). Effect sizes represent standardised mean differences (Hedges' g). Sample sizes (N observations) shown in parentheses on y-axis. I^2 statistics indicate percentage of variation across studies due to heterogeneity rather than chance, shown for random effects estimates. The "Overall" row presents effects pooled across all three outcome categories. Sample includes estimates from 79 formalisation policy evaluations in Latin America (1990-2020).

consistent positive and statistically significant effects across all outcomes, with formality impacts of 0.117 and monthly earnings effects of 0.114. Social protection programmes show large earnings effects of 1.264, although this estimate relies on limited observations. Information campaigns show modest but statistically significant formality effects of 0.015.

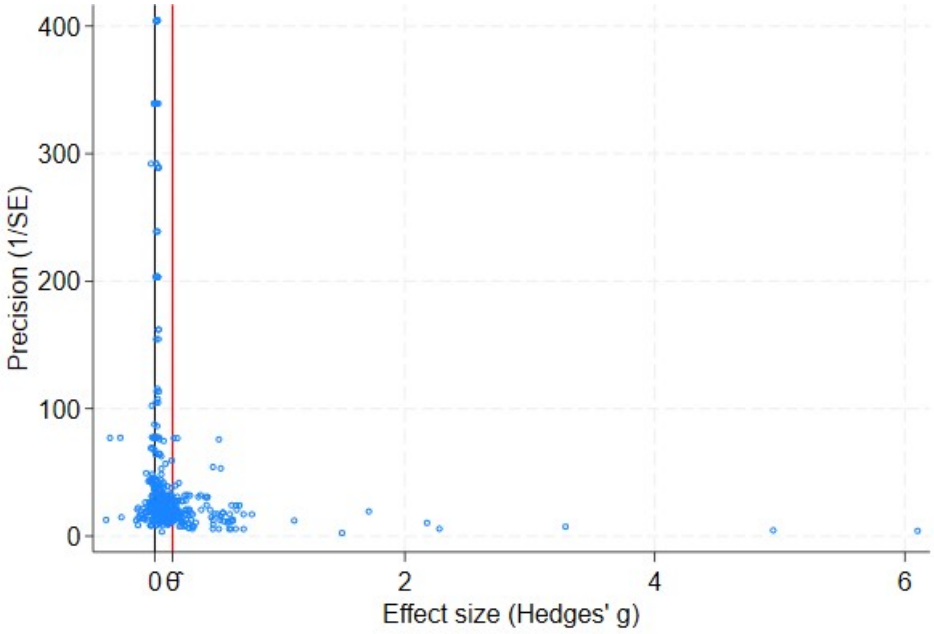
The outlier analysis following Havránek et al. (2020) criteria identifies 52 extreme observations (11.7% of the SMD sample) with effect sizes outside the $[-1, 1]$ range or precision measures beyond reasonable bounds. Removing these outliers reduces the overall mean effect from 0.155 to 0.115 and substantially decreases variability (the standard deviation falls from 0.464 to 0.176), indicating that the main results are not driven by extreme studies.

4.3 Publication Bias

We assess publication bias through funnel plot inspection and formal statistical tests. Figures 2 and 3 show funnel plots for pooled outcomes, with and without outliers. Both plots exhibit clear asymmetry, with studies concentrated in the bottom-right quadrant (large positive effects with low precision) and sparse representation in the bottom-left quadrant (small or negative effects with large standard errors). Approximately 73% of

the estimates fall in the right half of the funnel plot, compared to the 50% expected under symmetry. This pattern shows systematic under-representation of studies finding smaller effects when precision is low, consistent with publication selection favouring studies that report larger positive impacts. The asymmetry persists after outliers are excluded.

Figure 2: Funnel plot: Hedges' g with pooled estimate



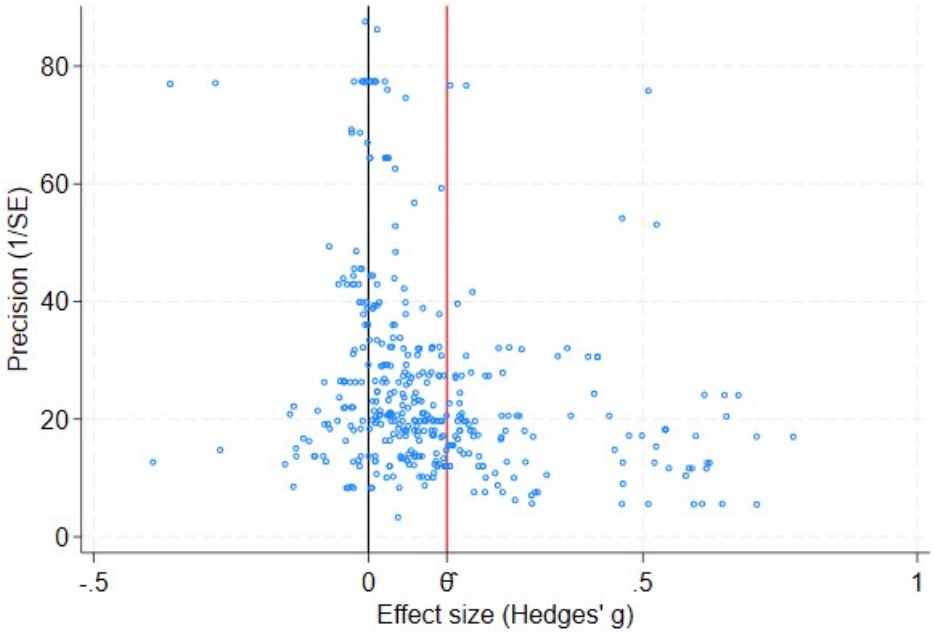
Notes: θ (red line) shows pooled random-effects estimate.

The Egger test provides formal confirmation of these visual patterns. For the pooled sample (N=444), both random effects and fixed effects models strongly reject the null hypothesis of no publication bias (Appendix B Table B.2). Publication bias appears in all outcome categories, with particularly strong evidence for monthly earnings.

Table 4 presents bias-corrected estimates using the multiple approaches described in Section 3.6. Each row represents a different method to estimate the FAT-PET model (Equation 3). The constant term provides the precision-effect test (PET), representing the bias-corrected average effect. The standard error coefficient provides the funnel asymmetry test (FAT), which tests for publication bias.

The results show sensitivity to methodological choices. The weighted least squares method following Stanley (2008) generates a bias-corrected effect of 0.004, representing a 97% reduction from the uncorrected pool estimate of 0.143. The standard error coefficient of 2.412 is statistically significant at the 1% level, providing strong evidence of publication bias. When accounting for study-level clustering through ordinary least squares with clustered standard errors, the bias-corrected estimate becomes −0.117, although statistically insignificant. The standard error coefficient remains large at 5.267 and significant at the 10% level, suggesting publication selection may entirely account for observed positive effects once within-study dependencies are addressed.

Figure 3: Funnel plot: Hedges' g with pooled estimate, outliers excluded



Notes: θ (red line) shows pooled random-effects estimate. Outliers excluded based on Havránek et al. (2020) criteria.

The fixed effects specification produces a bias-corrected estimate of −0.246 with a standard error coefficient of 7.767, significant at the 5% level. The instrumental variable approach, using sample size as an instrument for standard errors to address potential endogeneity, results in a bias-corrected estimate of 0.013 with a standard error coefficient of 2.752, significant at the 5% level. The precision-weighted approach produces a bias-corrected effect of −0.011 and a standard error coefficient of 3.219, significant at the 1% level.

All five approaches reject the null hypothesis of no publication bias at conventional significance levels through the FAT test, while none of the bias-corrected estimates from the PET test differ significantly from zero. The wide range of bias-corrected estimates (from −0.246 to 0.013) reflects the uncertainty of the model about the correct specification for publication bias correction.

Following Stanley and Doucouliagos (2014), when PET rejects the null of a zero effect, the PEESE estimator should be used, as it better approximates the quadratic relationship between effect sizes and their standard errors. Since our PET estimates do not reject zero in most specifications, we present PEESE primarily for comparison. The PEESE-corrected estimate of 0.071, significant at the 5% level with clustered standard errors, suggests effects approximately 50% smaller (0.071 versus 0.143) than the uncorrected random-effects pooling indicates. The PEESE standard error squared coefficient of 18.927 confirms the presence of publication bias, although with less precision than the linear FAT-PET specifications.

Table 4: FAT-PET-PEESE Analysis: All Outcomes Pooled

Model	**Constant (PET)**	**SE Coefficient (FAT)**
WLS (Stanley 2008)	0.004	2.412***
	(0.003)	(0.276)
OLS Clustered	-0.117	5.267*
	(0.113)	(2.895)
Fixed Effects	-0.246	7.767**
	(0.195)	(3.785)
IV (Sample Size)	0.013	2.752**
	(0.026)	(1.108)
Precision (1/SE)	-0.011	3.219***
	(0.018)	(1.146)
PEESE Analysis		
PEESE Clustered	0.071**	18.927
	(0.032)	(11.962)

Notes: N=444 from 79 studies. The table presents publication bias assessment following Equation 3 (rows 1–5) and Equation 4 (row 6). PET (Precision Effect Test): The constant term provides the bias-corrected average effect, representing the estimated effect when precision is perfect (SE=0). FAT (Funnel Asymmetry Test): The SE coefficient tests whether effect sizes systematically relate to their standard errors; a significant coefficient indicates publication bias. WLS implements inverse variance weighting following Stanley (2008). OLS Clustered uses study-level clustered standard errors. Fixed Effects includes study-level fixed effects. IV instruments standard errors with sample size. Precision weights by 1/SE. PEESE uses variance (SE^2) instead of SE and is recommended when PET rejects zero effect (Stanley and Doucouliagos, 2014). Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Bias patterns vary by outcome. Formality outcomes show PEESE-corrected effects of 0.087, significant at the 1% level (Appendix B Table B.3). Hours worked exhibit mixed signals with a PEESE-corrected estimate of 0.022, not statistically significant. Monthly earnings display the strongest bias, with PEESE correction resulting in −0.097, significant at the 10% level, suggesting publication selection may entirely account for observed positive effects in this outcome category. Detailed outcome-specific results and funnel plots appear in Appendix B (Tables B.3 to B.5, Figures B.1 to B.3).

These findings indicate that conclusions about formalisation policy effectiveness depend critically on addressing publication selection. The substantial reductions in effect sizes after bias correction suggest the uncorrected literature overstates policy impacts, particularly for earnings outcomes.

4.4 Conditional Effects from Meta-Regression

We now examine which programme features are associated with larger or smaller effects after correcting for publication bias and controlling for study characteristics. The meta-regression simultaneously estimates the independent association of each characteristic while holding others constant.

4.4.1 Specification

The meta-regression model follows Equation 6:

$$g_i = \alpha + \beta \times SE(g_i) + \gamma X_i + \varepsilon_i$$

where g_i is Hedges' g for estimate i, $SE(g_i)$ corrects for publication bias, X_i represents study characteristics (programme type, target population, evaluation design, duration, outcome category, and contextual factors), and ε_i is the error term. The subscript i indexes the 444 estimates with SMD data from 79 studies. Standard errors cluster at the study level to account for dependencies between multiple estimates from the same evaluation.

The baseline category for programme types is combined interventions that pair payroll tax reductions with administrative simplification of registration procedures (such as Brazil's SIMPLES programme, which reduced tax burdens while streamlining the registration process). The coefficients represent differences relative to this baseline.

4.4.2 Programme Type Effects

Table 5 presents meta-regression results with correction for publication bias. We focus first on programme type effects, then examine other dimensions of heterogeneity.

For formality outcomes (columns 3 to 4), active labour market policies show robust positive effects under both specifications: 0.238 ($p < 0.05$) for random effects and 0.169 ($p < 0.01$) for fixed effects. This makes ALMPs the most consistently effective policy tool relative to the baseline category. The convergence among the estimation methods reinforces this finding.

Labour inspections show the largest point estimates for formality: 0.320 ($p < 0.10$) for random effects and 0.146 ($p < 0.01$) for fixed effects, although the random effects estimate shows considerable uncertainty. The single observation in the SMD sample limits the interpretation. Information campaigns show positive formality effects under random effects of 0.226 ($p < 0.10$), although the fixed effects estimate of 0.019 is not statistically significant. This suggests that information-based approaches may produce modest improvements in formality, but the evidence is less robust than for ALMPs.

Payroll tax reductions alone show positive but statistically insignificant effects on formality under random effects of 0.186, while under fixed effects, the coefficient is significant but modest in magnitude (0.095, $p < 0.01$). The coefficients suggest payroll tax policies alone do not substantially outperform the combined baseline approach. Social protection programmes show no significant formality effects: 0.033 for random effects and -0.021 for fixed effects.

However, for earnings outcomes (columns 5 to 6), social protection programmes show

Table 5: Meta-regression results: SE publication bias correction

	All Outcomes		Formality		Earnings		Hours Worked	
	RE	FE	RE	FE	RE	FE	RE	FE
(i) Programme type (base: payroll + simplification)								
ALMPs	0.290	0.158***	0.238**	0.169***	0.352	0.105***	−0.203	−0.203
	(0.204)	(0.028)	(0.109)	(0.028)	(0.313)	(0.017)	(0.207)	(0.207)
Social protection	0.699***	0.140***	0.033	−0.021	1.258***	0.588***		
	(0.210)	(0.030)	(0.117)	(0.032)	(0.347)	(0.031)		
Labour inspections	0.499	0.149***	0.320*	0.146***				
	(0.347)	(0.052)	(0.183)	(0.052)				
Sensitization	0.204	0.023	0.226*	0.019				
	(0.233)	(0.042)	(0.127)	(0.042)				
Payroll tax	0.186	0.085***	0.186	0.095***				
	(0.236)	(0.028)	(0.137)	(0.030)				
(ii) Effect duration (base: long-term)								
Medium-term	−0.097**	0.000	−0.046	0.003	−0.317***	−0.085***	−0.010	−0.010
	(0.042)	(0.002)	(0.029)	(0.002)	(0.105)	(0.010)	(0.033)	(0.033)
Short-term	0.061	0.006***	−0.054*	0.006***	−0.194	−0.037***	0.017	0.017
	(0.044)	(0.002)	(0.029)	(0.002)	(0.130)	(0.010)	(0.040)	(0.040)
(iii) Outcome category (base: hours worked)								
Formality outcome	0.073	0.056***						
	(0.052)	(0.007)						
Earnings outcome	0.121**	0.053***						
	(0.054)	(0.007)						
(iv) Target group (base: men, pooled samples)								
Female	0.012	0.008***	0.007	0.012***	−0.058	−0.002	−0.002	−0.002
	(0.041)	(0.002)	(0.028)	(0.002)	(0.099)	(0.003)	(0.026)	(0.026)
Gender pooled	0.056	0.010***	0.011	0.012***	0.131	0.006**	−0.010	−0.010
	(0.038)	(0.001)	(0.026)	(0.002)	(0.093)	(0.003)	(0.030)	(0.030)
Youth(18-24)	−0.024	−0.011**	0.007	−0.028***	0.042	0.125***	0.012	0.012
	(0.039)	(0.004)	(0.025)	(0.005)	(0.108)	(0.014)	(0.048)	(0.048)
(v) Programme duration (base: long duration)								
Short duration	−0.040	−0.013*	−0.021	0.014	−0.129	−0.039***		
	(0.056)	(0.007)	(0.041)	(0.010)	(0.132)	(0.013)		
Medium duration	−0.265***	−0.009	−0.010	0.045***	−0.656***	−0.177***	−0.173	−0.173
	(0.052)	(0.007)	(0.039)	(0.010)	(0.126)	(0.015)	(0.147)	(0.147)
(vi) Evaluation design (base: non-experimental)								
Experimental design	−0.059	−0.064***	−0.090***	−0.070***	−0.324**	−0.138***	0.014	0.014
	(0.045)	(0.004)	(0.030)	(0.005)	(0.144)	(0.011)	(0.082)	(0.082)
(vii) Contextual factors								
GDP growth	0.023***	0.005***	0.004	0.007***	0.026*	0.008***	0.013	0.013
	(0.004)	(0.000)	(0.003)	(0.001)	(0.014)	(0.001)	(0.013)	(0.013)
(viii) Publication bias correction								
Standard Error (SE)	6.639***	3.435***	1.974***	3.265***	11.588***	5.388***	−0.314	−0.314
	(0.503)	(0.092)	(0.397)	(0.116)	(1.156)	(0.211)	(1.398)	(1.398)
Constant	−0.512**	−0.143***	−0.142	−0.109***	−0.356	−0.078***	0.306	0.306
	(0.227)	(0.030)	(0.125)	(0.031)	(0.335)	(0.023)	(0.261)	(0.261)
Pub. bias correction	*SE*	*SE*	*SE*	*SE*	*SE*	*SE*	*SE*	*SE*
Observations	392	392	233	233	119	119	40	40

Notes: Meta-regression with SE publication bias correction. RE = Random Effects, FE = Fixed Effects. Full specification includes all covariates and country fixed effects. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

strong positive effects under both specifications: 1.258 ($p < 0.01$) for random effects and 0.588 ($p < 0.01$) for fixed effects. This substantially exceeds impacts on formality itself. ALMPs show positive earnings impacts of 0.352 for random effects and 0.105 ($p < 0.01$) for fixed effects, although the random effects estimate does not reach statistical significance.

Hours worked outcomes (columns 7 to 8) show limited evidence of programme effectiveness differences. The small sample (N=40) and the absence of observations for most categories of programmes prevent reliable conclusions.

4.4.3 Effect Measurement Timing and Programme Duration

The timing of effect measurement and the duration of the programme show distinct patterns that require careful interpretation. These represent different dimensions: effect measurement timing refers to when outcomes are measured relative to programme completion (follow-up period), while programme duration refers to how long the intervention itself lasted.

The timing of the effect measurement shows important patterns. For formality outcomes, short-term effects (measured 0 to 12 months after programme completion) show negative associations relative to long-term measurement: -0.054 ($p < 0.10$) for random effects and 0.006 ($p < 0.01$) for fixed effects. Medium-term effects show smaller and statistically insignificant associations. This indicates that formality impacts take time to materialise. The longer follow-up periods tend to report larger effects.

For earnings, the pattern is more pronounced. Medium-term effects show large negative coefficients relative to long-term measurement: -0.317 ($p < 0.01$) for random effects and -0.085 ($p < 0.01$) for fixed effects. Short-term effects also show negative associations. This indicates earnings impacts also strengthen considerably over time.

These temporal patterns likely reflect several mechanisms. Skills accumulate and employers observe productivity over time. Job search takes time to find good matches. The stigma of unemployment or informality fades. Networks develop through formal employment experience. These processes suggest that short-term evaluations substantially underestimate programme impacts, particularly for earnings outcomes.

Programme duration shows strong associations with earnings outcomes. Medium-duration interventions show large negative coefficients relative to long-duration programmes: -0.656 ($p < 0.01$) for random effects and -0.177 ($p < 0.01$) for fixed effects. Short-duration programmes also show negative associations. This implies that sustained programme implementation is particularly important for achieving earnings gains.

For formality outcomes, programme duration shows mixed patterns with no clear relationship. The coefficients for short-duration and medium-duration programmes are small and statistically insignificant.

4.4.4 Demographic Patterns

Gender patterns in formality outcomes remain modest. Women-specific estimates show small positive effects: 0.007 for random effects and 0.012 ($p < 0.01$) for fixed effects. Gender-pooled estimates display similarly small positive associations. The programmes report similar formality effects for participants in women and men.

Youth-targeted estimates show coefficients of 0.007 for random effects and -0.028 ($p < 0.01$) for fixed effects. This indicates that youth-focused programmes report smaller formality effects relative to programmes targeting broader age groups, not that youth programmes are ineffective in absolute terms. The positive association in the sign-and-significance analysis (Appendix B) confirms that youth programmes produce positive effects; the meta-regression coefficient indicates that these effects are smaller than for general-population programmes.

For earnings results, youth-targeted estimates show positive associations of 0.042 for random effects and 0.125 ($p < 0.01$) for fixed effects, while female-specific estimates show small negative coefficients. This indicates systematic differences in how programmes affect earnings across demographic groups.

4.4.5 Methodological and Contextual Factors

Experimental evaluation designs show strong negative associations with effect sizes across all outcomes. For formality, the coefficient is -0.090 ($p < 0.01$) for random effects and -0.070 ($p < 0.01$) for fixed effects. For earnings, experimental designs show -0.324 ($p < 0.05$) for random effects and -0.138 ($p < 0.01$) for fixed effects. This likely reflects lower attrition bias and reduced selection on unobservables in experimental designs. Randomised experiments prevent systematic sorting of high-ability participants into treatment, which can inflate quasi-experimental estimates when unobserved characteristics correlate with both treatment assignment and outcomes. Experimental studies report effect sizes approximately 0.06 to 0.09 standard deviations smaller than comparable quasi-experimental evaluations.

GDP growth shows positive associations with effectiveness in the pooled sample: 0.023 ($p < 0.01$). For earnings specifically, the coefficient is 0.026 ($p < 0.10$) for random effects and 0.008 ($p < 0.01$) for fixed effects. This suggests that formalisation policies may be more effective in boosting earnings when implemented during periods of economic expansion.

The publication bias correction term (standard error) shows strong positive and statistically significant coefficients across all outcome categories, confirming the systematic overestimation of effects in the uncorrected literature. For formality, the coefficient is 1.974 ($p < 0.01$) for random effects and 3.265 ($p < 0.01$) for fixed effects. For earnings, the publication bias term reaches 11.588 ($p < 0.01$) for random effects, indicating a

particularly severe publication bias in the earnings literature.

4.4.6 Synthesis

Conditional effect estimates provide several policy-relevant insights. Active labour market policies emerge as the most effective tool for achieving formality outcomes. The effect size of approximately 0.24 standard deviations falls between conventional benchmarks for small (0.2) and medium (0.5) effects. For formality outcomes, where the underlying variable measures the probability of formal employment, this corresponds to an increase of roughly 12 to 14 percentage points in the likelihood of being formally employed. Social protection programmes with activation components show large earnings effects (1.26 standard deviations) but limited formality impacts, suggesting that these interventions primarily affect income rather than transitions to employment status. Labour inspections and information campaigns show positive formality effects, although the evidence is less robust given limited observations.

Programme benefits strengthen over time. Long-term effects exceed short- and medium-term impacts across outcome categories. This pattern is most pronounced for earnings, where medium-term effects are 0.32 standard deviations smaller than long-term effects. For formality, short-term evaluations report effects 0.05 standard deviations smaller than long-term evaluations. These temporal patterns underscore the importance of extended follow-up periods in impact evaluations.

The consistent pattern of smaller effects after correction for publication bias emphasises the need for realistic expectations about the impact of programmes. Uncorrected pooled estimates overstate true effects considerably. Bias correction reduces formality estimates by roughly one-third and earnings estimates by half or more. This has important implications for programme planning and resource allocation. Robustness checks using alternative publication bias correction methods and outcome-specific meta-regressions appear in Appendix B (Tables B.8 to B.11). These alternative specifications produce qualitatively similar results, with the relative ranking of programme effectiveness and the direction of key associations remaining consistent across methods. The descriptive heterogeneity analysis by programme type, demographic targeting, and implementation features also appears in Appendix B.

5 Discussion and conclusions

This meta-analysis provides the first comprehensive synthesis of formalisation policy evidence for Latin America, analysing 79 impact evaluations with 527 estimates from interventions implemented between 1990 and 2020. We address gaps in the literature by including Spanish-language studies and unpublished reports, examining interventions tar-

geting both worker and firm formalisation, and calculating standardised mean differences for 444 estimates while controlling for publication bias.

Main Findings

Three main findings emerge. First, formalisation policies predominantly report positive effects, with adverse outcomes being rare (5.1% of estimates). However, systematic publication bias substantially reduces effect sizes when corrected. Uncorrected pooled estimates substantially overstate true effects. After correction, formality effects fall by approximately one-third, while earnings effects are reduced by half or more depending on the correction method.

Second, programme effectiveness varies substantially across intervention types. Active labour market policies show the most consistent positive effects on formality, producing improvements that fall between small and medium effect sizes by conventional standards. In practical terms, these effects correspond to increases of 12 to 14 percentage points in the probability of formal employment. Information campaigns also produce positive effects, though more modest in magnitude and with weaker statistical confidence. Labour inspections show promising point estimates but rely on limited evidence. For earnings outcomes, social protection programmes with activation components produce the largest effects, substantially larger than formality impacts, but show negligible effects on transitions to formal employment. This divergence indicates that these programmes primarily affect income through direct transfers and improved job quality rather than by helping workers move into formal positions.

Third, the benefits of the programme strengthen over time. Long-term effects exceed short- and medium-term impacts across outcomes. For earnings, medium-term effects are 0.32 standard deviations smaller than long-term effects ($p < 0.01$). Programme duration matters: medium-duration interventions produce effects 0.27 standard deviations smaller than long-duration programmes ($p < 0.01$). Multi-component programmes outperform single-component interventions (0.173 versus 0.101 standard deviations).

Mechanisms and Interpretation

The observed patterns suggest different operating mechanisms across intervention types. Active labour market policies work through human capital accumulation, reduced search frictions, and lower hiring costs. Results show that these mechanisms translate more directly into formality outcomes than earnings gains. Information campaigns address knowledge gaps about registration procedures and formalisation benefits, explaining their modest but consistent formality effects. Labour inspections operate through deterrence, raising the expected costs of non-compliance.

Social protection programmes with activation components show large earnings effects without corresponding formality gains. Income support allows for a more effective job search and may finance investments in self-employment or skill acquisition. These mechanisms increase earnings in both formal and informal work, explaining why earnings impacts substantially exceed formality effects. The temporary nature of these programmes (typically 3 to 8 months) may limit formal employment transitions while allowing income improvements.

The heterogeneity in the effectiveness of the programme is aligned with integrated theoretical perspectives that recognise multiple segments of the informal sector (Fields, 2023; Bertranou, 2025). The effectiveness of ALMPs' effectiveness for formality supports human capital and employability mechanisms emphasised in dualist perspectives. The positive effects of information campaigns are consistent with exit theories that highlight knowledge barriers. The persistence of informality despite positive programme effects reflects structuralist insights about formal-informal links that economic growth alone cannot break.

The temporal strengthening of the effects likely reflects several mechanisms. Skills accumulate, and employers observe productivity over time. Job search takes time to find good matches. The stigma of unemployment or informality fades. Networks develop through formal employment experience. Multi-component programmes show larger effects potentially because they address multiple binding constraints simultaneously, creating complementarities between skills and job search support.

Our demographic findings differ from Escudero et al. (2019), who found stronger effects for women and youth. We observe modest gender differences (0.012 standard deviations for female-specific estimates, not significant) and negative associations for youth-targeted interventions (-0.028 standard deviations, $p < 0.01$ for formality). These divergences reflect our broader policy scope beyond ALMPs, explicit publication bias corrections, and a focus on formality rather than employment. The negative youth coefficient indicates that youth programmes report smaller formality effects than programmes targeting broader age groups, not that youth programmes are ineffective in absolute terms.

Experimental designs consistently report smaller effects than quasi-experimental studies (0.09 standard deviations smaller for formality, $p < 0.01$). This likely reflects lower attrition bias and reduced selection on unobservables in experimental designs. Randomised experiments prevent systematic sorting of high-ability participants into treatment, which can inflate quasi-experimental estimates when unobserved characteristics correlate with both treatment assignment and outcomes. GDP growth shows positive associations with effectiveness (0.023 standard deviations per percentage point, $p < 0.01$), confirming that formalisation policies work better during economic expansions.

Policy Implications

Several practical recommendations emerge. First, prioritise active labour market policies combining training with apprenticeships or labour intermediation for formality outcomes. These show consistent positive effects with substantial supporting evidence. Second, consider information campaigns as low-cost complements where knowledge gaps constitute barriers. Cost-effectiveness considerations favour information campaigns despite their modest effect sizes. These low-cost interventions may offer superior value relative to resource-intensive ALMPs or social protection programmes, although rigorous cost-benefit analysis remains scarce. Third, use social protection programmes with activation components to support vulnerable populations, recognising that these primarily affect earnings rather than formality status.

Fourth, invest in longer-term programmes. The benefits of the programme are strengthened over time, and sustained interventions are more effective than brief ones. Fifth, plan evaluation periods extending beyond 24 months to capture full impact trajectories. Short-term evaluations substantially underestimate programme impacts, particularly in terms of earnings. Sixth, maintain realistic expectations. After correcting for publication bias, even the most effective interventions produce modest formality effects. Policymakers should therefore view these programmes as useful components of a broader formalisation strategy rather than standalone solutions. No single intervention will dramatically reduce informality; sustained effort across multiple policy domains remains essential.

Finally, allocate resources for a rigorous evaluation of the intervention types understudied. The evidence is primarily concentrated on active labour market policies (79% of estimates). Rigorous evaluations of payroll tax reductions, labour inspections, and simplification measures remain scarce despite suggestive evidence of effectiveness.

Limitations and Future Research

Our analysis faces several limitations. Excluding studies without complete statistical information may introduce selection bias, although including unpublished studies and evaluations in Spanish mitigates this concern. The predominance of ALMP evaluations reflects historical priorities rather than necessarily indicating superior effectiveness. High heterogeneity ($I^2 > 99\%$) indicates that unobserved factors such as implementation quality, local labour market tightness, informal sector structure, enforcement capacity, and complementary policies play crucial roles that our covariates cannot fully capture. This reinforces the need for context-sensitive policy adaptation rather than one-size-fits-all approaches. Pooling estimates across different formality definitions may mask important distinctions between dimensions of formalisation.

Future research should address several priorities. First, conduct rigorous evaluations

of understudied intervention types, particularly payroll tax reductions, labour inspections, and simplification measures. Second, examine combined policy approaches to understand which combinations work best and whether effects are additive or synergistic. Third, investigate mechanisms through which interventions affect outcomes. Understanding why interventions work enables better adaptation across contexts. Fourth, assess implementation quality and its relationship to effectiveness. Implementation quality likely matters substantially but remains understudied. Finally, examine firm-level formalisation more extensively, as evidence concentrates on worker outcomes.

Concluding Remarks

Informality remains a defining challenge for Latin American labour markets. Our meta-analysis demonstrates that formalisation policies produce positive outcomes, but effect sizes are modest and vary across contexts and intervention types. The substantial publication bias we document underscores the need for realistic expectations. After bias correction, the effects remain positive, but smaller than the uncorrected estimates suggest.

Active labour market policies show consistent evidence of the effectiveness of formality outcomes. Social protection programmes produce large earnings effects but limited formality transitions. Information campaigns and labour inspections show promise, but require more rigorous evaluation. No single intervention provides a complete solution. Formalisation represents a gradual process that requires sustained effort in multiple policy domains, combining targeted interventions with broader macroeconomic policy, regulatory reform, and institutional strengthening.

The high heterogeneity in our results suggests substantial scope for understanding what makes interventions work in specific contexts. Context-specific factors, implementation quality, and complementary policies are likely to play crucial roles. Future progress depends on continued investment in rigorous impact evaluation, particularly for understudied intervention types, combined with efforts to understand mechanisms and implementation processes. The current evidence supports cautious optimism: formalisation policies work, but realistic expectations about impact magnitudes and attention to implementation remain essential for success.

References

Abel, M., Carranza, E.and Geronimo, K., and Ortega, M. (2022). Can temporary wage incentives increase formal employment? Experimental evidence from Mexico. Discussion Paper 15740, Institute of Labor Economics (IZA), Bonn.

Aedo, C. and Pizarro, M. (2004). Rentabilidad económica del programa de capacitación laboral de jóvenes Chile Joven. Report, INACAP and Mideplan, Santiago de Chile.

Alcázar, L. and Huerta, S. (2023). Mejorando la empleabilidad de mujeres urbanas vulnerables en tiempos de pandemia en el Perú: evaluación experimental del componente de capacitación virtual en un programa de empleo temporal. Documento de Investigación 127, Grupo de Análisis para el Desarrollo.

Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338.

Alzúa, M., Cruces, G., and López, C. (2016). Long-run effects of youth training programs: Experimental evidence from Argentina. *Economic Inquiry*, 54(4):1839–1859.

Alzuá, M. L. and Brassiolo, P. (2006). The impact of training policies in Argentina: An evaluation of Proyecto Joven. Working Paper 15/06, Inter-American Development Bank, Washington DC.

Amarante, V., Burdín, G., Ferrando, M.and Manacorda, M., and Vernengo, A.; Vigorito, A. (2009). Informe final de la evaluación de impacto del PANES. Working paper, Ministerio de Desarrollo Social - Universidad de la República, Montevideo.

Aparicio, G. (2014). Does formality improve firm performance? Evidence from a quasi-experiment in Mexico.

Attanasio, O., Guarín, A., Medina, C., and Meghir, C. (2017). Vocational training for disadvantaged youth in Colombia: A long-term follow-up. *American Economic Journal: Applied Economics*, 9(2):131–143.

Attanasio, O., Kugler, A., and Meghir, C. (2011). Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, 3(3):188–220.

Barrera-Osorio, F., Kugler, A., and Silliman, M. (2023). Hard and soft skills in vocational training: Experimental evidence from Colombia. *The World Bank Economic Review*, 37(3):409–436.

Barron, M. (2020). Business training programs and microenterprise formalization in Perú. *Cogent Economics & Finance*, 8:1791546.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Bernal, R., Eslava, M., Meléndez, M., and Pinzón, A. (2017). Switching from payroll taxes to corporate income taxes: Firms employment and wages after the 2012 Colombian tax reform. *Economia*, 18(1):41–74.

Berniell, L. and de la Mata, D. (2017). Starting on the right track? The effects of rst job experience on short and long term labor market outcomes. Working paper 2017/26, CAF.

Bertranou, F. (2025). Informalidad laboral: evolución conceptual, medición y enfoque para la transición hacia la formalización. Nota técnica, Organización Internacional del Trabajo, Santiago de Chile. Oficina de la OIT para el Cono Sur de América Latina.

Blanchard, P., Brum, M., Carrasco, P., Parada, C., and Perazzo, I. (2025). Employment effects of a social and labour inclusion programme. *Labour Economics*, 94:102717.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2021). *Introduction to Meta-Analysis*. John Wiley & Sons.

Brodeur, A., Lé, M., Sangnier, M., and Zylberberg, Y. (2016). Star wars: The empirics strike back. *American Economic Journal: Applied Economics*, 8(1):1–32.

Bruhn, M. (2013). A tale of two species: Revisiting the effect of registration reform on informal business owners in Mexico. *Journal of Development Economics*, 103:275–283.

Bruhn, M. and McKenzie, D. (2014). Entry regulation and the formalization of microenterprises in developing countries. *The World Bank Research Observer*, 29(2):186–201.

Cabrera, J. M., Cid, A., and Bernatzky, M. (2016). The effect of one-on-one assistance on the compliance with labor regulation. a field experiment in extremely vulnerable settings. Working Paper 84639, Munich Personal RePEc Archive (MPRA).

Calderón-Mejía, V. and Marinescu, I. E. (2012). The impact of Colombia's pension and health insurance systems on informality. Discussion Paper 6439, Institute of Labor Economics (IZA).

Calero, C., Gonzalez Diez, V., Soares, Y., Kluve, J., and Corseuil, C. (2017). Can arts-based interventions enhance labor market outcomes among youth? Evidence from a randomized trial in Rio de Janeiro. *Labour Economics*, 45:131–142.

Camargo, J., Lima, L., Riva, F., and Souza, A. P. (2021). Technical education, non-cognitive skills and labor market outcomes: Experimental evidence from Brazil. *IZA Journal of Labor Economics*, 10(1):1–34.

Card, D., Ibarrarán, P., Regalia, F., Rosas-Shady, D., and Soares, Y. (2011). The labor market impacts of youth training in the Dominican Republic. *Journal of Labor Economics*, 29(2):267300.

Card, D., Kluve, J., and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal*, 120(548):452–477.

Card, D., Kluve, J., and Weber, A. (2018). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3):894–931.

Carvalho, C. C., Corbi, R., and De-Losso, R. (2021). Estimating the employment and educational effects of vocational training: the role of school quality. Working Paper 2021-10, University of São Paulo (FEA-USP), São Paulo.

Chaparro, J., García, G., and Cardona, M. (2020). Evaluación de impacto de EMPLEANDO FUTURO. Primer bono de impacto social en América Latina. Evaluación de impacto, SIBs.CO.

Corseuil, C. H., Foguel, M. N., and Gonzaga, G. (2019). Apprenticeship as a stepping stone to better jobs: Evidence from Brazilan matched employer-employee data. *Labour Economics*, 57:177–194.

Costa, G. (2022). *Interventions on human capital formation among vulnerable populations: experimental evidence from two large-scale programs in Brazil*. PhD thesis, Doctoral dissertation. Sao Paulo Getulio Vargas Foundation.

Da Mata, D., Oliveira, R., and Silva, D. (2025). Who benefits from job training programs? Evidence from a high-dosage program in Brazil. *Journal of Development Economics*, 175:103476.

De Andrade, G. H., Bruhn, M., and McKenzie, D. (2016). A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms. *The World Bank Economic Review*, 30(1):24–54.

de Barros, R., Biderman, C., Costa, G. W., Lima, L., and Souza, A. (2023). Can at-risk youth be rescued? Experimental evidence of a nationwide youth empowerment program. Available at SSRN: https://ssrn.com/abstract=5223527.

De Mel, S., McKenzie, D., and Woodruff, C. (2013). The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, 5(2):122–150.

De Soto, H. (1989). *The Other Path: The Invisible Revolution in the Third World*. Harper & Row, New York.

Delajara, M., Freije-Rodriguez, S., and Soloaga, I. (2013). Evaluation of training for the unemployed in Mexico: learning by comparing methods. Working Paper 55210, Munich Personal RePEc Archive (MPRA).

Departamento Nacional de Planeación (2008). Subprograma Jóvenes en Acción: Consultoría para la evaluación de impacto del Subprograma Jóvenes en Acción. Evaluación de Políticas Públicas 9, Departamento Nacional de Planeación, Dirección de Evaluación de Políticas Públicas, Bogotá, D.C.

Díaz, J. and Rosas, D. (2016). Impact evaluation of the job youth training program Projoven. Working Paper 693, Inter-American Development Bank, Washington DC.

Díaz, J. J., Chacaltana, J., Rigolini, J., and Ruiz, C. (2018). Pathways to formalization: going beyond the formality dichotomy. Discussion Paper 11750, Institute of Labor Economics (IZA).

Díaz, J. J. and Jaramillo, M. (2006). An evaluation of the Peruvian "youth labor training program" Projoven. Working paper 10/06, Inter-American Development Bank.

Djankov, S. (2009). The regulation of entry: A survey. *The World Bank Research Observer*, 24(2):183–203.

Doerr, A. and Novella, R. (2024). The long-term effects of job training on labor market and skills outcomes in Chile. *Labour Economics*, 91:102619.

Durand, G. (2018). Bolivia and the program to support employment: an impact evaluation of its conditional cash transfer component. Master's thesis, School of Economics and Management, Lund University.

Egger, M., Smith, G. D., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109):629–634.

Escudero, V. (2018). Workfare programmes and their delivery system: Effectiveness of construyendo Perú. Working paper 39, International Labour Organization (ILO).

Escudero, V., Kluve, J., López Mourelo, E., and Pignatti, C. (2019). Active labour market programmes in Latin America and the Caribbean: Evidence from a meta analysis. *The Journal of Development Studies*, 55(12):2644–2661.

Fajnzylber, P., Maloney, W. F., and Montes-Rojas, G. V. (2011). Does formality improve micro-firm performance? Evidence from the Brazilan SIMPLES program. *Journal of Development Economics*, 94(2):262–276.

Feld, B. (2024). Direct and spillover effects of enforcing labor standards: Evidence from Argentina. *Journal of Human Resources*, 59(5):1503–1534.

Fields, G. S. (1990). Labour market modelling and the urban informal sector: Theory and evidence. In Fields, G. S. and Pieters, J., editors, *Employment and Development: How Work Can Lead From and Into Poverty*, pages 145–172. Oxford University Press.

Fields, G. S. (2009). Segmented labor market models in developing countries. In Kincaid, H. and Ross, D., editors, *The Oxford Handbook of Philosophy of Economics*, pages 476–508. Oxford University Press.

Fields, G. S. (2023). *The Job Ladder: Transforming Informal Work and Livelihoods in Developing Countries*. Oxford University Press.

Floridi, A., Demena, B. A., and Wagner, N. (2020). Shedding light on the shadows of informality: A meta-analysis of formalization interventions targeted at informal firms. *Labour Economics*, 67(101925).

Floridi, A., Demena, B. A., and Wagner, N. (2021). The bright side of formalization policies. meta-analysis of the benefits of policy-induced versus self-induced formalization. *Applied Economics Letters*, 28(20):1807–1812.

Galdo, J. and Chong, A. (2012). Does the quality of public-sponsored training programmes matter? Evidence from bidding processes data. *Labour Economics*, 19:970–998.

Galiani, S., Meléndez, M., and Ahumada, C. N. (2017). On the effect of the costs of operating formally: New experimental evidence. *Labour Economics*, 45:143–157.

Garsous, G., Corderi, D., and Velasco, M. (2015). Tax incentives and job creation in the tourism industry of Brazil. Working paper 644, Inter-American Development Bank.

Gasparini, L., Haimovich, F., and Olivieri, S. (2009). Labor informality bias of a poverty-alleviation program in Argentina. *Journal of Applied Economics*, 12(2):181–205.

Gómez, M. F. and González-Velosa, C. (2023). Can a pay-for-performance program help the vulnerable find jobs during a pandemic?: Experimental evidence from empleate in Colombia. Working paper 1478, Inter-American Development Bank.

Gruber, J. (1997). The incidence of payroll taxation: evidence from Chile. *Journal of Labour Economics*, 15(S3):72–101.

Havranek, T., Irsova, Z., Laslopova, L., and Zeynalova, O. (2024). Publication and attenuation biases in measuring skill substitution. *Review of Economics and Statistics*, 106(5):1187–1200.

Havránek, T., Stanley, T., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W., Rost, K., and Aert, R. (2020). Reporting guidelines for meta-analysis in economics. *Journal of Economic Surveys*, 34.

Hernani-Limarino, W. L. and Villarroel, P. M. (2015). Capacitación laboral y empleabilidad Evidencia de Mi Primer Empleo Digno. *Revista Latinoamericana de Desarrollo Económico*, 13:35–76.

Higgings, J. P., Li, T., and Deeks, J. J. (2023). Choosing effect measure and computing estimates of effect. In *Cochrane Handbook for Systematic Reviews of Interventions version 6.4*, chapter 6. Cochrane.

Ibarrarán, P., Kluve, J., Ripani, L., and Rosas-Shady, D. (2019). Experimental evidence on the long-term effects of a youth training program. *ILR Review*, 72(1):185–222.

Ibarrarán, P., Ripani, L., Taboada, B., Villa, J., and Garcia, B. (2014). Life skills, employability and training for disadvantaged youth: Evidence from a randomized evaluation design. *IZA Journal of Labor & Development*, 3(10).

ILO (2024). *2024-2030 Strategy for the Promotion of Formalization in Latin America and the Caribbean*. ILO Regional Office for Latin América and the Caribbean, Geneve.

Irsova, Z., Doucouliagos, H., Havranek, T., and Stanley, T. D. (2024). Meta-analysis of social science research: A practitioner's guide. *Journal of Economic Surveys*, 38(5):1547–1566.

Jessen, J. and Kluve, J. (2021). The effectiveness of interventions to reduce informality in low- and middle-income countries. *World Development*, 138(105256).

Katz, L. F. (1996). Wage subsidies for the disadvantaged. Working paper 5679, National Bureau of Economic Research.

Kluve, J. (2010). The effectiveness of European active labor market programs. *Labour Economics*, 17(6):904–918.

Kluve, J., Puerto, S., Robalino, D., Romero, J., Rother, F., Stöterau, J., Weidenkaff, F., and Witte, M. (2019). Do youth employment programs improve labor market outcomes? A quantitative review. *World Development*, 114:237–253.

Kugler, A., Kugler, M., and Prada, L. O. H. (2017). Do payroll tax breaks stimulate formality? Evidence from Colombias reform. Working paper 23308, National Bureau of Economic Research.

Kugler, A., Kugler, M., Saavedra, J. E., and Herrera-Prada, L. O. (2022). Long-term educational consequences of vocational training in colombia: Impacts on young trainees and their relatives. *Journal of Human Resources*, 57(1):178–216.

Layard, P. R. G., Layard, R., Nickell, S. J., and Jackman, R. (2005). *Unemployment: macroeconomic performance and the labour market*. Oxford University Press.

Lenz, A.-K. (2017). *Studies on entrepreneurship and formalization in Brazil*. PhD thesis, The School of Public Business and Administration (EBAPE)-Getulio Vargas Foundation, Rio de Janeiro.

Lerman, R. (2019). Do firms benefit from apprenticeship investments? *IZA World of Labor*.

Levy, S. (2008). *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*. Brookings Institution Press, Washington, D.C.

Llerena Pinto, C., Llerena Pinto, A., and Llerena Pinto, F. (2013). Evaluación de impacto del programa mi primer empleo en el Ecuador. Evaluación de impacto, ECONÓMICA CIC Centro de Investigación, Analítica y Desarrollo de Sistemas Informáticos para las Ciencias Sociales y Administración.

López Mourelo, E. and Escudero, V. (2017). Effectiveness of active labour market tools in conditional cash tranfers programmes:evidence for Argentina. *World Development*, 94:422–447.

Macroconsult, S. A. (2012). Evaluación de impacto del programa construyendo Perú. Working paper, Ministerio de Economía y Finanzas. Unidad de Coordinación de Préstamos Sectoriales (UCPS), Lima.

Madalozzo, R. C. and Bortoluzzo, A. B. (2011). The impact of tax exemptions on labor registration: The case of Brazilan domestic workers. Working Papers 232/2011, Instituto de Ensino e Pesquisa (Insper).

Maloney, W. F. (2004). Informality revisited. *World Development*, 32(7):1159–1178.

Marimon, R. and Zilibotti, F. (1999). Unemployment vs. mismatch of talents: Reconsidering unemployment benefits. *The Economic Journal*, 109(455):266–291.

Mera, M., Camisassa, J., Baliña, J., and Méndez, E. L. (2023). Evaluación de cursos para personas emprendedoras en la ciudad autónoma de Buenos Aires. Documento de Trabajo 217, CIPPEC.

Monteiro, J. C. and Assunção, J. J. (2012). Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilan microenterprises. *Journal of Development Economics*, 99(1):105–115.

Morales, L. F. and Medina, C. (2017). Assessing the effect of payroll taxes on formal employment: The case of the 2012 tax reform in Colombia. *Economía*, 18(1):75–124.

Morales-González, K., Ávila, W. M. d., and Cruz-Almanza, S. d. l. (2019). Evaluación del servicio público de empleo: sus efectos en la inserción laboral formal en el área Metropolitana de Barranquilla, Colombia. *Lecturas de Economía*, (91):211–239.

Moser, C. O. N. (1978). Informal sector or petty commodity production: Dualism or independence in urban development. *World Development*, 6(9-10):1041–1064.

Nogueira Puentes, R. (2018). Las políticas activas del mercado de trabajo: revisión de las experiencias de evaluación y una estimación del impacto del programa "Uruguay Trabaja". Master's thesis, Udelar. FCEA.

Novella, R., Rucci, G., Vazquez, C., and Kaplan, D. (2017). Training vouchers and labour market outcomes in Chile. *Labour*, 32(2):243–260.

Novella, R. and Valencia, H. (2022). Active labor market policies in a context of high informality: The effect of pae in Bolivia. *The Journal of Development Studies*, 58(12):2583–2603.

Ohaco, M. (2023). Registración del servicio doméstico: Evaluación de impacto de una política de difusión dirigida a empleadores. *Economica*, 69(1).

OConnell, S. D., Mation, L. F., Basto, J. B. T., and Dutz, M. A. (2019). Making public job training work: Evidence from decentralized targeting using firm input. Available at https://www.stephenoconnell.org/files/OM_MakingJobTrainingWork.pdf.

Pereda, P., Narita, R., Rocha, F., Diaz, M., Almeida, E., Borges, B. P. d. S., and Matsunaga, L. (2022). Formalization of female micro-entrepreneurs in Brazil. Available at SSRN: https://ssrn.com/abstract=4000220.

Perry, G. E., Maloney, W. F., Arias, O. S., Fajnzylber, P., Mason, A. D., and Saavedra-Chanduvi, J. (2007). *Informality: Exit and Exclusion*. World Bank, Washington, D.C.

Peña, H. E. (2010). *Impact Evaluation of a Job-Training Programme for Disadvantaged Youths: The Case of Projoven*. PhD thesis, Maastricht University, Maastricht, The Netherlands.

Picchio, M. (2020). Meta-analysis. In Zimmermann, K. F., editor, *Handbook of Labor, Human Resources and Population Economics*, pages 1–29. Springer International Publishing.

Pignatti, C. (2016). Do public employment services improve employment outcomes? Evidence from Colombia. Working Paper 10, International Labour Organization ILO, Geneva.

Pignatti, C. (2020). Compliance with Labour Legislation in Informal Labour Markets. Working paper. Available at SSRN: https://ssrn.com/abstract=4995256.

Piza, C. (2018). Out of the shadows? Revisiting the impact of the Brazilan SIMPLES program on firms formalization rates. *Journal of Development Economics*, 134:125–132.

Ponczek, V. and Ulyssea, G. (2022). Enforcement of labour regulation and the labour market effects of trade: Evidence from Brazil. *The Economic Journal*, 132(641):361–390.

Portes, A., Castells, M., and Benton, L. A., editors (1989). *The Informal Economy: Studies in Advanced and Less Developed Countries*. Johns Hopkins University Press, Baltimore, MD.

Ranis, G. and Stewart, F. (1999). V-goods and the role of the urban informal sector in development. *Economic Development and Cultural Change*, 47(2):259–288.

Reis, M. (2015). Vocational training and labor market outcomes in Brazil. *The BE Journal of Economic Analysis & Policy*, 15(1):377–405.

Rocha, R., Ulyssea, G., and Rachter, L. (2018). Do lower taxes reduce informality? Evidence from Brazil. *Journal of Development Economics*, 134:28–49.

Ronconi, L. (2010). Enforcement and compliance with labor regulations in Argentina. *ILR Review*, 63(4):719–736.

Ronconi, L. and Colina, J. (2011). Simplification of labor registration in Argentina: Achievements and pending issues. Working Paper 277, Inter-American Development Bank.

Ronconi, L., Sanguinetti, J., Fachelli, S., Casazza, V., and Franceschelli, I. (2006). Poverty and employability effects of workfare programs in Argentina. PMMA Working Paper 14, PEP.

Rosas-Shady, D. (2006). Impact evaluation of Projoven youth labor training program in Perú. Working Paper 04/06, Inter-American Development Bank.

Salazar-Xirinachs, J. M. and Chacaltana, J. (2018). *Políticas de Formalización en América Latina: Avances y Desafíos. Lima: OIT, Oficina Regional para América Latina y el Caribe*. ILO Regional Office for Latin América and the Caribbean, Lima.

Santa María, M., Olivera, M., Acosta, P., Vasquez, T., and Rodriguez, A. (2009). Evaluación de impacto programa jóvenes con futuro. Working Paper 44, Fedesarrollo–Centro de Investigación Económica y Social.

Schmid, G., Speckesser, S., and Hilbert, C. (2001). Does active labour market policy matter? An aggregate impact analysis for germany. In *Labour market policy and unemployment: An Evaluation of Active Measures in France, Germany, the Netherlands, Spain and Sweden*, pages 77–114. Edward Elgar Cheltenham.

Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, 70(1):103–127.

Stanley, T. D. and Doucouliagos, H. (2011). *Meta-Regression Analysis in Economics and Business*. Routledge.

Stanley, T. D. and Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, 5(1):60–78.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, 108(8):2015–2047.

Valdivia, M. (2015). Business training plus for female entrepreneurship? Short and medium-term experimental evidence from Perú. *Journal of Development Economics*, 113:33–51.

Van Gameren, E. (2010). Evaluación de impacto del programa de apoyo al empleo. Evaluación de impacto, Secretaría de Trabajo y Previsión Social (STPS), Ciudad de México.

Viollaz, M. (2018a). Are labour inspections effective when labour regulations vary according to the size of the firm? Evidence from Perú. *ILR Review*, 157(2):213–242.

Viollaz, M. (2018b). Enforcement of labour market regulations: heterogeneous compliance and adjustment across gender. *IZA Journal of Labour Policy*, 7(1):1–28.

Zucco, C., Lenz, A.-K., Goldszmidt, R., and Valdivia, M. (2020). Face-to-face vs. virtual assistance to entrepreneurs: Evidence from a field experiment in Brazil. *Economics Letters*, 188(108922).

Zucco, C., Lenz, A.-K., Goldszmidt, R., and Valdivia, M. (2023). Do political preferences affect policy learning and uptake? Evidence from a field experiment with informal entrepreneurs. *The Journal of Politics*, 85(3):1003–1016.

Appendices

Appendix A

Sample Selection and Construction Process

Paper-level screening first identified 45 studies that did not meet our fundamental inclusion criteria, removing 150 estimates. The largest exclusion category was social protection programmes without activation components (67 estimates), followed by studies that did not present estimations for probability of being formal (38 estimates), studies that did not evaluate formalisation policy effectiveness as a primary objective (30 estimates), and studies employing non-experimental identification methods (15 estimates). This initial screening left 112 relevant studies containing 1,810 estimates.

Estimate-level refinement applied granular selection criteria to these remaining estimates, excluding 1,302 for various methodological and definitional reasons. The largest exclusion categories reflected our analytical focus: 247 estimates related to cohort or call variations not incorporated, 247 estimates for outcome variables outside our scope, 180 estimates for employment outcomes we decided not to analyse, and 166 estimates where only combined treatments were available when disaggregated effects were required. Methodological exclusions included 110 estimates lacking standard errors, 95 estimates representing duplicate formality definitions, and various other technical exclusions totalling 267 estimates.

Data construction adjustments subsequently created 19 additional pooled gender estimates for five studies that reported separate male and female effects but lacked overall estimates. This refinement process resulted in 527 estimates from 79 studies, ensuring methodological consistency while capturing the full breadth of formalisation policy evaluations in Latin American.

Data Construction and Technical Adjustments

We made several technical adjustments during data construction to ensure consistency and avoid double-counting across studies. For studies presenting multiple treatment variations targeting the same outcome, we prioritised aggregate treatment effects when available. Nine studies reported both individual and aggregate treatment estimates (such as different types of training programmes or sensitisation mechanisms), and we consistently selected the aggregate estimate to maintain comparability while avoiding overrepresentation of individual studies. Six studies did not provide aggregate estimates for multiple treatments and therefore contribute multiple estimates per outcome category. These include studies examining different training modalities or varied sensitisation approaches within the same evaluation framework.

Table A.1: Study and estimate selection process

Selection Stage	Studies	Estimates
Initial search results	152	1,960
Paper-level exclusions:	*-45*	*-150*
Does not evaluate formalization policy effectiveness	-13	
Does not present estimations for probability of being formal	-11	
Non-experimental identification	-13	
Social protection without activation component	-8	
Remaining after paper exclusions	**107**	**1,810**
Estimate-level exclusions from relevant papers:		*-1,302*
Cohort/call not incorporated		-247
Outcome variable not considered		-247
Employment outcome not considered		-180
Only combined treatment is considered		-166
No standard errors		-110
Duplicate formality (definition not considered)		-95
Age categories not used		-57
Informal employment outcome not considered		-43
Duplicate earnings (definition not considered)		-42
Estimates not considered		-29
Only pooled definition of firm considered		-25
Matching method not selected		-18
Paper analyzed in new version		-12
Subgroup not considered		-12
Only pooled definition of worker considered		-8
Estimate with non-experimental identification		-6
Control group not considered		-3
Duplicate employment (definition not considered)		-2
Estimate-level inclusions from relevant papers:		*19*
Pooled gender estimates created		19
Final analytical sample	**79**	**527**

Notes: The table shows the complete selection process from 1,960 initial estimates to 527 final estimates. Data construction adjustments created pooled gender estimates for studies reporting only gender-disaggregated results. Paper-level exclusions removed entire studies that did not meet inclusion criteria. Estimate-level exclusions removed specific estimates within retained studies for methodological and definitional reasons. When all estimates from a study are not considered, the study is not considered.

One study, Pignatti (2016), presented estimates for four different control groups. We excluded the "employers" control group as it was not comparable with other labour intermediation programmes in our sample, retaining only the three control groups representing job seekers.

Three studies contributed estimates for both firm-level and worker-level outcomes within the same intervention, reflecting the dual nature of formalisation processes. These studies examined both employer registration and employee social security coverage, providing complementary perspectives on policy effectiveness. We retained both types of estimates as they measure distinct aspects of formalisation.

Our sample includes eleven programmes that have been evaluated by multiple independent studies, representing 27 total evaluations. These include large-scale national programmes such as Colombia's *Jóvenes en Acción* (evaluated by four separate studies) and Peru's *Projoven programme* (assessed in five different evaluations). This overlap strengthens our evidence base by providing multiple perspectives on the same interventions while allowing us to examine consistency across different evaluation approaches and time periods.

Handling Interaction Terms and Gender Pooling

We implemented two additional data construction procedures to maximise the analytical value of available estimates. First, six studies reported results using interaction terms rather than separate subgroup estimates. Following the approach recommended by Irsova et al. (2024), we calculated the total treatment effects for specific demographic groups by computing partial derivatives of the treatment effect. For gender interactions, for instance, we derived effects as $\beta_{male} = \beta_{treatment}$ and $\beta_{female} = \beta_{treatment} + \beta_{interaction}$, with standard errors adjusted using the delta method: $SE_{female} = \sqrt{SE^2_{treatment} + SE^2_{interaction}}$, assuming zero covariance between coefficients due to the typical unavailability of covariance estimates (Irsova et al., 2024; Havranek et al., 2024). This approach allowed us to derive gender-specific and age-specific effects from studies by Alzuá and Brassiolo (2006), Barrera-Osorio et al. (2023), López Mourelo and Escudero (2017), and Doerr and Novella (2024).

Second, we created pooled gender estimates for study-outcome combinations that reported separate male and female effects but lacked an overall estimate. Following Borenstein et al. (2021), we employed inverse-variance weighting to combine mutually exclusive subgroups within studies:

$$\hat{\beta}_{pooled} = \frac{w_m \hat{\beta}_m + w_f \hat{\beta}_f}{w_m + w_f} \tag{7}$$

where weights are defined as $w = 1/SE^2$, and pooled standard errors as $SE_{pooled} =$

$1/\sqrt{w_m + w_f}$. This inverse-variance approach is preferred over simple sample-size weighting as it better accounts for variance driven by study design and outcome measurement precision (Borenstein et al., 2021). This procedure generated 19 additional estimates in five studies, allowing us to include studies with gender-disaggregated results in analyses that examine the overall effectiveness of the programme.

Sample and Control Group Observations for SMD Construction

The calculation of standardised mean differences requires precise treatment and control group sample sizes for each estimate. However, most studies in our sample report regression-based results where the analytical sample may differ substantially from the initial design sample due to attrition, missing covariates, or subgroup-specific analyses. We therefore constructed treatment and control group sizes based on the actual regression samples rather than design samples to ensure accurate weighting in our meta-analytical models.

For studies reporting overall treatment and control group distributions, we applied these proportions to subgroup analyses when demographic breakdowns were not explicitly provided. When studies presented descriptive statistics showing the demographic composition of treatment and control groups separately (e.g., 40% of the treatment group and 45% of the control group were women), we maintained these differential distributions in our sample size calculations. In cases where only general demographic proportions were reported (e.g., 70% of the total sample were women), we assumed equal representation between treatment and control groups while acknowledging that this can introduce measurement error.

This approach was particularly relevant for quasi-experimental studies, where the concept of distinct treatment and control groups is less straightforward than in randomised trials. In difference-in-differences, regression discontinuity, and instrumental variable designs, we constructed treatment and control samples based on the analytical framework of each study, typically defining the control group as observations contributing to the baseline comparison in the regression specification. While this methodology introduces some imprecision, particularly for subgroup analyses, we considered it preferable to using aggregate sample sizes that would not reflect the actual precision of demographic-specific estimates.

Our approach ensures that the weights assigned to each estimate in our meta-analysis correspond to the statistical precision of the underlying regression rather than potentially misleading design sample sizes, thereby improving the accuracy of our standardised mean difference calculations and subsequent meta-analytical results.

Appendix B

This appendix provides supplementary analyses referenced in Section 4. We present outcome-specific publication bias diagnostics, a preliminary sign-and-significance analysis, descriptive heterogeneity patterns, and robustness checks for the meta-regression results.

Outcome Variable Definitions

Table B.1 presents the complete list of formality, earnings, and hours worked definitions used across all studies in our sample, along with programme types analysed. This demonstrates the heterogeneity in outcome measurement that motivates our use of standardised mean differences.

Table B.1: Outcome Definitions Across Studies

Author (Year)	Programme	Formality	Hours	Earnings
Abel et al. (2022)	HS	(E) Social security registration	–	Monthly - formal, level
Aedo and Pizarro (2004)	T + A	(E) Work contract	–	Monthly, level
Alcázar and Huerta (2023)	T	(E) Formal salaried worker	–	Monthly, level
Alzuá and Brassiolo (2006)	T + A	(E) Employer pays social security	–	Monthly, level
Alzúa et al. (2016)	T + A	(E) SIPA registration	–	Monthly - formal, level
Amarante et al. (2009)	SP	(E) Social security access	Weekly, level	Monthly, log
Aparicio (2014)*	S	(F) Government registration	–	–
Attanasio et al. (2011)*	T + A	(E) Health, pension, injury benefits	Weekly, level	Monthly - formal, level
Attanasio et al. (2017)	T + A	(E) PILA registration	–	Monthly - formal, level
Barrera-Osorio et al. (2023)	T	(E) Formal contract	Weekly, level	Monthly, level
Barron (2020)	T	(F) Tax ID number	–	–
Bernal et al. (2017)	PTR	–	–	Monthly average wage, log (firm level)
Berniell and de la Mata (2017)	HS	(E) SIPA registration	–	Monthly - formal, level
Blanchard et al. (2025)*	SP	(E) Employer has paid licence	Weekly, level	–
Bruhn (2013)	S	(F) Government registration / (E) Signed contract	–	–
Cabrera et al. (2016)	SE	(F) Legal work permit	–	–
Calderón-Mejía and Marinescu (2012)*	S	(E) Health and pension contributions	–	–
Calero et al. (2017)	T	(E) Official work card	Weekly, level	Monthly, level
Camargo et al. (2021)	T	(E) Official work card	–	Monthly, level

Continued on next page

Table B.1 – continued from previous page

Author (Year)	Programme	Formality	Hours	Earnings
Card et al. (2011)	T + A	(E) Employer health insurance	Weekly, level	Monthly, level
Carvalho et al. (2021)	T	(E) Formal sector worker	–	–
Chaparro et al. (2020)	T + LI	(E) PILA registration	–	–
Corseuil et al. (2019)	T + A	(E) RAIS registration	–	–
Costa (2022)	T	(E) RAIS registration	–	–
Da Mata et al. (2025)	T + A	(E) RAIS registration	–	Monthly - formal, level
De Andrade et al. (2016)	SE/LIN	(F) Government registration	–	–
de Barros et al. (2023)	T	(E) RAIS registration	–	–
Delajara et al. (2013)	T + LI	(E) IMSS registration	–	Monthly - formal, level
Departamento Nacional de Planeación (2008)*	T + A	(E) Formal sector employment	Weekly, level	Monthly, log
Díaz et al. (2018)*	PTR + S	(F) Municipal licenses / (E) Electronic Payroll Registry	–	–
Díaz and Jaramillo (2006)	T + A	(E) Formal contract/insurance/social security	Weekly, level	Monthly, level
Díaz and Rosas (2016)*	T + A	(E) Retirement pension	Weekly, log	Monthly, log
Doerr and Novella (2024)	T	(E) Social security/unemployment insurance	Hours worked	Monthly, level
Durand (2018)	A	(E) Healthcare access	–	Monthly, log
Escudero (2018)	SP	(E) Pension system affiliation	Weekly, log	–
Fajnzylber et al. (2011)	PTR + S	(F) Tax registration	–	–
Feld (2024)	SE + LIN	(E) Pension contributions	Weekly, log	Monthly, log
Galdo and Chong (2012)	T + A	(E) Formal contract	–	Monthly, level

Continued on next page

Table B.1 – continued from previous page

Author (Year)	Programme	Formality	Hours	Earnings
Galiani et al. (2017)	SE + PTR	(F) License	–	–
Garsous et al. (2015)	PTR	–	–	Monthly average wage, log (municipal)
Gasparini et al. (2009)	SP	(E) SIJP registration	–	–
Gómez and González-Velosa (2023)	T + LI	(E) Social security contributions	–	Monthly - formal, level
Hernani-Limarino and Villarroel (2015)	T + A	(E) Social protection compliance	–	Monthly, level
Ibarrarán et al. (2019)*	T + A	(E) Employer health insurance	–	Monthly, log
Ibarrarán et al. (2014)*	T + A	(E) Employer health insurance	–	Monthly, log
Kugler et al. (2017)*	PTR	(E) Health and pension contribution	–	–
Kugler et al. (2022)	T + A	(E) Health/pension coverage	–	–
Lenz (2017)	SE	(F) MEI registration	–	–
Llerena Pinto et al. (2013)	A	(E) Formal sector employment	–	–
López Mourelo and Escudero (2017)	T + LI	(E) Formal sector employment	Weekly, level	–
Macroconsult (2012)	SP	(E) Legal entity/contract	Weekly, level	Monthly, level
Madalozzo and Bortoluzzo (2011)	PTR	(F) Registered employment	–	–
Mera et al. (2023)	T	(F) Tax authority registration	–	–
Monteiro and Assunção (2012)	PTR + S	(F) Official license	–	–
Morales and Medina (2017)	PTR	–	–	Monthly average wage (firm level)
Morales-González et al. (2019)	LI	(E) Social security benefits	–	–
Nogueira Puentes (2018)	SP	(E) BPS contributions	–	–
Novella and Valencia (2022)	HS + LI	(E) Social security contributions	–	Monthly, log

Continued on next page

Table B.1 – continued from previous page

Author (Year)	Programme	Formality	Hours	Earnings
Novella et al. (2017)	T	(E) Unemployment insurance registration	–	Monthly, log
OConnell et al. (2019)	T	(E) RAIS registration	–	Monthly, level
Ohaco (2023)	SE	(F) Domestic worker registration	–	–
Peña (2010)	T + A	(E) Youth training contract	–	Monthly, log
Pereda et al. (2022)	PTR + S	(F) MEI registration	–	–
Pignatti (2020)*	LIN	(E) Health social security	Weekly, level	Monthly, level
Pignatti (2016)	LI	(E) Social security coverage	–	–
Piza (2018)	PTR + S	(F) Municipal/state license	–	–
Reis (2015)	T	(E) Formal sector employment	–	Monthly, level
Rocha et al. (2018)	PTR + S	(F) RAIS registration	–	–
Ronconi and Colina (2011)	S	(E) Pension contributions	–	–
Ronconi et al. (2006)	SP	(E) Social security coverage	–	Monthly, level
Rosas-Shady (2006)	T + A	(E) Formal sector employment	Weekly, level	Monthly, level
Santa María et al. (2009)	T + A	(E) Social security/pension affiliation	–	Monthly, log
Ponczek and Ulyssea (2022)	LIN	–	–	Monthly average, level (micro-region)
Valdivia (2015)	T	(F) Tax registration	–	–
Van Gameren (2010)	LI/T/A	(E) Labour benefits	–	Monthly, level
Viollaz (2018a,b)*	LIN	(E) Employment benefits	–	–

Zucco et al. (2020, 2023)	SE	(F) MEI formalization	–	–

Notes: (E) = Employee formality; (F) = Firm formality. Programme type abbreviations: HS = Hiring subsidies; T = Training; A = Apprenticeship or Internship; SP = Social protection; S = Simplification; SE = Sensitization; PTR = Payroll tax reduction; LI = Labour intermediation; LIN = Labour inspections. Dashes (–) indicate outcome not measured in the study. Preferred estimate was selected based on priority in studies with multiple definitions, marked with a * following - 1) registration with authorities, 2) employment benefits, 3) pension benefits, 4) social security benefits, 5) signed contract.

Outcome-Specific Publication Bias Analysis

We complement the pooled publication bias analysis in Section 4.3 with outcome-specific assessments. Table B.2 presents Egger test results for each outcome category. All outcomes show evidence of publication bias, with particularly strong evidence for monthly earnings (coefficient: 11.237, $p < 0.01$).

Table B.2: Summary of Egger Test Results by Outcome

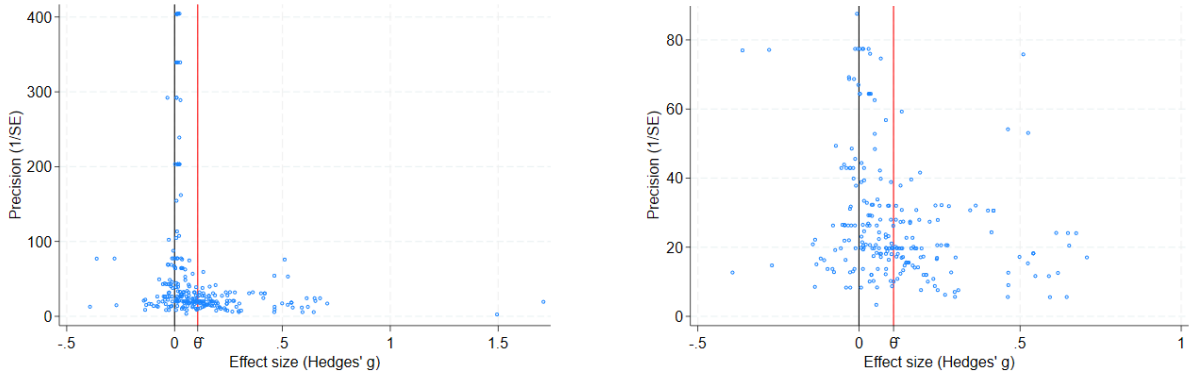
Outcome	Test	Coefficient	Interpretation
All Outcomes	RE	4.921 ***	Pub. Bias
		(0.413)	
	FE	2.412 ***	Pub. Bias
		(0.055)	
Formality	RE	1.766 ***	Pub. Bias
		(0.349)	
	FE	2.332 ***	Pub. Bias
		(0.071)	
Hours Worked	RE	1.396 **	Pub. Bias
		(0.576)	
	FE	1.651 ***	Pub. Bias
		(0.397)	
Monthly Earnings	RE	11.237 ***	Pub. Bias
		(1.109)	
	FE	3.186 ***	Pub. Bias
		(0.100)	

Notes: Egger's test is a linear regression of the intervention effect estimates on their standard errors weighted by their inverse variance. RE refers to Random Effects, FE refers to Fixed Effects. Pub. Bias refers to Publication Bias. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figures B.1 to B.3 display funnel plots for each outcome category. Visual inspection confirms the asymmetry patterns identified in the pooled analysis, with studies concentrated in regions indicating positive effects with low precision across all outcome types.

Tables B.3 to B.5 present detailed FAT-PET-PEESE bias correction estimates for each outcome category. Bias correction substantially reduces effect sizes across all outcomes, with the largest reductions for earnings outcomes. For formality outcomes, the PEESE-corrected estimate is 0.087 ($p < 0.01$). Hours worked shows a PEESE-corrected estimate of 0.022, though this does not reach statistical significance. Monthly earnings displays the strongest bias, with PEESE correction yielding -0.097, suggesting publication selection may entirely account for observed positive effects in this outcome category.

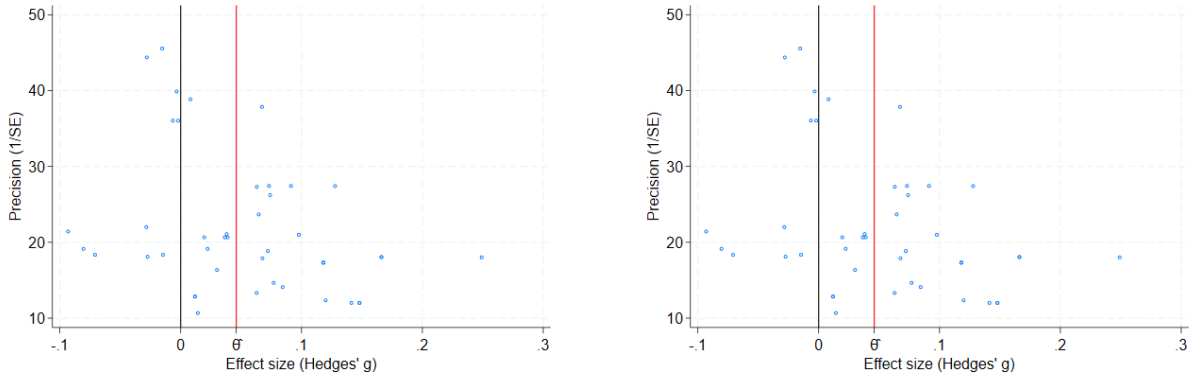
Figure B.1: Funnel plots for formality outcomes



(a) With pooled estimate (b) Outliers excluded

Notes: θ (red dashed lines) show pooled random-effects estimates. N=269.

Figure B.2: Funnel plots for hours worked outcomes



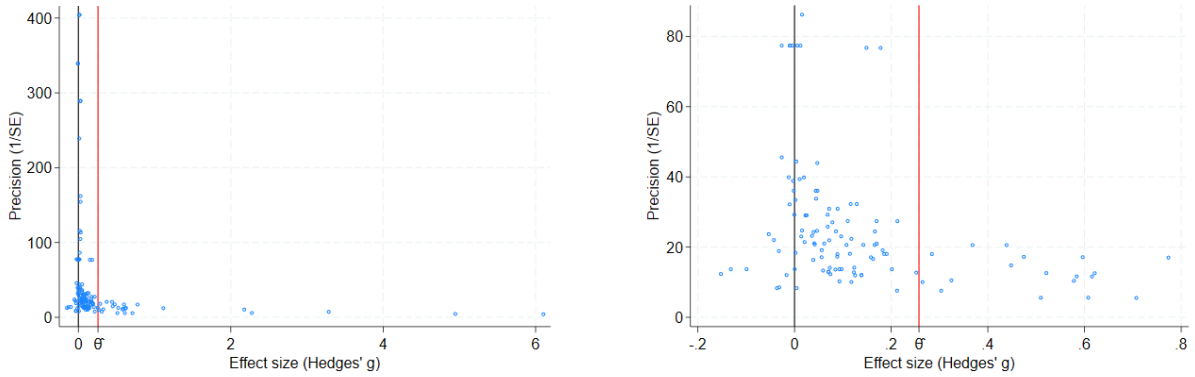
(a) With pooled estimate (b) Outliers excluded

Notes: θ (red dashed lines) show pooled random-effects estimates. N=44.

Sign and Significance Analysis

Before presenting our main standardised mean difference meta-analysis in Section 4, we conducted a preliminary sign and significance analysis following Escudero et al. (2019), Card et al. (2018), and Jessen and Kluve (2021). This complementary method examines patterns in the direction and statistical significance of treatment effects without requiring complete sample size information. The approach requires only the reported sign and statistical significance of each estimate, making it applicable to the full sample of 527 estimates. This analysis provides a descriptive overview of patterns in the literature but has important limitations: it discards information about effect magnitudes, treats all statistically significant positive effects equally regardless of size, and does not correct for publication bias. These limitations motivate our primary focus on standardised mean difference meta-analysis with explicit bias correction in the main text.

Figure B.3: Funnel plots for monthly earnings outcomes



(a) With pooled estimate

(b) Outliers excluded

Notes: θ (red dashed lines) show pooled random-effects estimates. N=131.

For each estimate, we extracted information on effect direction (positive or negative) and statistical significance based on conventional thresholds (typically $p < 0.05$ or $p < 0.10$). Our sample presents 287 (54.5%) statistically significant positive, 213 (40.4%) statistically insignificant, and 27 (5.1%) statistically significant negative effects. Given the low proportion of significant negative effects, we follow Escudero et al. (2019) in creating a binary outcome variable coded as 1 if the estimate is statistically significant positive and 0 otherwise.

Table B.6 presents results from linear probability models examining the relationship between programme effectiveness (binary indicator for positive and significant effects) and study characteristics. Explanatory variables are added sequentially across eight specifications, with standard errors clustered at the study level. Models 7 to 8 include country fixed effects.

The analysis reveals that most study characteristics are not significantly correlated with the likelihood of obtaining positive significant results. Three patterns emerge with statistical support. First, youth-targeted programmes show robust positive associations in specifications with country fixed effects (Models 7 to 8), with coefficients of 0.176 ($p < 0.05$) and 0.181 ($p < 0.05$). However, youth coefficients are small and statistically insignificant in specifications without country controls, indicating the relationship is sensitive to model specification. Second, experimental evaluation designs show consistently large negative associations across Models 6 to 8, with coefficients ranging from -0.301 to -0.349 (all $p < 0.01$). This likely reflects stricter methodological standards in experimental studies rather than worse policy performance. Third, information campaigns show positive coefficients in Models 6 to 8, ranging from 0.505 to 0.543 (all $p < 0.10$).

Most other factors show no significant relationships. Programme type differences between ALMPs, social protection, payroll tax reductions, and simplification measures

Table B.3: FAT-PET-PEESE Analysis: Formality Outcomes

Model	Constant (PET)	SE Coeff (FAT)	Interpretation
WLS (Stanley 2005)	0.006	2.332***	Pub. Bias
	(0.003)	(0.389)	
OLS Clustered	0.018	2.024***	Pub. Bias
	(0.035)	(0.597)	
Fixed Effects	-0.058	3.584***	Pub. Bias
	(0.046)	(0.945)	
IV (Sample Size)	0.032	1.741**	Pub. Bias
	(0.034)	(0.752)	
Weighted (1/SE)	0.007	2.234***	Pub. Bias
	(0.009)	(0.483)	
PEESE Analysis			
PEESE Clustered	0.087	6.871***	
	(0.025)	(1.493)	

Notes: N=444 from 59 studies. PET (Precision Effect Test) provides bias-corrected effects through the constant term. FAT (Funnel Asymmetry Test) tests the SE coefficient for publication bias. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

are statistically insignificant across specifications. Formality and earnings outcomes show positive but insignificant coefficients relative to hours worked. Both effect duration (short versus medium versus long-term measurement timing) and programme duration show inconsistent patterns with no statistical significance. Women-focused estimates and gender-pooled estimates show small negative coefficients that rarely reach significance. GDP growth shows a positive but insignificant association (0.014).

The inclusion of country fixed effects in Models 7 to 8 substantially increases the ^{2} from 0.130 to 0.223 to 0.230, indicating that country-level factors explain considerable variation in results beyond observable characteristics.

Table B.7 presents results using study-level weights where each study receives equal weight regardless of the number of estimates it contributes. Results are qualitatively similar to the main specification.

This preliminary analysis identifies experimental designs as systematically reporting more conservative findings and shows that country-level factors explain substantial variation. However, the approach discards information about effect magnitudes and does not correct for publication bias. These limitations motivate our main standardised mean difference meta-analysis presented in Section 4.

Table B.4: FAT-PET-PEESE Analysis: Hours Worked Outcomes

Model	Constant (PET)	SE Coeff (FAT)	Interpretation
WLS (Stanley 2005)	-0.031	1.651***	Pub. Bias
	(0.023)	(0.546)	
OLS Clustered	-0.011	1.232*	No bias
	(0.041)	(0.723)	
Fixed Effects	0.010	0.836	No bias
	(0.056)	(1.062)	
IV (Sample Size)	-0.012	1.255*	No bias
	(0.036)	(0.700)	
Weighted (1/SE)	-0.021	1.417**	Pub. Bias
	(0.032)	(0.595)	
PEESE Analysis			
PEESE Clustered	0.022	10.178	
	(0.030)	(7.378)	

Notes: N=444 from 59 studies. PET (Precision Effect Test) provides bias-corrected effects through the constant term. FAT (Funnel Asymmetry Test) tests the SE coefficient for publication bias. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Heterogeneity by Programme Type and Implementation Features

The high heterogeneity in the pooled analysis ($I^2 = 99.9\%$) indicates substantial variation in treatment effects across studies. We examine effect sizes across key characteristics to understand sources of this variation. The results in this section are simple weighted averages within each subgroup. They are not meta-regression estimates, which appear in Section 4.4 and control for multiple characteristics simultaneously while correcting for publication bias.

Programme Type Heterogeneity Figure B.4 presents effect sizes for each programme category with sufficient observations. ALMPs show the most consistent effectiveness. With 366 observations, random effects meta-analysis shows an effect size of 0.112 ($p < 0.01$). Between-study heterogeneity is high ($I^2 = 99.6\%$), but the estimate is statistically significant, indicating robust positive effects across diverse implementation contexts. All three estimation methods converge around similar point estimates.

Social protection programmes with activation components show the largest point estimates (RE: 0.394, $p < 0.05$). Precision is low, with wide confidence intervals reflecting substantial variation across seven contributing studies. Heterogeneity is high ($I^2 = 99.98\%$). The fixed effects estimate approaches zero, underscoring sensitivity to assumptions about between-study variation.

Information campaigns show modest but consistent effects (RE: 0.015, $p < 0.01$).

Table B.5: FAT-PET-PEESE Analysis: Monthly Earnings Outcomes

Model	Constant (PET)	SE Coeff (FAT)	Interpretation
WLS (Stanley 2005)	-0.002	3.186 ***	Pub. Bias
	(0.005)	(0.482)	
OLS Clustered	-0.410	11.850 *	No bias
	(0.263)	(6.315)	
Fixed Effects	-0.634	15.776 **	Pub. Bias
	(0.327)	(5.726)	
IV (Sample Size)	-0.025	5.114 **	Pub. Bias
	(0.027)	(2.606)	
Weighted (1/SE)	-0.051	5.564 *	No bias
	(0.047)	(2.920)	
PEESE Analysis			
PEESE Clustered	-0.097	70.313 ***	
	(0.051)	(24.673)	

Notes: N=444 from 59 studies. PET (Precision Effect Test) provides bias-corrected effects through the constant term. FAT (Funnel Asymmetry Test) tests the SE coefficient for publication bias. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Heterogeneity is relatively low ($I^2 = 33.2\%$), suggesting information-based approaches produce small but predictable improvements across contexts. All three estimation methods cluster tightly together.

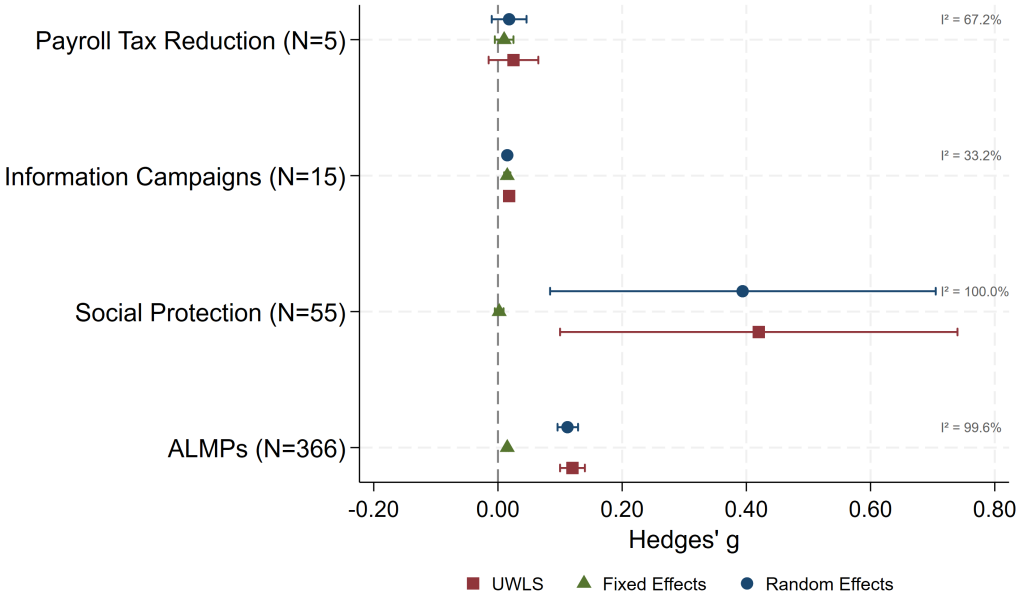
Evidence on payroll tax reductions remains limited. With only five observations from four studies, the random effects estimate (0.018) does not reach statistical significance. The single observation for labour inspections and the absence of simplification estimates reflect evidence gaps.

Demographic Targeting Figure B.5 summarises results across participant characteristics. Male participants show larger random effects estimates (0.180, $p < 0.01$) compared to females (0.110, $p < 0.01$), though confidence intervals overlap substantially. Both groups display I^2 values above 99%, indicating high heterogeneity. The pooled gender estimate (0.143, $p < 0.01$) falls between the gender-specific estimates.

Youth-focused interventions using strict ILO definitions (<30 years) show effect sizes of 0.122 ($p < 0.01$). Broader criteria (≤35 years) show similar effects of 0.118 ($p < 0.01$). Confidence intervals are narrow, and consistency across both youth definitions suggests robust effects for younger workers. Programmes targeting general populations show larger point estimates (0.169, $p < 0.01$) but substantially wider confidence intervals, reflecting greater variation in effects.

Implementation Features Figure B.6 examines implementation dimensions. The upper panel shows effect measurement timing (follow-up period after programme comple-

Figure B.4: Meta-Analytic Effect Sizes by Programme Type



Notes: Meta-analytic effect sizes with 95% confidence intervals by programme type, comparing three estimation methods: unrestricted weighted least squares (UWLS), fixed effects (FE), and random effects (RE). Effect sizes represent standardised mean differences (Hedges' g). Sample sizes shown in parentheses on y-axis. I^2 statistics indicate between-study heterogeneity, shown for random effects estimates. Labour inspections (N=1) and simplification measures (N=0) excluded due to insufficient observations.

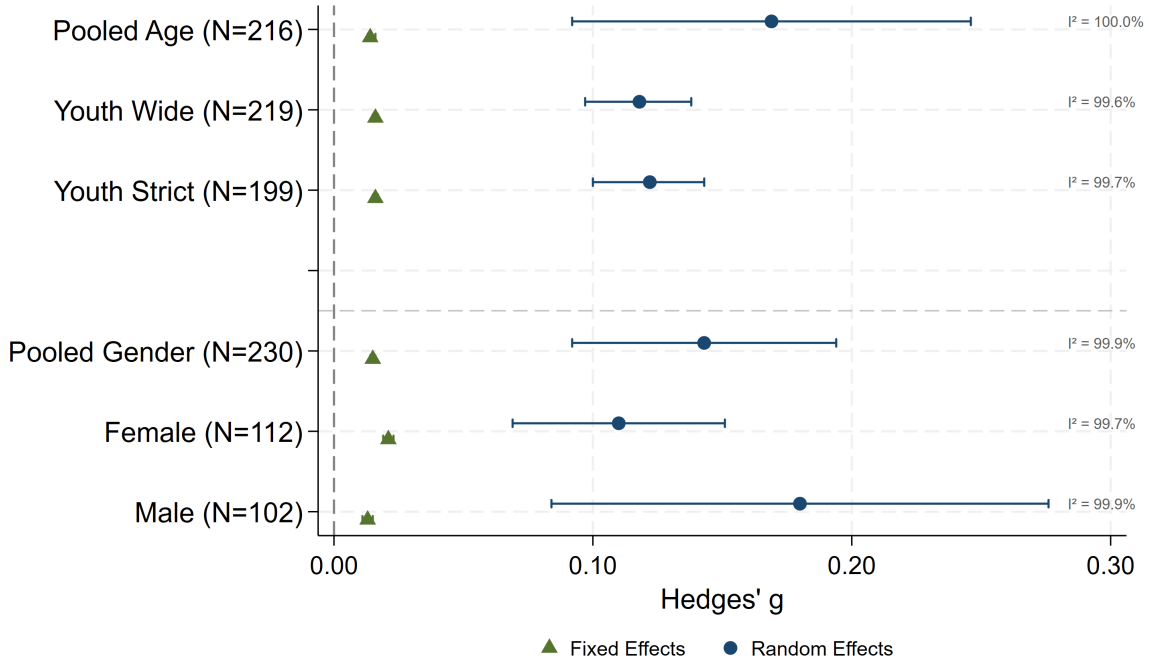
tion). Short-term measurement ($\leq$12 months after completion) shows the largest random effects estimates (0.196, $p < 0.01$). Medium-term measurement follows (0.110, $p < 0.01$). Long-term measurement shows the smallest impacts (0.079, $p < 0.01$). Confidence intervals do not overlap, suggesting these differences are meaningful.

This pattern requires careful interpretation. It does not mean short programmes are more effective than long programmes. Rather, it reflects when outcomes are measured. Short-term measurements may capture immediate compliance responses or programme participation effects that fade over time. Long-term measurements may reflect more sustainable but smaller behavioural changes. Alternatively, different types of interventions may be measured at different time horizons. The meta-regression analysis in Section 4.4 addresses this by controlling for programme type, duration, and other characteristics simultaneously, revealing that effects actually strengthen over time when other factors are held constant.

The lower panel compares programme complexity. Multi-component programmes show larger random effects estimates (0.173, $p < 0.01$) than single-component interventions (0.101, $p < 0.01$), suggesting potential complementarities between policy tools. Both categories display high heterogeneity ($I^2 > 99\%$). The difference between multi-component and single-component approaches is statistically significant.

Fixed effects and random effects estimates differ substantially across all categories, indicating considerable between-study heterogeneity. Despite this, point estimates are

Figure B.5: Meta-Analytic Effect Sizes by Target Population



Notes: Meta-analytic effect sizes with 95% confidence intervals by demographic targeting, comparing fixed effects (FE) and random effects (RE) methods. Effect sizes represent standardised mean differences (Hedges' g). Sample sizes shown in parentheses on y-axis. I^2 statistics indicate between-study heterogeneity for random effects estimates. Upper panel shows gender-specific results; lower panel shows age-targeted results. Youth Strict uses ILO definition (<30 years); Youth Wide includes participants ≤35 years.

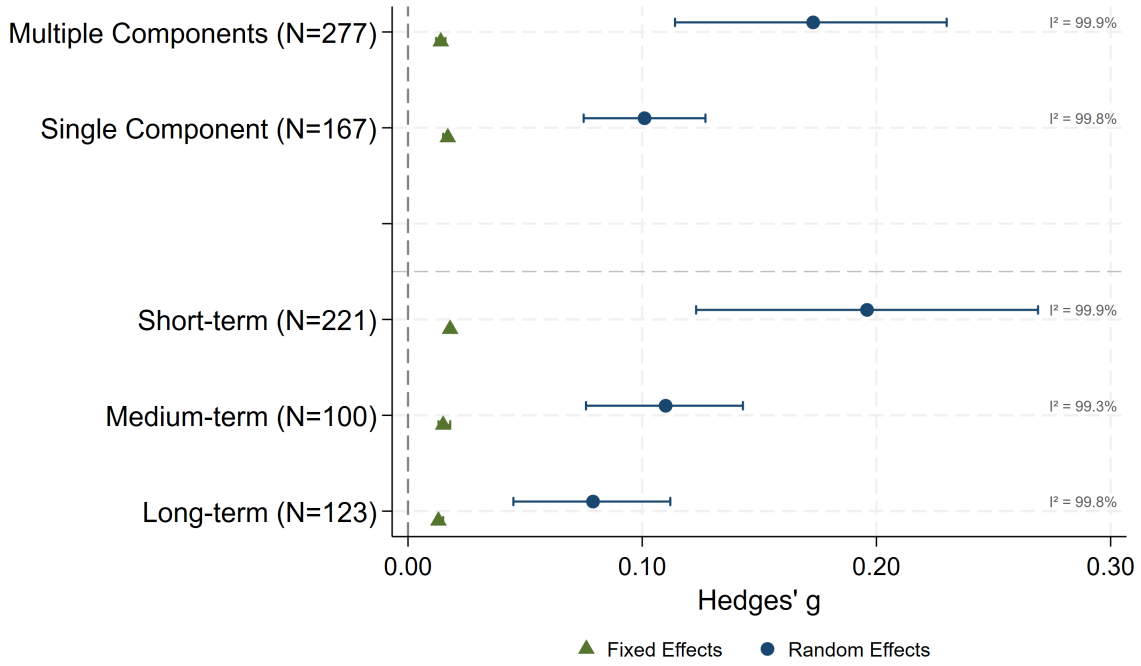
consistently positive across subgroups, and statistical significance holds in most cases. This suggests formalisation policies produce genuine positive effects, though effect magnitudes vary considerably across contexts.

Robustness Checks and Alternative Specifications

Table B.8 presents meta-regression results using PEESE publication bias correction with variance rather than standard error terms. The specification shows qualitatively similar patterns across programme types compared to the main results in Table 5, though with generally smaller effect magnitudes. The relative ranking of programme effectiveness remains consistent. Social protection programmes show the largest effects over the baseline category across both correction methods.

Tables B.9 to B.11 provide detailed meta-regression results for each outcome category separately. These confirm the heterogeneous impact patterns identified in the pooled model presented in Section 4.4.

Figure B.6: Effect Measurement Timing and Programme Complexity



Notes: Meta-analytic effect sizes with 95% confidence intervals by implementation features, comparing fixed effects (FE) and random effects (RE) methods. Effect sizes represent standardised mean differences (Hedges' g). Sample sizes shown in parentheses on y-axis. I^2 statistics indicate between-study heterogeneity for random effects estimates. Upper panel shows effect measurement timing (follow-up period): Short-term ($\leq$12 months after programme completion), Medium-term (12 to 24 months), Long-term (>24 months). Lower panel shows programme complexity: Single Component programmes implement one intervention type; Multiple Components combine different types.

Table B.6: Meta-regression results: Baseline models for total sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(i) Programme type (base: payroll + simplification)								
ALMPs	0.178	0.203	0.237	0.186	0.123	0.251	0.090	0.032
	(0.245)	(0.249)	(0.249)	(0.255)	(0.256)	(0.244)	(0.262)	(0.279)
Social protection	−0.145	−0.126	−0.081	−0.084	−0.191	−0.168	−0.247	−0.270
	(0.270)	(0.275)	(0.276)	(0.276)	(0.283)	(0.273)	(0.295)	(0.298)
Labour inspections	0.055	0.091	0.134	0.140	0.081	0.055	−0.210	−0.253
	(0.259)	(0.269)	(0.272)	(0.270)	(0.281)	(0.273)	(0.301)	(0.317)
Sensitization	0.287	0.283	0.283	0.284	0.209	0.543*	0.525*	0.505*
	(0.267)	(0.268)	(0.269)	(0.266)	(0.268)	(0.282)	(0.286)	(0.301)
Payroll tax	0.292	0.296	0.314	0.311	0.292	0.329	0.058	−0.033
	(0.267)	(0.266)	(0.266)	(0.262)	(0.266)	(0.257)	(0.281)	(0.291)
Information campaign	0.200	0.221	0.225	0.231	0.227	0.189	−0.201	−0.280
	(0.368)	(0.381)	(0.383)	(0.383)	(0.380)	(0.361)	(0.362)	(0.366)
(ii) Effect duration (base: short-term)								
Medium-term		0.016	0.029	0.019	0.004	−0.042	−0.070	−0.054
		(0.095)	(0.095)	(0.094)	(0.095)	(0.091)	(0.090)	(0.087)
Short-term		0.065	0.073	0.080	0.062	−0.058	−0.156*	−0.136
		(0.096)	(0.096)	(0.094)	(0.095)	(0.099)	(0.088)	(0.082)
(iii) Outcome category (base: hours worked, other outcomes)								
Formality outcome			0.097	0.114	0.154	0.126	0.136	0.151
			(0.111)	(0.109)	(0.108)	(0.101)	(0.099)	(0.096)
Earnings outcome			0.025	0.042	0.064	0.063	0.078	0.078
			(0.098)	(0.099)	(0.098)	(0.094)	(0.091)	(0.087)
(iv) Target group (base: men, older workers, general population)								
Female				−0.072	−0.075	−0.077	−0.080	−0.079
				(0.093)	(0.093)	(0.094)	(0.094)	(0.097)
Gender pooled				−0.064	−0.075	−0.094*	−0.063	−0.059
				(0.060)	(0.059)	(0.056)	(0.057)	(0.058)
Youth(18-24)				0.087	0.061	0.100	0.176**	0.181**
				(0.105)	(0.100)	(0.090)	(0.069)	(0.072)
(v) Programme duration (base: long duration)								
Short duration					−0.133	0.014	0.153	0.127
					(0.120)	(0.098)	(0.112)	(0.118)
Medium duration					0.001	0.106	0.089	0.034
					(0.118)	(0.103)	(0.112)	(0.102)
(vi) Evaluation design (base: non-experimental)								
Experimental design						−0.301***	−0.349***	−0.314***
						(0.095)	(0.091)	(0.090)
(vii) Contextual factors								
GDP growth								0.014
								(0.011)
Constant	0.400*	0.340	0.235	0.271	0.389	0.400	0.664**	0.637**
	(0.238)	(0.257)	(0.286)	(0.292)	(0.314)	(0.287)	(0.309)	(0.317)
Country effects	*No*	*No*	*No*	*No*	*No*	*No*	*Yes*	*Yes*
R^2	0.046	0.049	0.055	0.064	0.078	0.130	0.223	0.230
Observations	527	527	527	527	527	527	527	518

Notes: Dependent variable is binary indicator for positive and significant effect. Standard errors in parentheses clustered at study level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Meta-regression results: Baseline models for total sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(i) Programme type (base: payroll + simplification)								
ALMPs	0.205	0.204	0.222	0.254	0.217	0.280	0.114	0.084
	(0.211)	(0.215)	(0.219)	(0.223)	(0.226)	(0.217)	(0.230)	(0.244)
Social protection	−0.050	−0.052	−0.022	−0.046	−0.111	−0.097	−0.176	−0.177
	(0.245)	(0.245)	(0.250)	(0.251)	(0.258)	(0.247)	(0.292)	(0.287)
Labour inspections	−0.020	−0.025	−0.001	−0.010	−0.048	−0.051	−0.245	−0.256
	(0.238)	(0.249)	(0.256)	(0.253)	(0.260)	(0.257)	(0.263)	(0.270)
Sensitization	0.346	0.347	0.347	0.349	0.317	0.515*	0.581**	0.573**
	(0.238)	(0.237)	(0.237)	(0.235)	(0.239)	(0.266)	(0.275)	(0.282)
Payroll tax	0.226	0.225	0.240	0.239	0.227	0.249	0.022	−0.007
	(0.268)	(0.267)	(0.269)	(0.266)	(0.270)	(0.266)	(0.259)	(0.265)
Simplification	0.386	0.383	0.383	0.386	0.385	0.379	0.073	0.036
	(0.300)	(0.312)	(0.312)	(0.309)	(0.309)	(0.304)	(0.281)	(0.288)
(ii) Effect duration (base: short-term)								
Medium-term		−0.006	−0.007	−0.004	−0.007	−0.006	0.003	0.008
		(0.122)	(0.123)	(0.120)	(0.118)	(0.117)	(0.107)	(0.105)
Short-term		−0.006	−0.008	−0.017	−0.020	−0.049	−0.148	−0.141
		(0.103)	(0.104)	(0.104)	(0.103)	(0.110)	(0.100)	(0.098)
(iii) Outcome category (base: hours worked, other outcomes)								
Formality outcome			0.119	0.108	0.139	0.128	0.162	0.175
			(0.124)	(0.122)	(0.118)	(0.115)	(0.111)	(0.107)
Earnings outcome			0.086	0.078	0.099	0.097	0.126	0.126
			(0.122)	(0.120)	(0.119)	(0.115)	(0.108)	(0.102)
(iv) Target group (base: men, older workers, general population)								
Female				−0.053	−0.052	−0.048	−0.039	−0.031
				(0.092)	(0.092)	(0.093)	(0.095)	(0.101)
Gender pooled				−0.070	−0.076	−0.097	−0.077	−0.079
				(0.072)	(0.073)	(0.074)	(0.071)	(0.074)
Youth(18-24)				−0.090	−0.110	−0.073	0.024	0.024
				(0.099)	(0.098)	(0.101)	(0.097)	(0.099)
(v) Programme duration (base: long duration)								
Short duration					−0.079	−0.020	0.004	−0.023
					(0.107)	(0.104)	(0.104)	(0.103)
Medium duration					0.031	0.107	−0.002	−0.046
					(0.115)	(0.117)	(0.122)	(0.108)
(vi) Evaluation design (base: non-experimental)								
Experimental design						−0.197*	−0.247**	−0.213**
						(0.102)	(0.099)	(0.099)
(vii) Contextual factors								
GDP growth								0.010
								(0.011)
Constant	0.364*	0.370	0.253	0.335	0.392	0.386	0.636**	0.607**
	(0.205)	(0.230)	(0.271)	(0.282)	(0.299)	(0.282)	(0.296)	(0.303)
Country effects	*No*	*No*	*No*	*No*	*No*	*No*	*Yes*	*Yes*
R^2	0.057	0.057	0.061	0.067	0.075	0.096	0.195	0.192
Observations	527	527	527	527	527	527	527	518

Notes: Dependent variable is binary indicator for positive and significant effect. Standard errors in parentheses clustered at study level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.8: Meta-regression results: PEESE publication bias correction

	All Outcomes		Formality		Earnings		Hours Worked	
	RE	FE	RE	FE	RE	FE	RE	FE
(i) Programme type (base: payroll + simplification)								
ALMPs	0.183	0.124***	0.208*	0.132***	0.169	0.041**	−0.140	−0.140
	(0.188)	(0.028)	(0.109)	(0.028)	(0.244)	(0.016)	(0.205)	(0.205)
Social protection	0.598***	0.119***	0.013	−0.035	0.981***	0.488***		
	(0.194)	(0.030)	(0.118)	(0.032)	(0.274)	(0.031)		
Labour inspections	0.324	0.155***	0.275	0.153***				
	(0.320)	(0.052)	(0.184)	(0.052)				
Sensitization	0.113	0.000	0.216*	0.047				
	(0.217)	(0.042)	(0.128)	(0.042)				
Payroll tax	0.090	0.046	0.159	0.046				
	(0.218)	(0.028)	(0.138)	(0.030)				
(ii) Effect duration (base: long-term)								
Medium-term	−0.051	0.002	−0.039	0.005***	−0.160**	−0.059***	−0.017	−0.017
	(0.038)	(0.002)	(0.029)	(0.002)	(0.080)	(0.010)	(0.033)	(0.033)
Short-term	0.086**	0.007***	−0.050*	0.008***	−0.037	−0.003	0.009	0.009
	(0.041)	(0.002)	(0.030)	(0.002)	(0.101)	(0.010)	(0.039)	(0.039)
(iii) Outcome category (base: hours worked)								
Formality outcome	0.081*	0.045***						
	(0.048)	(0.007)						
Earnings outcome	0.125**	0.040***						
	(0.050)	(0.007)						
(iv) Target group (base: men, pooled samples)								
Female	0.001	0.008***	0.001	0.011***	−0.050	−0.003	0.005	0.005
	(0.038)	(0.002)	(0.028)	(0.002)	(0.079)	(0.003)	(0.025)	(0.025)
Gender pooled	0.038	0.006***	0.005	0.008***	0.108	0.000	0.001	0.001
	(0.035)	(0.001)	(0.026)	(0.002)	(0.073)	(0.003)	(0.027)	(0.027)
Youth(18-24)	0.009	−0.018***	0.014	−0.037***	0.109	0.142***	0.004	0.004
	(0.036)	(0.004)	(0.025)	(0.005)	(0.084)	(0.013)	(0.049)	(0.049)
(v) Programme duration (base: long duration)								
Short duration	−0.167***	−0.056***	−0.065*	−0.047***	−0.205**	−0.100***		
	(0.050)	(0.007)	(0.039)	(0.009)	(0.102)	(0.012)		
Medium duration	−0.263***	−0.036***	−0.021	0.005	−0.542***	−0.178***	−0.123	−0.123
	(0.048)	(0.007)	(0.040)	(0.010)	(0.100)	(0.015)	(0.143)	(0.143)
(vi) Evaluation design (base: non-experimental)								
Experimental design	−0.039	−0.070***	−0.085***	−0.080***	−0.194*	−0.114***	−0.013	−0.013
	(0.041)	(0.004)	(0.031)	(0.005)	(0.112)	(0.011)	(0.080)	(0.080)
(vii) Contextual factors								
GDP growth	0.020***	0.007***	0.004	0.008***	0.026**	0.011***	0.012	0.012
	(0.004)	(0.000)	(0.003)	(0.001)	(0.011)	(0.001)	(0.012)	(0.012)
Variance(PEESE)	29.695***	31.878***	9.408***	20.549***	60.499***	50.586***	4.498	4.497
	(2.019)	(0.910)	(1.960)	(1.203)	(4.304)	(1.586)	(13.432)	(13.432)
Constant	−0.181	−0.038	−0.044	0.011	−0.070	0.012	0.205	0.205
	(0.206)	(0.030)	(0.122)	(0.030)	(0.254)	(0.021)	(0.228)	(0.228)
Pub. bias correction	*PEESE*	*PEESE*	*PEESE*	*PEESE*	*PEESE*	*PEESE*	*PEESE*	*PEESE*
R^2								
Observations	392	392	233	233	119	119	40	40

Notes: Meta-regression with PEESE publication bias correction using variance. RE = Random Effects, FE = Fixed Effects. Full specification includes all covariates and country fixed effects. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.9: Meta-regression results: Formality outcomes (Random Effects with Publication Bias Correction)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(i) Programme type (base: payroll + simplification)							
ALMPs	0.242*	0.273**	0.294**	0.321**	0.312**	0.240*	0.238**
	(0.134)	(0.134)	(0.134)	(0.129)	(0.131)	(0.125)	(0.109)
Social protection	0.183	0.215	0.241*	0.230*	0.218	0.026	0.033
	(0.139)	(0.138)	(0.138)	(0.133)	(0.135)	(0.134)	(0.117)
Labour inspections	0.206	0.207	0.207	0.326	0.302	0.388*	0.320*
	(0.230)	(0.228)	(0.227)	(0.220)	(0.223)	(0.211)	(0.183)
Sensitization	0.192	0.192	0.196	0.314**	0.299**	0.306**	0.226*
	(0.143)	(0.142)	(0.141)	(0.139)	(0.142)	(0.135)	(0.127)
Payroll tax	0.190	0.212	0.212	0.282*	0.265	0.221	0.186
	(0.172)	(0.171)	(0.171)	(0.165)	(0.167)	(0.157)	(0.137)
(vii) Publication bias correction							
SE - Publication Bias	1.759***	1.871***	1.949***	1.952***	1.958***	1.858***	1.974***
	(0.348)	(0.351)	(0.355)	(0.345)	(0.360)	(0.352)	(0.397)
(ii) Effect duration (base: long-term)							
Medium-term		−0.011	−0.015	−0.041	−0.041	−0.072**	−0.046
		(0.034)	(0.034)	(0.033)	(0.033)	(0.033)	(0.029)
Short-term		0.054**	0.050*	0.002	−0.002	−0.112***	−0.054*
		(0.027)	(0.028)	(0.029)	(0.029)	(0.032)	(0.029)
(iii) Target group (base: men, pooled samples)							
Female			0.003	−0.000	0.001	−0.004	0.007
			(0.034)	(0.032)	(0.032)	(0.030)	(0.028)
Gender pooled			0.052*	0.037	0.036	0.020	0.011
			(0.030)	(0.029)	(0.029)	(0.028)	(0.026)
Youth(18-24)			0.006	0.028	0.034	0.022	0.007
			(0.027)	(0.026)	(0.027)	(0.027)	(0.025)
(iv) Evaluation design (base: non-experimental)							
Experimental design				−0.120***	−0.125***	−0.127***	−0.090***
				(0.027)	(0.030)	(0.031)	(0.030)
(v) Programme duration (base: long duration)							
Short duration					0.018	−0.001	−0.021
					(0.037)	(0.044)	(0.041)
Medium duration					0.047	−0.016	−0.010
					(0.036)	(0.042)	(0.039)
(vi) Contextual factors							
GDP growth							0.004
							(0.003)
Constant	−0.201	−0.260*	−0.312**	−0.249*	−0.262*	−0.099	−0.142
	(0.135)	(0.137)	(0.140)	(0.135)	(0.141)	(0.140)	(0.125)
Meta-analysis method	*RE*	*RE*	*RE*	*RE*	*RE*	*RE*	*RE*
Country effects	*No*	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Publication bias correction	*SE*	*SE*	*SE*	*SE*	*SE*	*SE*	*SE*
R^2							
Observations	269	269	269	269	269	269	233

Notes: Sample restricted to formality outcomes. Dependent variable is Hedges' g (standardised mean difference). Random effects meta-regression with publication bias correction using standard error. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.10: Meta-regression results: Earnings outcomes(Random Effects with Publication Bias Correction)

	(1)	(2)	(3)	(4)	(5)	(6)
(i) Programme type (base: payroll + simplification)						
ALMPs	−0.232	−0.179	0.003	0.095	0.281	0.352
	(0.347)	(0.336)	(0.337)	(0.328)	(0.294)	(0.313)
Social protection	0.501	0.481	0.575	0.782**	1.365***	1.258***
	(0.370)	(0.360)	(0.356)	(0.352)	(0.315)	(0.347)
(vii) Publication bias correction						
SE - Publication Bias	9.902***	10.775***	11.619***	11.278***	10.408***	11.588***
	(1.046)	(1.054)	(1.085)	(1.094)	(1.094)	(1.156)
(ii) Effect duration (base: long-term)						
Medium-term		−0.286**	−0.319**	−0.334***	−0.317***	−0.317***
		(0.124)	(0.125)	(0.120)	(0.105)	(0.105)
Short-term		0.047	−0.054	−0.096	−0.206*	−0.194
		(0.109)	(0.117)	(0.114)	(0.123)	(0.130)
(iii) Target group (base: men, pooled samples)						
Female			−0.053	−0.063	−0.028	−0.058
			(0.119)	(0.113)	(0.094)	(0.099)
Gender pooled			0.155	0.174*	0.115	0.131
			(0.106)	(0.102)	(0.088)	(0.093)
Youth(18-24)			−0.005	−0.021	0.180**	0.042
			(0.093)	(0.094)	(0.088)	(0.108)
(iv) Evaluation design (base: non-experimental)						
Experimental design			−0.212**	−0.179*	−0.382***	−0.324**
			(0.099)	(0.107)	(0.129)	(0.144)
(v) Programme duration (base: long duration)						
Short duration				−0.045	−0.215*	−0.129
				(0.119)	(0.120)	(0.132)
Medium duration				−0.349***	−0.615***	−0.656***
				(0.106)	(0.120)	(0.126)
(vi) Contextual factors						
GDP growth						0.026*
						(0.014)
Constant	−0.163	−0.202	−0.321	−0.268	−0.130	−0.356
	(0.343)	(0.335)	(0.346)	(0.351)	(0.314)	(0.335)
Meta-analysis method	*RE*	*RE*	*RE*	*RE*	*RE*	*RE*
Country effects	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Publication bias correction	*SE*	*SE*	*SE*	*SE*	*SE*	*SE*
R^2						
Observations	131	131	131	131	131	119

Notes: Sample restricted to earnings outcomes. Dependent variable is Hedges' g (standardised mean difference). Random effects meta-regression with publication bias correction using standard error. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.11: Meta-regression results: Hours worked (Random Effects with Publication Bias Correction))

	(1)	(2)	(3)	(4)	(5)	(6)
(i) Programme type (base: payroll + simplification)						
ALMPs	0.021	0.023	0.080***	0.100*	−0.087	−0.203
	(0.023)	(0.023)	(0.029)	(0.055)	(0.171)	(0.207)
(vii) Publication bias correction						
SE - Publication Bias	1.117*	1.079	1.598***	0.807	−0.183	−0.314
	(0.654)	(0.657)	(0.601)	(0.744)	(1.323)	(1.398)
(ii) Effect duration (base: long-term)						
Medium-term		−0.030	−0.026	−0.030	0.006	−0.010
		(0.027)	(0.022)	(0.022)	(0.029)	(0.033)
Short-term		−0.026	−0.037*	−0.050**	0.019	0.017
		(0.024)	(0.019)	(0.021)	(0.035)	(0.040)
(iii) Target group (base: men, pooled samples)						
Female			0.008	0.003	−0.000	−0.002
			(0.023)	(0.023)	(0.025)	(0.026)
Gender pooled			0.007	−0.000	−0.008	−0.010
			(0.023)	(0.023)	(0.029)	(0.030)
Youth(18-24)			−0.016	−0.002	0.006	0.012
			(0.026)	(0.030)	(0.045)	(0.048)
(iv) Evaluation design (base: non-experimental)						
Experimental design			−0.103***	−0.112***	−0.030	0.014
			(0.022)	(0.030)	(0.065)	(0.082)
(v) Programme duration (base: long duration)						
Short duration				0.099*	0.037	
				(0.058)	(0.085)	
Medium duration				0.019	−0.099	−0.173
				(0.035)	(0.124)	(0.147)
(vi) Contextual factors						
GDP growth						0.013
						(0.013)
Constant	−0.019	0.001	−0.016	0.002	0.219	0.306
	(0.028)	(0.033)	(0.037)	(0.048)	(0.236)	(0.261)
Meta-analysis method	*RE*	*RE*	*RE*	*RE*	*RE*	*RE*
Country effects	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Publication bias correction	*SE*	*SE*	*SE*	*SE*	*SE*	*SE*
R^2						
Observations	44	44	44	44	44	40

Notes: Sample restricted to hours worked outcomes. Dependent variable is Hedges' g (standardised mean difference). Random effects meta-regression with publication bias correction using standard error. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Conclusion

This thesis has examined three questions in applied labour economics: how artificial intelligence adoption affects employer demand for personality traits, why job advertisements mentioning personality traits are associated with lower wages, and how effective formalisation policies have been in Latin America. Although the chapters differ in their focus and methodology, they share a commitment to using careful empirical methods and novel data sources to address policy-relevant questions.

Summary of Findings

Chapter 1 shows that personality traits are highly demanded in UK job advertisements, appearing in 72 per cent of vacancies. Employers' demand for Openness increases with AI adoption, both within specific AI-related positions and across firms' broader hiring following their first use of AI technologies. This finding suggests that as AI-skills expands, employers increasingly value traits associated with creativity and adaptability.

Chapter 2 documents and explains the wage penalty paradox. Jobs mentioning personality traits offer lower wages on average, but this relationship reflects multiple mechanisms rather than a single cause. Occupation characteristics explain part of the penalty, but significant effects persist after controlling for observable job features. Compensating differentials operate when traits are bundled with attractive amenities. Linguistic intensity matters: low-emphasis trait mentions carry minimal penalties while high-emphasis mentions show substantial penalties. These patterns indicate that employers use personality language strategically for different purposes in different contexts.

Chapter 3 finds that formalisation policies in Latin America generally produce positive effects, though publication bias has led to overestimation of impact sizes in the literature. After correction, active labour market policies show the most consistent effects on formality, while social protection programmes with activation components primarily affect earnings rather than employment status. Programme benefits strengthen over time, and multi-component interventions outperform single-component approaches.

Implications for Research and Policy

These findings have several implications. For researchers studying skill demand and technological change, the results from Chapters 1 and 2 suggest that personality traits deserve greater attention alongside cognitive and technical skills. Models of human capital formation should incorporate non-cognitive attributes as distinct factors that interact with technology in specific ways. The strong complementarity between AI and Openness indicates that AI may affect personality requirements as systematically as it affects skill

demands.

For educators and workforce development professionals, the evidence that personality traits appear in 72 per cent of job advertisements, far exceeding formal education requirements, suggests that traditional academic qualifications may be insufficient preparation for the labour market. Development of creative thinking and adaptability may be particularly valuable as AI adoption continues to expand.

For policymakers in developing countries seeking to reduce informality, the meta-analysis in Chapter 3 offers several lessons. First, expectations should be realistic: even well-designed programmes produce modest effects, and publication bias has inflated estimates in the existing literature. Second, sustained intervention matters: effects strengthen over time, and programmes of longer duration produce larger impacts. Third, comprehensive approaches work better than single interventions: multi-component programmes that address multiple barriers to formalisation outperform narrowly targeted policies. Fourth, different policies serve different purposes: active labour market policies are most effective for achieving formality outcomes, while social protection programmes primarily affect income.

Limitations and Future Directions

Several limitations affect the interpretation of these results. Chapters 1 and 2 rely on online job advertisement data, which captures posted wages rather than realised compensation and may not represent all labour market activity. Chapter 3 synthesises existing evaluations, each with its own methodological limitations, and cannot fully overcome the constraints of the underlying studies.

Future research could extend this work in several directions. Linking job advertisement data with administrative records on employment outcomes would allow examination of whether trait requirements predict worker-job match quality and wage trajectories. Experimental studies could test how job seekers respond to different intensities and combinations of personality language in advertisements. For formalisation policies, longer-term follow-up and more systematic evaluation of multi-component programmes would strengthen the evidence base.

Despite these limitations, this thesis contributes new evidence on important questions in labour economics. Understanding how employers value personality traits, how AI-skills affects their demand, and how policy can promote formal employment are essential for designing education systems, workforce programmes, and labour market institutions that serve workers in a changing economy.