

THREE ESSAYS ON RETAIL GLOBALISATION AND INTERNATIONAL TRADE

Kyung In Hwang

A thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

September 2018

Abstract

Using the empirical context of Korea's retail sector, this thesis brings together three empirical essays that examine the significance of the retail sector in the global economy.

The first essay examines how retail market powers influence the impact of tariff cuts on consumer prices in the context of Korea's tariff reduction since the mid-1980s. Using consumption-item-level consumer price index (CPI) and product-level tariff data with a new matching strategy, we empirically find that tariff cuts incompletely pass through to consumer prices. Guided by a simple theoretical framework based on an oligopolistic competition model with a nested CES demand, this essay considers markups charged by retailers as a new channel by which consumer prices do not fully react to tariff cuts. To identify the retail markup channel, we firstly find that tariff pass-through to consumer prices are deteriorated in regions with greater retail market concentrations. Further, this essay measures retail markups proxied by price-cost margins, and quantifies the degree to which retail markups exacerbate tariff pass-through using a new instrumental variable strategy. The instrument variable estimation reveals that a 10 percentage point increase in retail markups contributes to 1.73-1.96 percent lower tariff pass-through. Our results are robust with estimation based on limited four-year census data by which we account for store-level variations and geographical market heterogeneities in retail markup measure. We also estimate the retail markup channel augmented by spatial autoregressive (SAR) regression model to control for CPI's spatially cross-section dependence, and the results are robust.

The second essay aims to answer the question about whether strategic interactions affect FDI (Foreign Direct Investment)'s market entry. From the literature, the answer is less clear as host-country markets are generally considered as being perfectly competitive. Under the imperfectly competitive market, however, even multinational enterprises cannot act in a vacuum. In this essay, we study the strategic aspect of FDI's entry into host-country markets. To study it, we use an empirical context of the Korean discounting retail industry where, as the first and the second biggest firms, a domestic retailer (Emart) competes with a multinational retailer (Tesco) by opening multiple stores across local markets. We empirically show that retail FDI enters host-country markets strategically by reduced-form estimations. Further, we build an entry game-theoretical model in which the strategic interaction is fully incorporated, and estimate the structural parameters of the model. The existence of strategic entry is clearly confirmed by structural estimation results. We also find that asymmetric information between

foreign and domestic retailers play a key role in forming FDI's strategic market entry.

Lastly, the third essay studies whether multinational enterprises affect domestic firms' markups. Since Korea liberalised retail FDI in 1996, foreign stores have proliferated over the Korean market whereas average markup of local stores has declined to a large degree. To assess the causal impact of retail FDI on domestic markups, we use the inverse propensity-score weighting regression (IPW) model combined with difference-in-difference (DD) approach. In this essay, we find that multinational retailers have a causal impact on reduction in average local markups. Through the lens of IPW and DD, we also find that the decrease in domestic markups is thoroughly driven by pro-competitive effects through the entry of foreign retailers, but productivity spillover effects induced by multinational retailers are limited. This essay also reveals that Korean large discount stores, of which scales, functions and product assortments are very similar to multinational stores, have no impact on average markups, indicating that only FDI plays a role in decline in markups.

Acknowledgements

I would like to express my warmest gratitude to Professor Sourafel Girma who encouraged and directed me. His exceptional guidance, patience and challege brought my thesis towards a completion. I am also deeply indebted to Dr. Zhihong Yu for his great passion for teaching and research. I have greatly benefited from conversations and discussions with him. I would also like to thank all my PhD colleagues, with whom I have shared moments of deep anxiety but also of big excitement. I must express my gratitude to my mother who has never lost in faith in me. Her unceasing prayer has enabled me to reach today's position. I want to thank my children, Hyein and Yuha, for inspiring me every day and being my best supporters. My thesis is completed and I am coming home. I would like to thank my wife, Ji Young Lee, for continuously and patiently supporting me in every phase of the work on this thesis and not letting me give up and giving me all the encouragement I needed to continue. My thesis has not been completed without her. Finally, I thank my God. Even here, Your hand guided me and was holding me fast.

Contents

Abstract **2**

Acknowledgements **4**

1 Introduction **10**

1.1 Context and Issues 10

1.2 Overview 15

1.3 Thesis Outline 17

2 Tariffs, Retailers, and Consumer Prices **19**

2.1 Introduction 19

2.2 Data 24

2.2.1 Trade Policy Background 24

2.2.2 Price and Tariff Data 25

2.2.3 Matching CPI with Tariffs 27

2.2.4 Data Summary 29

2.3 Tariffs and Consumer Prices 31

2.3.1 Regression Specification 31

2.3.2 The Impact of Tariffs on Consumer Prices 33

2.4 Retail Markup Channel 36

2.4.1 A Simple Theoretical Framework 36

2.4.2 Retail Market Concentrations and Tariff Pass-Through . . . 39

2.4.3 The Role of Retail Markups in Tariff Pass-Through 41

2.5 Accounting for Heterogeneities and Cross-Section Dependence . . . 47

2.5.1 Market Heterogeneities in Retail Markups 47

2.5.2 Cross-Section Dependence in Consumer Prices 50

2.6 Conclusion 51

3 Foreign Retail Entry Under Domestic Rivals **71**

3.1 Introduction 71

3.2 Background 75

3.2.1 Large Discount Store Market in Korea 75

3.2.2 Illustrative Analyses on Retail Spatial Competition 77
3.3 Data 78
3.4 Reduced-Form Evidence 81
3.4.1 Baseline Regression Specification 81
3.4.2 Baseline Results 82
3.4.3 Endogeneity and IV Estimation 83
3.4.4 IV Results 85
3.5 Structural Estimation 87
3.5.1 An Empirical Model for Market Entry 87
3.5.2 Structural Estimation 90
3.5.3 Incomplete Information versus Complete Information 93
3.6 Conclusion 97

4 Do Multinationals Reduce Domestic Markups? 109
4.1 Introduction 109
4.2 Background 112
4.3 Data and Preliminary analysis 114
4.4 Identification 117
4.4.1 Empirical Strategy 117
4.4.2 Diff-in-Diff Inverse Propensity-Score Weighting estimator . . 118
4.4.3 Constructing Pseudo-Randomised Sample 121
4.5 Results 123
4.5.1 Treatment Effects on Average Domestic Markups 123
4.5.2 Do Productivity Spillovers Influence Treatment Effects? . . 126
4.5.3 Market Definition and Treatment homogeneity 127
4.5.4 Korean Large Discount Stores 129
4.6 Conclusion 130

5 Conclusion 143

List of Figures

1.1 The importance of the retail sector in Korea's economy (2013) . . . 11
1.2 The Evolutions of Total Retail Sales . . . 12
1.3 The Evolution of Total Number of Retailers . . . 14

2.1 Matching procedure between CPI and Tariffs . . . 63
2.2 Time Trend in Log Consumer Price Index . . . 64
2.3 Time trend in Average Tariff Rates . . . 65
2.4 Tariff Dispersions across Sectors . . . 66
2.5 Association Between Retail Markups and Wholesale Herfindahl Index . . . 67
2.6 Large Discount Stores in Korea (2010) . . . 68

3.1 Spatial Expansions of Large Discount Stores . . . 105
3.2 Spatial Expansions of Large Discount Stores in Seoul . . . 106
3.3 Spatial Expansions of Large Discount Stores in Kyunggi . . . 107
3.4 Tesco's Market Entry by Emart's Pre-occupation . . . 108

4.1 Price-Cost Margin and Retail FDI over Years . . . 141
4.2 Average PCM: County with FDI vs. County without FDI . . . 141
4.3 An Example of Market Definitions and Treatment . . . 142

List of Tables

2.1 Average Agricultural Tariff Rates 52
2.2 Target Average Tariff Rates 52
2.3 Descriptive Statistics 52
2.4 Sector-level Log Consumer Price Indices over Years 53
2.5 Province-level Log Consumer Price Indices over Years 54
2.6 Tariff Rates and Consumer Price Index 55
2.7 Assessment of Bias Due to Unobserved Confounding Effects 56
2.8 Log(1+Tariff) and CPI 57
2.9 Retail Market Concentration and Tariff Pass-Through 58
2.10 Ten-Year Lagged Populations and Consumer Prices 59
2.11 Retail Markup Channel 60
2.12 Retail Markup Channel, *Analysed by Census Data* 61
2.13 Retail Markup Channel, *Spatial Autoregressive (SAR) Model* 62

3.1 Total Sales in Korean Large Discount Stores from 1988 to 2008 . . 98
3.2 Distribution of Counties by Tesco's and Emart's Entry 98
3.3 Baseline Results: Linear Probability Model 99
3.4 Baseline Results: Binary Choice (Probit) Model 100
3.5 IV Results: Linear Probability Model 101
3.6 IV Results: Binary Choice (Probit) Model 102
3.7 Structural Parameter Estimates: Incomplete Information Game . . 103
3.8 Structural Parameter Estimates: Complete Information Game . . . 104

4.1 Numbers of Stores by Retail Multinationals in Korea 132
4.2 Association between Price-Cost Margin and Retail FDI 132
4.3 Probit Estimation Results 133
4.4 Balancing Test 134
4.5 Baseline Treatment Definition 134
4.6 Baseline DDIPW Estimation Results 135
4.7 Treatment Definition in Shorter Time Horizon: 2002-2004 135
4.8 Treatment Definition in Shorter Time Horizon: 2006-2009 135
4.9 DDIPW Results with Different Periods of Treatments 136

4.10 Treatment Effects of Retail FDI on Productivities 137
4.11 Treatment Effects of Retail FDI on Markups: County-level Analyses 138
4.12 Town-level Analyses Excluding Direct Treated Towns 139
4.13 Korean Large Discount Stores(KLDS) and Average Market Markups 140

Chapter 1

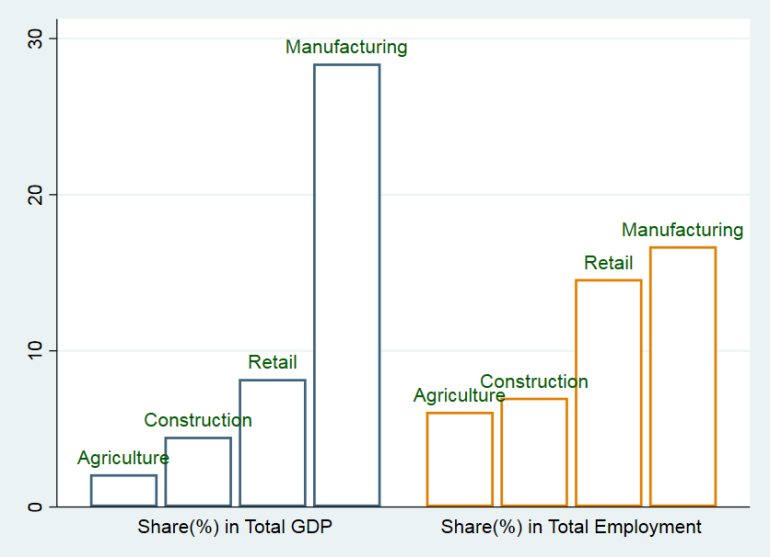
Introduction

1.1 Context and Issues

Standard international trade theories and international macro models typically assume that producers sell goods directly to final consumer markets. In the real world, however, a very large portion of the seller-buyer transactions are conducted through retail distribution networks. The retail intermediation is, thus, likely to act as a trade friction and would substantially affect welfare gains of trade. Despite its making practical sense, international economics literature has little paid attention to the retail sector. In addition, it is essential to shed a new light on retail firms in the international economics research, because multinational retail enterprises have entered into many countries actively over the past two decades, and play a crucial role in transforming the landscape of the retail sector in host-county markets. Although the retail firms take a key position in the global service FDI which accounts for a half value of global service greenfield investments, studies on the impact of retail FDI on host countries are scanty. In this thesis, we investigate the economic implications of the retail sector in a globalised world, focusing on the context of the Republic of Korea.

In Korea, the retail sector offers a substantial contribution to the national economy. As Figure 1.1 shows, the sector is the second largest sector in Korea's economy. The share of the retail sector in Korea's total GDP is 8.2% in 2013. While it is about one third of manufacturing sector(28.4%), GDP share of the sector exceeds both the construction(4.5%) and the agricultural(2.1%) sectors in Korea. With regard to employments, the retail sector is far more important in Korea's economy. Figure 1.1 also presents that the retail sector accounts for 14.6% of Korea's total employments. Although it is fewer than the manufacturing employments, the differential in employment shares between the manufacturing and the retail sector is only 2.1% points. Still, the retail sector hires even more people than the construction sector(7.0%) as well as the agriculture sector(6.1%).

Figure 1.1: The importance of the retail sector in Korea's economy (2013)



Sources: Bank of Korea, National Account; KOSTAT

Notes: This figure shows the importance of the retail sector relative to other sectors. The left four bars illustrate sectoral GDP shares(%), whereas the four bars on the right hand demonstrate sectoral employment shares(%). All the values are based on statistics in 2013.

It was not until 1996 that Korea's retail sector turned to being a part of the global market. Foreign direct investment in Korea's retail market was liberalised in 1996[1], and multinational retail firms have entered Korea. As a matter of fact, the retail sector prior to opening the market had been one of the least developed sectors in Korea. Before 1996, mom-and-pop stores and small local supermarkets accounted for a majority of Korean retail sale (Cho et al., 2015). According to the Korean national office of statistics (KOSTAT), the average number of workers per retail store, which indexes the size of retail stores, is 1.9 - 2.1 before 1996, and it is remarkably lower than other countries.[2] After 1996, however, the Korean retail sector began to experience dramatic changes. There are three noticeable structural changes in the Korean retail sector.

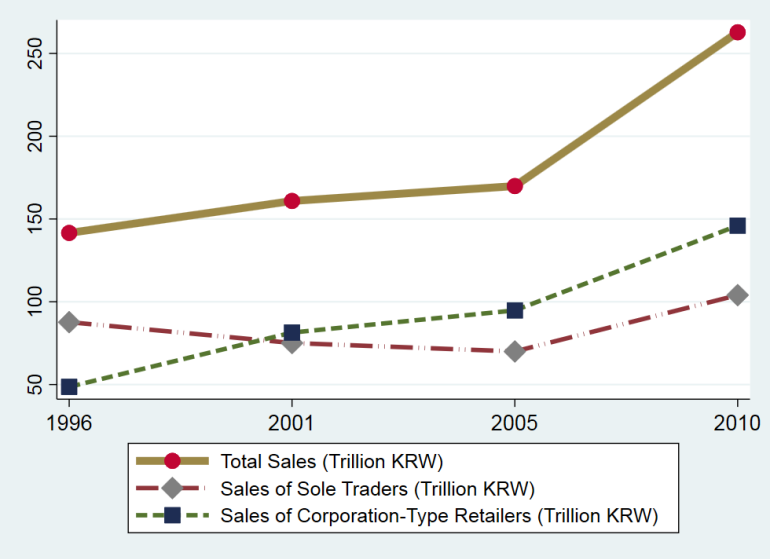
The first structural change appeared in Korea's retail sector is that total sectoral sales become almost doubled from 1996 to 2010, and that the expansion was mainly driven by large-scaled retail firms. Figure 1.2 illustrates the evolutions of total retail sales from 1996 to 2010. As shown by panel (a), total retail sales were 141.6 trillion Korean won (KRW) but it increases to 262.8 trillion KRW.[3]

[1]The retail FDI liberalisation process is explained in Chapter 4(Section 4.2)

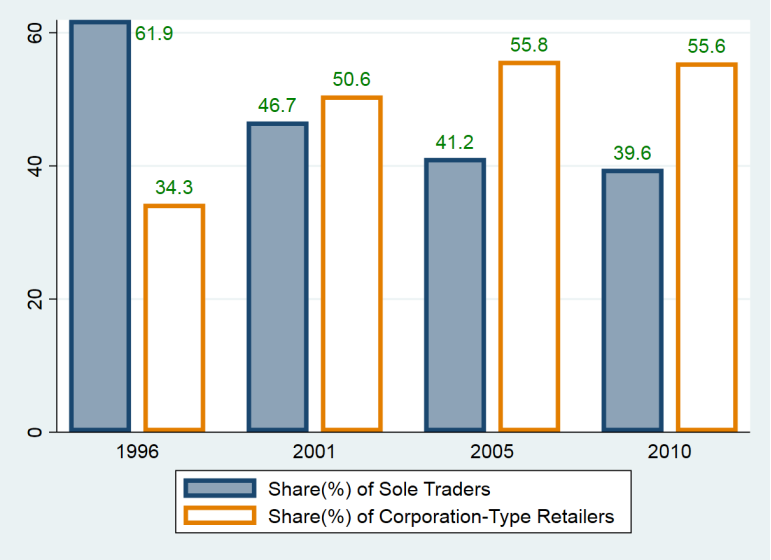
[2]For instance, the average worker per retail establishments in Japan already reached to 3.9 in 1985 (Kim and Kim, 2003).

[3]The sales amounts in Panel (a) of Figure 1.2 are all expressed in real values deflated by Korea's CPI(consumer price index).

Figure 1.2: The Evolutions of Total Retail Sales



(a) Time Trends in Total Sales



(b) Time Trends in Sales Shares by Business Type

Notes: Panel (a) demonstrates the evolutions of total sales. I further show trends in total sales of sole traders (i.e., self- and family-owned retailers) as well as those of corporation-type retailers. Panel (b) presents sales shares of both sole traders and corporation-type retailers over years. Sales amounts in Panel (a) are expressed in trillion Korean won (KRW) and are real terms deflated by Korea's consumer price index.

Panel (a), however, demonstrates that time trends in total sales differ between corporation-type retail firms and sole traders(i.e., self- and family-owned retailers). Specifically, while corporation-type retailers show an almost identical trend with total sectoral sales, sole traders representing small-scaled retail firms have a decreasing trend during the period of 1996 to 2005. In terms of market shares, retail corporations show an increasing trend as depicted by Panel (b), whereas sole traders move in the opposite direction. In particular, market shares of retail corporations began to surpass those of sole traders from 2001.

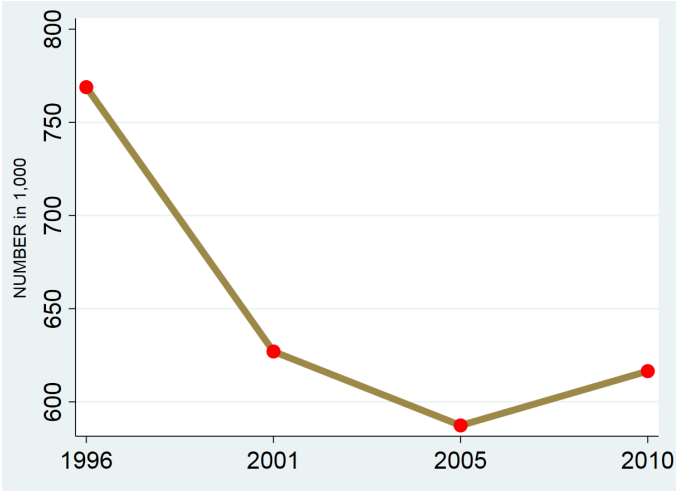
Second, although total sales have expanded since 1996, the number of retailers has shown a dramatic plunge over the period. In Figure 1.3, I respectively plot total numbers of entire retailers (Panel (a)), corporation-type retailers (Panel (b)) and sole traders (Panel (c)) across years. Panel (a) of Figure 1.3 presents that the number of retail firms was 768,875 in 1996 whereas it falls to 579,356 in 2010. The massive reduction is driven by sole traders as illustrated in Panel (c). On the other hand, the number of corporation-type retailers even soars from 13,828 in 1996 to 32,974 in 2010.

The third salient feature is the increasing prevalence of large discount stores. Large discount stores are a retail store format firstly introduced in the mid of 1990s in Korea. A typical store belonging the format has an area of 40,000 square metres of floor spaces, sells 50,000-100,000 product items and hires 100-150 employees (Cho et al., 2015). Large discount stores are analogous to hypermarket stores or superstores that sells general merchandises at low prices. Since multinational retailers have mainly operated stores of this format in the global market, abolishing the limits on foreign retail entry in 1996 contributed to the proliferation of the stores established by multinationals. At the same time, being forced to face a new circumstance, Korean conglomerates (so-called 'chaebol') that have run their businesses in the consumer-goods sector or department stores actively introduced and increased the new store format. As a result, there were only 17 large discount stores in 1996, but the total number of large discount stores increases to 420 in 2016.

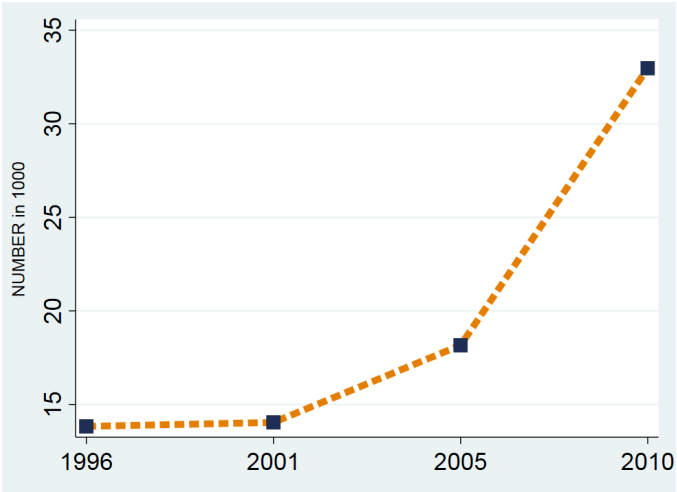
The aforementioned three features, which have characterised the Korean retail sector since 1996, are indicative of economic consequences of retail globalisation in two folds. First, retail FDI may induce increase in competition in host-country retail markets. Figure 1.2 suggests that Korea's FDI liberalisation contribute to reallocating market shares from small retailers towards large-scaled retailers. Figure 1.3 also shows that there has been a large degree of firm churning in the retail sector since 1996. The structural changes indicates that fiercer competition induced by retail FDI liberalisation transformed the landscape of the retail sector.

Second, retail globalisation is associated with changes in market structure. We

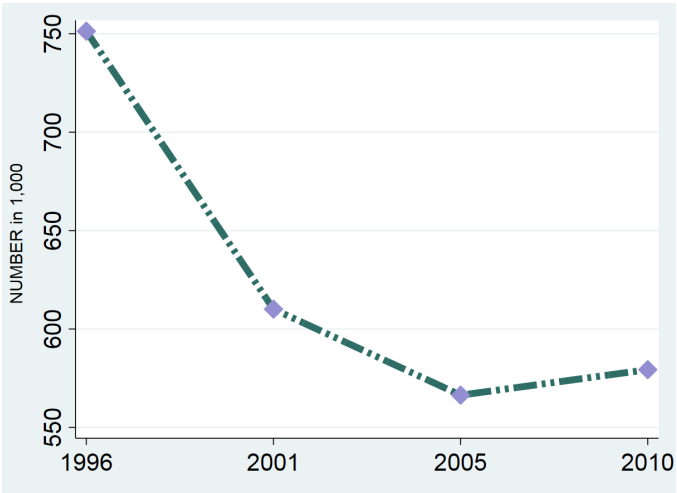
Figure 1.3: The Evolution of Total Number of Retailers



(a) Total Number of Retailers



(b) Total Number of Retail Corporations



(c) Total Number of Sole Traders

Notes: Panel (a) presents total numbers of retailers, Panel (b) shows total numbers of corporation-type retailers, and Panel (c) illustrates total number of sole traders. We plot them during the period of 1996 to 2010.

have noted that while total retail sales have considerably increased, the number of firms has highly decreased since 1996. Therefore, it is easily predicted that the entry of foreign retailers contributes to changes in retail market structure, a transition towards increase in market concentrations. Furthermore, if considering that large discount stores profoundly affect local retail markets as extensively shown by the so-called 'Wal-Mart effect' literature, we would expect that degrees of concentrations may differently be determined across local markets, depending on whether or not foreign stores are present in local markets.

Nevertheless, existing research has rarely studied the relationship between retail globalisation and the two interlinked facets, the pro-competitive effect and the impact of retail market structure. Therefore, this thesis brings together three related, yet independent, essays that aims to study the importance of the retail sector in view of international economics, focusing on competition effects of retail FDI liberalisation and its impact on market structures. Specifically, the first essay explores how the market structure of retail sector is interconnected to outcomes of international trade, the second essay investigates retail FDI's entry determinants when retail markets are imperfectly competitive, and we study, in the last essay, whether the retail globalization leads to the pro-competitive effects.

1.2 Overview

The first essay examines how retail market structure influences the effect of trade liberalisation on consumer prices. Contrary to what people commonly believe, it is often observed that consumer prices do not decline as much as a country removes trade barriers. It is important to explain why consumer prices do not fully react to tariff reductions, because welfare gains of trade crucially depends on the price effect. In this study, we pay attention to the role of retail sector.

We find that tariff cuts imperfectly pass through to retail prices, using Korea's consumer price index (CPI) and tariff data. Guided by a simple theoretical framework built on an oligopolistic competition model with a nested CES demand system (Atkeson and Burstein, 2008), this essay considers market structure of the retail sector as a new channel through which tariffs imperfectly pass through to consumer prices. In words, our prediction based on the model is that, if retail markets are highly concentrated, and thus retailers charge higher markups, then tariff cuts are not fully reflected in consumer prices. To identify the retail channel, we measure the retail markups proxied by price-cost margins, and quantifies the degree to which retail markups exacerbate tariff pass-through using a new instrumental variable (IV) strategy. Exploiting both the Herfindahl index for the wholesale sector and ten-year lagged populations as IVs for retail markups, we

find that a 10 percentage point increase in retail markups contributes to 1.73-1.96 percent lower tariff pass-through. Our results are robust with estimation based on four-year census data by which we account for store-level variations and geographical market heterogeneities in the retail markup measure. The results are also robust when spatially cross-sectional dependences in CPI are controlled for by spatial autoregressive (SAR) regression estimation.

This first essay contributes to the literature in two respects. First, this study provides a new channel of the imperfect tariff pass-through. To our best knowledge, this research is the first study to link retail markups with trade outcomes. Second, we successfully relate consumption-item-level CPI data to product-level(HS 6 digit) tariffs by using a new data matching strategy. We match two data sets by using United Nations Statistics Division(UNSD)'s concordance tables across various product classifications. We believe that the matching result between CPI and tariffs is firstly utilised in this study.

The second essay investigates whether and why multinationals care about their competitors in foreign market entry. In standard FDI theories, industry within which FDI occurs is assumed to be inherently perfectly competitive. Yet, we have seen various oligopolistic industries where multinational enterprises are major players in the global market. Under such imperfectly competitive markets, even multinational enterprises cannot act in a vacuum. They may deeply care about entry choices of domestic competitors, when they decide whether or not to enter host-country markets. This assay aims to empirically study the impact of domestic rivalry on FDI's entry.

This study relies on an empirical context of Korea's retail sector where, as the first and the second biggest retailers, a foreign firm (Tesco) competes with a Korean firm (Emart) by opening multiple large discount stores over Korea's local markets. To shed light on FDI's strategic market entry, we use data on the two firms' store-level information and Korea's retail census data. We find that retail FDI makes strategically enter the Korean market by reduced-form estimations. Specifically, we address the endogeneity problem embodied in foreign entry variable by bivariate probit model with an instrumental variable approach. Our reduced-form identification reveals that the average probability of foreign (Tesco) entry is lower by 10.4% in markets where Korean (Emart) store is present.

In addition, we build an empirical game-theoretical model in which the strategic interaction is fully incorporated in two firms' entry choices. We estimate the structural parameters of the model, and the existence of strategic market entry is clearly confirmed by structural model estimations. Using the structural estimation approach, we further find that asymmetric information between foreign and domestic retailers play a key role in forming FDI's strategic market entry.

This research links the literature on empirical models of entry games to FDI literature. While there is a large strand of industrial organisation(IO) literature on game estimations, competition between foreign and domestic firms has scarcely investigated. In this regard, this study contributes to the literature. This study also relates to literature on FDI's entry determinants. We present that even multinationals take into account their domestic competitors' actions under oligopolistic markets, and we believe that this is a novel point in the literature.

The third essay assesses the impact of entry of multinationals on average markups of local firms in host-country markets using a context of the Korean retail sector. Trade economists recently develop international trade models with endogenous markups so that they can incorporate the pro-competitive effects of trade in the model. Also, there are a body of papers to empirically show that opening up to trade causes fall in markups. On the other hand, the relationship between inward FDI and local markups has been almost unexplored from the literature, both theoretically and empirically. The third assay aims to fill the gap.

A crucial concern when estimating the pro-competitive effect of retail FDI is how to address the selection issue embodied in the entry choice of foreign retail stores. In this essay, we make use of the inverse propensity-score weighting (IPW) regression model in combination with difference-in-difference (DD) method to tackle the selection problem. We find that multinational retailers have a causal impact on decrease in domestic average markups. There is two further issues in identifying the causality: (1) the markup effect may be entangled with productivity spillovers via multinationals; (2) Korean large discount stores may also affect markups charged by other Korean firms. Through the same approach of IPW and DD, we find that both the spillover effect and the impact of Korean large discount stores are limited, and thus the decline in local markups is thoroughly driven by the pro-competitive effects.

We expect this study to contribute to the literature in three folds. First, as mentioned above, this research newly present evidence of FDI's pro-competitive effect. Second, while the IPW and DD approach are exclusively used in estimating direct effects of multinationals (i.e., foreign firms' markup premium over local firms), we adopt the methodology to identify the indirect effect of multinationals on local firms. Third, this essay studies the influence of multinational retailers, rather than manufacturing FDI that existing research has mainly focused on.

1.3 Thesis Outline

The thesis is organised as follows. Chapter 2 explores the relationship between Korea's retail sector and its influence on consumer price effect of trade reforms.

Chapter 3 investigates whether and when multinationals make strategic market entry in response to their competitors' choices. Chapter 4 studies whether FDI causes local firms to decrease their markups using the context of the Korean retail sector. Lastly, Chapter 5 discusses the main conclusions from the three empirical essays and potential extensions of the current thesis.

Chapter 2

Tariffs, Retailers, and Consumer Prices

2.1 Introduction

International trade theories typically suggest that domestic price level declines when a country implements trade liberalization. It is verified by the fact that the removal of trade barriers leads to decrease in prices of foreign products and thus results in fall in domestic prices. Due to the clarity of the relationship, it has been little studied to quantify to what extent tariff reduction affects consumer prices. Yet, it is essential to shed a new light on the question of *tariff pass-through into consumer prices*.

Tariff pass-through directly relates to macro perspective topics which have long been central questions in international trade. How large are welfare gains from trade liberalization? How does trade reform affect inequality and poverty? The answers for these questions crucially depend on how trade liberalization influences consumer gains through the *price channel*. The price channel of trade liberalization is also a critical issue from a policy perspective. Although firms have been at the center of attention from trade economists, consumer welfare is also affected by trade liberalizations, as much as the degree to which consumer prices respond to reductions in tariffs.

It is widely regarded that sensitivity of consumer prices is proportional to changes in tariff rates. However, it is only feasible if there is no market friction from borders to consumer markets, which has been frequently rejected by a recent body of international economics literature. In particular, previous literature does not account for the role of retail firms in determining tariff pass-through. Considering that retailers distribute products from manufacturers or borders to consumer markets, tariff pass-through is highly likely to be influenced by pricing strategies of these intermediaries. This chapter investigates the role of retailers as a channel through which consumer prices incompletely respond to tariff cuts.

To estimate the extent to which tariff affects consumer prices, we consider tariff

reductions that took place in Korea since the mid-1980s. Korea is an idealized case for this study for two reasons. First, Korea has aggressively opened the economy by reducing its tariff barriers. Interestingly, both degrees and timing in tariff reductions are varied across sectors in Korea. In fact, low time frequency in tariff data (i.e., Tariffs change at most once a year) is considered as one of barriers to identify the price effect of trade reforms, which is by contrast no problem to identify pass-through of exchange rates varying very frequently. In this sense, heterogeneous degrees and timing across sectors shown in Korean tariff cuts help to overcome the obstacle. Second, in parallel with tariff reductions, Korea's retail sector has experienced a structural change as well. One of distinctive features is the expansion of large discount stores, which may influence retail market structure differently across regions. For these reasons, we exploit the Korean case in this chapter.

This chapter uses Korean consumer price index (CPI) data. Korean CPI has one great merit that it provides not only sector-level but also region-level price information since 1985. We take advantage of the three-way variations (i.e., sector, region, and year) both to estimate tariff pass-through and to identify whether retailers play a role in imperfect sensitivity of consumer prices to tariffs. We also make use of tariff data gathered over the period of 1988 to 2012. Since tariffs are provided at the disaggregated product level (HS 6-digit), it is important to deal with the matching of product-level tariff data with sector-level CPI data. We match two data sets by using a bunch of existing concordances across product classifications established by the United Nation Statistics Division(UNSD). Through the matching procedure, we link CPI to tariffs, not relying on our own arbitrary decisions but relying on qualified and precise UNSD's concordances.

To test the effect of tariffs on consumer prices, we run a regression with sector, province and year fixed effects along with observable controls to capture time-varying confounders. We find that changes in tariffs are not completely reflected in changes in consumer prices: a 10 percentage point decline in tariff rates is associated with 0.60 percent decrease in CPI. The effect does not seem to be sizeable, because it indicates the estimated decrement in CPI is, on average, no more than 0.42 percent from 1989 to 2012. Interestingly, our finding is consistent with previous empirical studies such as De Loecker et al. (2016) and Topalova (2010), but the magnitude is smaller than those papers. The discrepancy seems to indirectly indicate the existence of retail channel through which tariff pass-through is exacerbated between importers and consumer markets. It is because CPI is the far downstream price while wholesale prices explored by Topalova (2010) and factory gate prices investigated by De Loecker et al. (2016) are upstream relative to consumer prices.

As commonly concerned by the literature, tariffs may be determined by unobserved political economics factors, which would in turn lead to the endogeneity issue. That is, a sector whose outcomes are relatively high would be able to lobby aggressively to the government, and the sector would have favourable tariff rates. If considering that initial Korea's tariff reductions were driven by governmental policy decisions, the endogeneity embodied in tariffs should be seriously taken into account for in the Korean context. We assess the bias from such confounders by conducting a sensitivity analysis proposed by Oster (2017). We find from the sensitivity analysis that the entire effect of tariffs on CPI is unlikely to be driven by the bias from political economics factors. Thus, sensitivity analysis confirms that, while the precise estimate is still unknown, there should exist a causal relationship between tariffs and consumer prices.

Why are tariff reductions imperfectly implemented in consumer prices? Why do tariff cuts pass through to consumer prices less than what is expected? From a theoretical point of view, the answer is less clear. In this chapter, we construct a simple theoretical model to explain why retailers negatively affect tariff pass through. Tariff reductions would not completely pass through to consumer prices if retail firms respond to the marginal cost shocks in a manner that they increase their markups while decreasing retail prices modestly. To uncover the mechanism in a theoretical framework, we consider an oligopolistic competition model with two-level CES demand system based on Atkeson and Burstein (2008). The theoretical framework allows for *variable markups* across retailers, and shows that higher market share is related to lower demand elasticity and thus higher markups. Relying on the setting, we demonstrate that retail markups lower tariff pass-through.

Being guided by the theoretical framework, we empirically investigate the relationship between retail market concentrations and tariff pass-through. Since retail firms, on average, charge higher markups in highly concentrated markets, it is worth testing whether the effect of tariffs on CPI is deteriorated in markets where retail sector is highly concentrated. To test it, we construct the Herfindahl index for the retail sector which varies across provinces and years. Then, we augment the Herfindahl index and its interaction with tariffs to the tariff pass-through regression specification. As emphasised before, the empirical challenge is that tariffs are endogenous. To deal with the issue, we focus on the change within sector-year clusters by including sector-year fixed effects in the regression. Put differently, sector-year fixed effects are able to eliminate any kinds of sector-year specific confounding effects certainly including political economics components. We find that a 10 percentage point higher Herfindahl index causes 0.18 percent lower tariff pass-through rates.

While we show that retail market structure has a significant impact on CPI's sensitivity to tariffs, the question of whether *markups charged by retailers* affect tariff pass-through has not solved yet. To reveal correctly the role of retail channel, the focus should be on the impact of *retail markups* on tariff pass-through. However, it is not straightforward to measure province-level retail markups mainly due to lack of available data. While recent papers proposed a methodology to construct econometrically estimated markups (e.g., De Loecker and Warzynski (2012)), it is extremely data demanding that it is hard to apply it to this study. Our markup measures are proxied by price-cost margins. To construct them, we use KOSTAT (Korea's office of national statistics)'s data on the retail sector where we get access to province-level aggregate data necessary for measuring price-cost margins.

Previously, we only care about tariff's endogeneity, whereas we do not for the Herfindahl index because market structure is assumed to be given exogenously as is done in the literature. On the other hand, when retail markups are adopted, they should be regarded as an endogenous variable as retail markups are not randomly determined in the regression of CPI on tariffs. Specifically, retail firms are more likely to open stores in regions where CPI is higher, thus leading to higher competitions and lower retail markups in those regions. To address the endogeneity, we propose two new instrument variable strategies. The first instrument variable is province-level Herfindahl index for the *wholesale* sector. The instrument strategy crucially relies on the vertical relationship between retailers and wholesalers. While we focus on a horizontal market structure as a retail channel of imperfect pass-through, retail firms' vertical relationship with wholesalers certainly affect their pricing behaviours. In this sense, we use the wholesale Herfindahl index to instrument retail markups. Moreover, it is clear that wholesale market structures influence consumer prices only via retailer's pricing strategies because wholesalers are disconnected to consumer markets. The second instrument variable for retail markups is ten-year lagged populations. The instrument is based on the insight that market sizes are one of crucial determinants of retailers' market entry and location decisions. While populations are sometimes considered as being exogenously given in the economics model, we make use of very long lagged values because they make the variable far more uncorrelated to current consumer prices. From the IV estimation, we find strong evidence that retail markups have negative impacts on tariff pass-through: 10 percentage point increase in intermediary markups cause 1.73-1.96 percent lower tariff pass-through.

Empirical findings so far may rest on how reliable retail markup measures (i.e., price-cost margins) are. Two issues may be raised. First, our price-cost margins are measured by simple ratio of aggregate profits to aggregate sales at

the province level. That is, there is no consideration of store-level variations. Second, geographical heterogeneities within a province are not accounted for, even though provinces are too spacious to be a single geographical market for a retailer. To overcome the issues, we investigate the retail markup mechanism focusing on KOSTAT's census data for the years of 1996, 2001, 2005 and 2010. For the four years, store-level information is accessible, so we first construct weighted average price-cost margins across within-province stores using census data. Census data provides detailed geographical identifiers such as county and town codes. Using county codes, we further build weighted mean markups with county-level market heterogeneities controlled. Limiting to four census years and exploiting two new price-cost margins, we confirm that retail markups truly affect tariff pass-through negatively. Moreover, we incorporate spatial autoregressive (SAR) model to retail markup channel to control for CPI's cross-section dependence at the region level. As expected, the results are very robust.

This chapter is related to empirical studies on trade and prices. Porto (2006) investigates price effects of tariff reforms using consumer expenditure survey data in Argentina. Broda and Romalis (2008) study whether imports from China reduce cost of living in the U.S. using home-scanner data. Faber (2014) finds cheaper access to the U.S. inputs reduces relative prices of high quality products in Mexico. Our study differs from them in that we relate CPI which is comprehensive of all industries to product-level (HS 6-digit) tariffs, and, to our knowledge, this chapter is the first study to match CPI with disaggregated tariff data.

This chapter also complements a literature by providing another evidence that changes in tariffs are not perfectly reflected in prices. Feenstra (1989) shows incomplete tariff pass-through happens if foreign exporters are under imperfect competitions. Nicita (2009) studies the impact of tariff reductions on local prices and shows the price effect differs across regions. Atkin and Donaldson (2015) also present that geographical distance to borders matters to tariff pass-through. Han et al. (2016) show market structure (i.e., the size of private sector) influences the sensitivity of the price effects of tariff reductions in China. De Loecker et al. (2016) find that firms' factory gate prices incompletely react to tariff reduction, because they absorb decline in marginal costs as their markups. Apart from previous literature, we newly suggest that retail markups are a crucial channel of imperfect tariff pass-through into consumer prices.

From a perspective of the retail sector, research on exchange rate pass-through are related to our work. For instance, Burstein et al. (2003) and Goldberg and Campa (2010) show that incorporating the retail sector counterfactually improves the predicted pass-through rates. The exchange rate pass-through literature has paid attentions to two channels related to retail sector. First, local distribution

costs may play an important role in exchange rate pass-through. (e.g., Burstein et al. (2003), Corsetti and Dedola (2005)) Second, markup adjustments of retailers exacerbate exchange rate pass-through. (Nakamura and Zerom, 2010) Even in the exchange rate pass-through, while there is ample evidence that low exchange rate pass-through can be explained by local distribution costs, there is less support for retail markups. (Burstein and Gopinath, 2014) This chapter contributes to the literature by adding new empirical evidence that retail markups matter for marginal cost pass-through.

The rest of the chapter proceeds as follows. Section 2.2 describes policy background and data used in this chapter. Section 2.3 empirically uncovers the causal relationship between tariff reductions and consumer prices. Section 2.4 provides evidence of the impact of retail markups on tariff pass-through. Section 2.5 shows results of robustness checks. Section 2.6 concludes.

2.2 Data

2.2.1 Trade Policy Background

It was not until 1984 that Korea turned to trade reform. The trade reform aimed to achieve two objectives: uniform tariff rates across sectors and considerable reductions in tariffs. Before 1984, Korea's tariff scheme had been very differential across sectors and had maintained high rates. Accordingly, Korea's economy had been dependent on non-traded sectors. The tariff reform, targeting at accomplishing two purposes mentioned above, improved resource allocations and contributed to the growth of traded manufacturing sectors. (Cheung and Hongsik, 2010)

Korea's tariff reform has been carried out in two phases. Each phase has its own end and feature. The first phase of Korea's tariff reduction was conducted from 1984 to 1994. It was a government-driven tariff reduction and focused on manufacturing sectors. Since tariff rates had been substantially differential across sectors, the tariff reform in the first period was implemented in a manner that the Korean government sets pre-determined target rates which are uniform across sectors. As a result of the first reform, Korea's average tariff rate of manufacturing products declined from 40% to 14%.[1]

In 1995, Korea joined the World Trade Organization (WTO), so that Korea had to reduce importing tariffs in order to keep WTO requirements. The second

[1]The first phase of Korea's tariff reforms can be further divided into two waves. The first wave in this phase was conducted from 1984 to 1988. Tariffs had been reduced over the period, targeting at 20% uniformly across final products. The second wave took place from 1989 to 1994. Target rates were far decreased in this period to 8% across final products. (Cheung and Hongsik, 2010)

phase of tariff reforms was driven by the Korea's joining WTO. Interestingly, tariff cuts were mostly implemented in agricultural products in the second phase because manufacturing products were already low enough to satisfy WTO's requirements. The average tariff rate of agricultural products dropped by 12.3 percentage points from 1996 to 2005. (Choi et al., 2012) Even since 1996, manufacturing sector, though, has experienced additional tariff reductions. Unlike the pre-1995 trade reforms by which uniform target rates were applied to the whole manufacturing industries, tariff reduction in the second phase focused on selective manufacturing industries. For instance, in 1997, tariff rates in 182 raw materials and intermediary inputs, which had accounted for a large share in total Korea's imports, declined. In 2000, the Korean government also reduced 43 IT (Information technology) products, because Korea had to comply with agreements by ITA (Information Technology Agreement).

In short, Korea's tariff cuts took place over the long period, and can be divided into two phases, each of which has a respective purpose. As a matter of fact, it is unlikely that tariff reductions took place exogenously, which indicate that endogeneity issue may be addressed when estimating the impact of tariffs on consumer prices. We will elaborate how We deal with the endogeneity in tariffs in section 2.3 and 2.4.

2.2.2 Price and Tariff Data

We make use of Korea's consumer price index (CPI) collected by Statistics Korea (KOSTAT). The KOSTAT is a governmental agency which is responsible for national statistics. Korean CPI data provides not only sector-level but also province-level price information. More specifically, Korean CPI contains price indices across 72 sectors and 16 provinces since 1985.[2]

We use Korean CPI data for three reasons. Firstly, CPI is directly linked to this chapter's research question. Previous literature focuses mainly on export or import unit prices to investigate tariff pass-through. In this study, however, we aim to evaluate the causal effect of tariff cuts on prices that *consumers* face. Considering that CPI is constructed by surveying sales prices at retail stores, it is highly suitable for this end. Secondly, regional variations in Korea's CPI

[2]In fact, *sectors* may not be an appropriate terminology for CPI categories, in the sense that CPI categories are much closer to consumption items that are components of consumption basket, rather than physical products like Harmonised system (HS) codes. Nevertheless, we use the terminology, sectors, for CPI categories for the convenience reason. Above all, as we will explain in section 2.2.3, we successfully match each CPI category to HS codes with a novel matching strategy that we devise. In this regard, we think there is no problem in calling CPI consumption categories *sectors*. Throughout this chapter, we use the terminology CPI sectors instead of CPI consumption item.

play an essential role in this chapter. The variation across regions allows us to investigate the role of retail intermediaries in tariff pass-through, because retail market structure measures used in this chapter are established at the region level. Thirdly, Korea's CPI has an advantage that it provides substantial amount of sectoral price information as well. The sector-level variation is also crucial for this study in the sense that tariff data varies across sectors. As far as we know, there are few countries that publicly release such a substantial amount of sector-level CPI information.

Korean CPI data is not a balanced panel. It is because the number of provinces surveyed for CPI has incremented over years. It started with 6 provinces at first, rose to 15 in 1990, and became 16 from 1995 up to the present. Although the KOSTAT began with only 6 provinces for the first 5 years(1985-1989), the 6 provinces actually accounted for 43.2% of Korea's total GDP.[3] Hence, it seems that even CPI during initial five years is sufficient enough to represent the Korean economy.

We also utilize Korean a tariff dataset drawn from UNCTAD-TRAINS (United Nations Conference on Trade and Development-Trade Analysis and Information System). It contains not only MFN applied tariff rates but also import flows since 1988, both of which are based on six-digit HS (Harmonized System) product level. Unfortunately, it provides Korean tariffs with missing values in several years (e.g., 1991, 1993-94, 1997-98, 2001, 2003, 2005, 2008). We fill the gap with another data source, WTO-IDB (World Trade Organization-Integrated Data Base) which also contains both MFN applied tariffs and imports from 1996 to 2012. We confirmed that two sources have identical information during the overlapped years by both datasets. So, we combine both data sources and use it as one dataset. Still, even using two sources of data does not cover three-year tariffs (1991, 1993 and 1994). In short, the tariff data used for this analysis are applied MFN tariffs across all products at the HS six-digit level and it ranges from 1988 to 2012 with three-year missing values.

Among a total of 72 sectors classified by Korean CPI, this chapter restricts to 26 traded manufacturing sectors which are able to relate to tariffs. In CPI sector classifications, 30 sectors are classified as service sectors(e.g., housing services, insurances, restaurants and hotels), so we drop them because they have nothing to do with tariffs. We further exclude three manufacturing sectors - electricity, gas, and newspapers and periodicals, since there are few records in which products in those sectors have been imported from foreign countries in our dataset. That is, we consider the three sectors as non-traded ones.

[3]The share of 6 provinces in Korea's total GDP is calculated on the basis of 2003. We compute the proportion from KOSTAT's regional GDP data. Since the KOSTAT releases province-level GDP from 2003, we inevitably use the share of 6 provinces as of 2003.

We also exclude 13 traded agricultural sectors for two reasons. Firstly, it is clear that pre-1996 tariffs of agricultural products are inappropriate for an indicator of protection. According to Table 2.1, there appears a sudden and abnormal surge in agricultural tariffs between 1995 and 1996. As a matter of fact, it results from the so-called tariffication. Tariffication was one of UR (Uruguay Round)'s agreements, which required developing countries to convert agricultural non-tariff barriers(NTBs) into tariffs. The high increase in tariffs in 1996, therefore, has nothing to do with increase in protection level. Rather, it is an outcome of the action intended to transform NTBs into quantifiable measures in order for Korea to render its agricultural trade policy more transparent. Secondly, in Korea, special safeguard mechanism (SSM) has often been carried out in agricultural sectors over years. SSM is a tool that allows developing countries to raise tariffs temporarily to deal with import surges or price falls. It explains why there are some agricultural tariff observations that have sky-scraping values in our dataset. Therefore, we do not include agricultural sectors in this chapter.

As noted above, it is unavailable to get access to tariff data in 1991, 1993 and 1994. Recall that a large portion of decrease in manufacturing tariffs took place before 1996. Thus, some may argue that pre-1996 tariffs contain meaningful information and that missing in those data would have a harmful influence on regression results. For the reason, we fill in the missing observations by a linear interpolation. The rationale for using linear interpolation is that the missing seems to be "Missing at Random (MAR)" case, which means it is unlikely to have specific intentions by data providers. More importantly, considering that before 1996, Korea carried out tariff reductions in a way that it just followed a *pre-determined* reduction plan which has a linearly monotonic decreasing trend, we expect that the linear interpolation would be able to revive missing data as correctly as possible. Table 2.2 shows the target tariff rates set by the Korean government from 1988 to 1994.

2.2.3 Matching CPI with Tariffs

Our empirical analysis critically rests on matching CPI data with tariffs. Matching of both data is significantly demanding work in the sense that two sources of data have different levels of aggregation. In particular, CPI is highly aggregated (i.e., CPI has totally 72 sector-level categorises), while tariffs are very disaggregated (i.e., Tariff data has individual values across HS 6 digit level products). More seriously, each data follows different classification standards. Trade data exists on the basis of HS codes, whereas CPI builds on a different consumption-item classification standard, called COICOP (Classification of Individual Consumption

According to Purpose). It would seem that such matching difficulties lead scholars to rarely estimating consumer price effects of trade reforms.[4]

In this chapter, we construct a concordance between CPI and tariffs, which has, to our knowledge, not been generated before. The bottom line of our matching procedure is that we exploit existing United Nations Statistics Division (UNSD)'s concordances across related product classification standards. In fact, the UNSD has established and maintained various kinds of product classification standards, and it offers concordances between them as well. Unfortunately, the UNSD has not generated a concordance between CPI and HS codes. In this chapter, we establish the concordance between them by using UNSD's various concordances as a kind of *bridges* to connect CPI to tariffs.

Figure 1 illustrates the matching procedure that we exploit to relate CPI to tariffs. The procedure consists of four steps. The first step is to match CPI to COICOP. According to the KOSTAT, classification of consumer price index is fundamentally based on COICOP, which in priciple allows for one-to-one matching between CPI and COICOP.[5] Secondly, we relate the matched COICOP with CPC using UNSD's concordance table. In this step, all the matching is automatically conducted only relying on UNSD's concordance table. The next step is to match CPC with HS codes utilising related UNSD's concordance table. As is well-known, HS system has changed its verstion four times so far. For this reason, we firstly match CPC products with HS nomenclature 1996 version (HS96) because CPC is able to match only with HS96 through UNSD's concordance. Then, we match the results to other HS versions, using UNSD's concordances across HS nomenclature versions.[6]

The final step is to construct weighted average tariff rates across CPI sectors. Completing the matching procedure up to the third step, each CPI sector becomes having a great deal of corresponding product (HS six-digit) level tariff rates. Thus, we calculate the average of tariff rates across HS six-digit level products within each

[4]Broda and Romalis (2008) is an exception. They construct a concordance between barcode-based product level prices and HS codes. However, the approach is different from this chapter in the sense that they rely on the arbitrary matching based on authors' subject judgement and product descriptions.

[5]There are, though, four CPI sectors that cannot be directly matched to COICOP. The sectors all belong to the textile industry. The matching is difficult because we need to do match four CPI sectors(*Men's Clothes, Women's Clothes, Casual Clothes* and *Children Clothes*) to two COICOP categories(*clothing material* and *garments*). In this case, we link CPI to CPC (Central Product Classification), without using COICOP as a bridge. Note that, as illustrated in Figure 1, CPC is used as a bridge to connect COICOP to HS codes, which we explain in detail in the main text. Using UNSD's concordance between COICOP and CPC, we find CPC products corresponding to the two COICOP items. Then, we allocate each of matched CPC products to the four CPI sectors by examining descriptions of both CPI and CPC in detail.

[6]COICOP is a reference classification particular for consumption items. CPC is a product classfication based on physical characteristics of products. All the concordances developed by UNSD can be found in UNSD classification registry (http://unstats.un.org/unsd/cr/registry).

CPI sector. The weights that we utilize are product-level(HS 6-digit level) share of one-year lagged imports in total sectoral imports one year before. Although many of related papers has used current imports as weights for average tariff measure, our preferred weights for this study are imports one-year before. The reason is that determination of tariff rates may be affected by foreign imports, which potentially raises an endogeneity issue. Consequently, the formal specification of weighted average tariff of sector i at year t is

$$\tau_{it} = \sum_{j \in I_i} (weight_{ijt} \times \tau_{ijt})$$

where j indexes a 6-digit HS level product, I_i is a set of product j matched with sector i by our concordance table. As described above, the weight of product j and sector i at year t is defined as

$$weight_{ijt} = \frac{Imports_{ijt-1}}{\sum_{j \in I_i} Import_{ijt-1}}$$

2.2.4 Data Summary

Table 2.3 shows the summary statistics of variables in the dataset. The dataset used in this study contains a total of 10,400 observations (26 sectors $\times$ 16 provinces $\times$ 25 years = 10,400). Table 2.3 presents that there are 377 missing values in CPI variable as Korean CPI initially started with only 6 provinces and the KOSTAT has gradually increased the number of provinces surveyed as noted above. The missing values in weighted average tariff data is due to lack of import data in 1987, one year before the sample period (1988-2002). Weighted average tariff rates in 1988 are unable to be established as we use one-year lagged imports as weights. Figure 2.2 illustrates a time trend in the logarithm of aggregate CPI, with a base year 2010 (CPI_{2010}=100). It shows that log CPI increases at the beginning (1988-1998), sluggishly rising or even flat in the middle (1999-2007), and rises again in the late (2008-2012) of sample period.

In Table 2.4, we report sector-level log consumer price indices over years.[7] The table illustrates that most sectors show increasing trends, while three sectors are exceptionally decreasing.[8] Interestingly, the three sectors are all classified into information technology (IT) industry. It seems that decrease in CPI of the three sectors is due to rapid IT technology innovations. In addition, the reduction of

[7] A sector-level log CPI is calculated by averaging out all provincial log CPI within each sector. A sector-level standard deviation of log CPI, thus, measures the degree to which province-level indices are dispersed within each sector.

[8] Three sectors where log CPI has a downward time trend are *equipment for sounds and pictures*, *information processing equipment*, and *telephone and telefax equipment*.

log CPI in three sectors may be a cause of the flat trend of aggregate (log) CPI between 1999 and 2007. In other words, the CPI decrease in IT sectors is so profound that it substantially influences the aggregate CPI.

Table 2.5 describes provincial log consumer price indices over years.[9] Unlike Table 2.4, every province has experienced on-going increase in log CPI over the sample period. A marked feature appeared in Table 2.5 is that differences between provincial indices are minor. This feature contrasts with the high-level of differences between sector-level indices shown in Table 2.4. It clearly suggests that sector-level heterogeneity needs to be more seriously taken into account when regressing log CPI on tariffs. To the end, we use a sector fixed effect in the regression, and we also exploit sector-year specific control variables to account for the heterogeneity issue. We explain details of the fixed effect as well as controls in section 2.3.

Figure 2.3 presents time trends in average tariff rates over the sample period. We draw time trends for weighted average tariffs and simple average tariffs, separately. As well illustrated in Figure 2.3, two average tariff measures have almost identical trends over the sample period. While average tariff rates in general show steadily downward trends, a slight increase appears in 1998. The abnormal rise in 1998 resulted from a temporary trade policy which reacted to Korea's financial crisis in 1997. Afterwards, tariffs has again been dropped up to 2012. As discussed in section 2.2.1, tariff reductions significantly took place before 1995, the year when Korea joined WTO member countries. Although it is still downward trend, tariff reductions after 1995 look very modest. The pattern also accords well with what we explain earlier on Korea's trade reforms. In section 2.2.1, we mentioned that the Korean government selectively reduced tariff rates for some of sectors. In terms of aggregate rates, the degree of reductions does not seem to be substantial. Yet, the selective changes in sector-level tariffs are reflected in the increase in deviations across sectors. Figure 4 illustrates tariff dispersions across sectors. Figure 4 shows that, while sectoral dispersions continuously decline up to 1995, it has been enlarged since 1996 because of the differential tariff policy across sectors.

[9]Each province-level index is the average of sectoral log consumer price indices within the province. Standard deviations present degree of variations in log CPI at the province level.

2.3 Tariffs and Consumer Prices

2.3.1 Regression Specification

Let i index sectors, r denote provinces and t imply years. Equation (2.1) describes our preferred regression specification to estimate the degree to which importing tariffs have an influence on consumer prices in domestic markets.

$$\ln p_{irt} = \beta\tau_{it} + Z_{it}{}'\gamma + \delta_i + \eta_r + \lambda_t + \varepsilon_{irt} \tag{2.1}$$

where lnp_{irt} is the logarithm of the consumer price index (CPI) of sector i in province r and year t. τ_{it} is the tariff rate of sector i in year t, which is the weighted average across products within the sector as described in section 2.2.3. β is a coefficient of interest. It captures tariff pass-through rate which implies how much percentage of CPI would decrease (increase) when tariff rate declines (rises) by 1 percentage point. δ_i is sector fixed effects to account for unobserved time-invariant characteristics across sectors. The province fixed effect η_r is to capture differences across provinces such as geographical heterogeneities and labour costs. We control for macroeconomic fluctuations through λ_t.

In section 2.2.4, we emphasised the necessity of controlling for sectoral heterogeneity by showing Table 2.4. While we include sector-specific fixed effects (δ_i) in the regression equation above, it might be concerned that there are still time-varying unobserved sectoral confounding effects. It is not suitable to include the sector-by-year fixed effects, because the fixed effects will make the variations in tariff rates disappeared. To account for the unobservables, we use three sector-year specific control variables (Z_{it}). The control variables aim to capture time-varying confounders included in the error term. Accordingly, the three variables are expected to be correlated with the independent variable of our interest, tariffs.[10]

First, we employ import penetration ratio. Import penetration ratio, which measures the level of import competition, is defined as the ratio of imports to imports plus domestic productions in a given sector and a given year. Recall that a sector, in this chapter, is defined as a consumption item which is basically a component of consumption baskets used in CPI construction and that we call it CPI sector. This implies that, to measure sector-level import penetrations, both imports and domestic productions are required to be calculated on the basis of CPI

[10] As expected, it is plausible that the control variables are also endogenous variables. Thus, the estimated coefficients for them are likely to be biased. Recall that the objective of the controls are to account for unobserved confounding components, not to examine the impact of control variables on CPI. Moreover, even if they may be endogenous, it is less likely to affect the estimation result for CPI effect of tariffs. We rather expect that adding the controls increase the correctness of the estimate for the tariff variable as they would eliminate the concern about time-varying sectoral confounders.

sectors. For the reason, we assign product-level imports to CPI sectors through our concordance table between HS codes and CPI, and then we compute sector-level imports by summing up all product-level imports belonging to the corresponding CPI sector. However, with regard to domestic productions, it is impossible to measure them across CPI sector-level due to lack of available data. Thus, we use consumption data as a proxy for the denominator of import penetration ratios, i.e., the sum of imports and domestic productions.[11] Although the proxy has a limitation that exports are disregared when we replace imports plus domestic productions by consumptions[12], we view consumptions as a proxy as good as we can, especially under lack of data in Korea.

We use world prices as the second control variable. International trade theories typically suggest that, in small (or price-taking) countries, domestic prices of foreign goods are affected by world prices. In this sense, many related studies include world prices in the tariff pass-through regression.(e.g., Nicita (2009), Feenstra (1989), Han et al. (2016))[13] As it is difficult to measure world prices, we exploit the U.S. export prices for the proxy of world prices, which is widely used in the literature.(Han et al., 2016)) We obtain both US export values and export quantities from UN Comtrade dataset, calculate export unit values (i.e. the ratio of export value to export quantity), and use it as the proxy for world prices. Since the U.S. export data exists at the HS 6-digit level in UN Comtrade, we calculate an average world price of each CPI sector weighted by product-level export share within each sector, and use it as sector-level world price. When relating product-level exports to CPI sectors, we use our concordance described before.

Third, we exploit imports from China to Korea. Since 2001 when China joined WTO, share of China's products in Korea's total imports has rapidly increased. Specifically, the share rose from 17.4% in 1997 to 41.5% in 2007. Like many countries in the world, it has been a common phenomenon that *made-in-China*

[11]The sources of consumption data are Household Income Expenditure Survey(HIES) and population census, all of which are collected by the KOSTAT. Regarding the classification issue, it is very suitable for this study, because the KOSTAT has constructed data at the CPI sector level. As a matter of fact, HIES data does not contain total value of sectoral consumption but *average* value of it, which is a little weird to us. However, population census provides the number of households over years. We construct CPI sector specific consumptions, multiplying sectoral average by the number of households.

[12]When considering the relationship between domestic production and consumption, it can be expressed, in a very simple term, that $consumption = domestic\ production - (exports + imports)$. Therefore, replacing domestic productions plus imports by consumption may be problematic in that exports are not taken into account.

[13]In order for world prices to capture the confounding effect in the regression, there should be a relationship between Korean tariffs and world prices. It is not straightforward to answer whether world prices are correlated to Korean tariffs. A plausible answer is that the global trend toward trade liberalisations, which certainly influences world prices, has affected Korean tariffs. For instance, WTO regime, which acted as a driving force of Korea's trade reforms, have enormously affected on global prices of traded goods.

products are displayed in every corner of Korean retail stores. Thus, it is highly likely that Korea's CPI is influenced by the growth of imports from China. Further, it is clear that Korea's tariff policy is associated with the rapid expansion of imports from China. China's exports to Korea would not have increased that much if it had not been for Korea's tariff reductions. Put together, imports from China to Korea would be a good control variable.

In terms of standard errors for estimates, we use clustered standard errors at the sector-by-year level, which aim to control for spatially serial correlations in the error term. When using a spatial panel data, estimates may be at risk of being contaminated by the spatial error correlation. In fact, Table 2.5, where we report province-level average (log) CPI over years, suggests that there may be a certain co-movement of Korean regional consumer prices over time. On the other hand, there seems to be less concern about sector-level correlation among prices as presented in Table 2.4. Hence, if residual autocorrelation exists, province-level serial correlations, arguably, should be more taken into account than sector-level serial correlations when regressing log CPI on tariffs. In this regard, we choose clustered standard errors at the sector-by-year level so as to capture within-cluster (i.e., province-level) auto correlations.

2.3.2 The Impact of Tariffs on Consumer Prices

Table 2.6 presents the estimation results of Equation (2.1). The parameter of interest in Equation (2.1) is β which implies the effect of tariffs on log CPI. We report the estimates for β in the first row of the table. The pooling OLS results without any controls presented in column (1) show that $\hat{\beta}$ is statistically insignificant. Also, applying either sector fixed effects or province fixed effects to the regression makes no difference with the pooling OLS as reported in columns (2)-(3). However, $\hat{\beta}$ turns into a positive and significant estimate when we include year fixed effects in addition to sector and province fixed effects (column (4)). It may indicate that it is important to control for variations in years because, as shown in section 2.2.4, CPI and tariffs have monotonic time trends. Column (5), where three sector-year specific controls are additionally included, is the result of our preferred regression specification. The estimate for β in column (5) is highly significant and positive.[14]

In terms of magnitudes, column (5) shows that a 10 percentage point fall in

[14]As mentioned in the previous subsection, standard errors are computed by clustering at the sector-by-year level. As a robustness check, we use standard errors clustered at different levels and see if results are changed. We find that, whether clustering at the province-by-year level or clustering at the sector-by-province level, estimated standard errors for every coefficient in Equation (2.1) are almost identical to those in Table 2.6.

tariff rates is associated with 0.60 percent decrease in consumer prices. Interestingly, using the context of India's trade reform from 1989 to 1997, De Loecker et al. (2016) presents a 10 percentage point fall in tariffs leads to a 1.29 percent reduction in factory gate prices of manufacturing firms. Topalova (2010) also shows that 10 percentage point decline in tariff rates in India results in fall in wholesale prices by 0.96 percent during that period. Two aspects are noteworthy. First, although we analyse a different country, Korea, in this study, our results do not look far from research about Indian trade reforms. Second, the size of the effect in this chapter is slightly smaller than De Loecker et al. (2016) and Topalova (2010), but the discrepancy rather seems to accord to what this chapter aims to investigate. In this chapter, we evaluate tariff pass-through to CPI that are downstream prices paid by consumers at the retail level. On the other hand, De Loecker et al. (2016) exploit factory gate prices and Topalova(2010) utilizes wholesale prices, both of which are upstream prices relative to CPI. Therefore, the smaller effect in this chapter than in De Loecker et al. (2016) and Topalova(2010) may indicate that there are retail market frictions by which tariff pass-through rates are deteriorated.

Even though the impact of tariffs on (log) consumer prices is estimated to be significant and positive as expected, it is hard to believe that $\hat{\beta}$ in column (5) is correctly estimated. It is because the estimation so far does not address an econometric challenge when tariff rates are used as a treatment variable: the endogeneity issue embodied in tariffs. In terms of this study, when estimating tariff pass-through, it is essential to control for endogenous determinations of tariff rates with respect to consumer prices. In particular, it is concerned from the trade liberalisation literature that tariff determinations are likely to be influenced by sector-level political powers. This seems to be quite relevant to the context of Korea's trade reforms. Korea's tariff cuts, especially from 1988 to 1995, were driven by governmental policy decision as explained in section 2.2. It is, thus, probable that tariffs are affected by sectoral political pressures or aggressive lobbies on policy decision makers. Moreover, the sector-level political powers highly depend on sectoral outcomes which affect sector-level consumer prices. In this regard, political powers across sectors can be a confounding factor in our regression model. While we use sector-year specific control variables(Z_{it}) plus sector, province and year fixed effects in Equation (2.1), the estimate $\hat{\beta}$ in column (5) of Table 2.6 would still be biased if the political confounding factor exists as an unobservable varying across sectors and years.

In fact, it is not easy to find appropriate econometric methodologies to address the endogeneity issue especially in the context of Korea's tariff cuts. In section 2.2, we explained that Korea's tariff reductions have been carried out in two phases,

each of which has its own purpose and feature. The dichotomy would make it hard to find econometric approaches suitable for the entire sample period. Instead, we conduct a sensitivity analysis to evaluate an upper bound of unobservables-induced bias when estimating Equation 2.1. While it does not deliver information on the unbiased estimate directly, it can be used as the way to examine the robustness of the estimated coefficient of our interest ($\hat{\beta}$). In particular, we exploit Oster (2017)'s formal bounding argument for the bais attributable to unobserved confounding effects.[15]

We want to note that the hypothetical impact of unobserved confounding effects does not necessarily indicate soley the political economics factor. It would rather indicate the sum of the entire omitted variable bias. Nevertheless, if the calculated bias due to unobserved factors does not exceed the estimated treatment effect, then it suggests that our estimation result is to a certain degree robust to unobserved confounders. In other words, it can be said that our treatment effect estimated is unlikely to be thoroughly driven by unobserved confounders which certainly includes the political economics factor. Table 2.7 shows the result of sensitivity analysis. Column (1) shows the regression results of Equation (2.1) from which we exclude observables (δ_i, η_r, λ_t, Z_{it}). Column (2) presents regression results with full sets of controls. Based on the formula defined by Oster (2017) (see footnote 14) and estimation results reported in columns (1)-(2), we calculate an upper bound of the bias as -0.0208. The sensitivity analysis indicates that there is no evidence that CPI effects of tariffs is entirely driven by unobserved political powers. It is because the bias (-0.0208) only accounts for 34.7% of the point estimate for β (0.0599) in column (5) of Table 2.6.

We experiment upon an alternative econometric specification used in several

[15]Let W index observable explanatory variables and $W^{'}$ denote unobservable confounding factors. Let us assume that selection on observables is proportional to selection on unobservable confounding factors, indicating that $\frac{Cov(X,W^{'})}{Var(W^{'})} = \delta\frac{Cov(X,W)}{Var(W)}$, where X is the explanatory variable of interest (i.e., tariffs in this chapter), and δ is a measure of how much correlated both selections are. Now, we consider regressions as below,

$$y = \beta^F X + W + W^{'} + \varepsilon_F$$
$$y = \beta^1 X + \varepsilon_1$$
$$y = \beta^2 X + W + \varepsilon_2$$

and $R^2{}_{max}$ denotes a hypothetical R-squared of the first regression which fully includes not only observables (W) but also unobservables ($W^{'}$). $R^2{}_1$ is R-squared of the model without both observables and unobservables (the second regression), and $R^2{}_2$ R-squared with obeservables (the third regression). In this set-up, Oster (2017) suggests that the unobservable confounding effect can be measured by

$$\text{Confounding Effect} = \delta\frac{(\beta^1 - \beta^2)(R^2{}_{max} - R^2{}_2)}{(R^2{}_2 - R^2{}_1)}$$

empirical papers on tariff pass-through. Initiated by Feenstra (1989), there are many studies (e.g., Nicita (2009), Marchand (2012), Han et al. (2016)) which exploit a regression specification to relate prices to tariffs, based on the following idea: in small countries, the relationship between tariffs and domestic prices can be basically expressed as $p_{it} = p^*_{it}(1+\tau_{it})$, where p_{it} is domestic prices, p^*_{it} is world prices and τ_{it} is tariff rates. In sequence, if we take a logarithm on both sides of the equation, then it would generate $\ln p_{it} = \ln p^*_{it} + \ln(1+\tau_{it})$. In this sense, we re-estimate Equation (2.1) with the tariff variable replaced by $\ln(1+\tau_{it})$ as bellow

$$\ln p_{irt} = \beta \ln(1+\tau_{it}) + Z_{it}{}'\gamma + \delta_i + \eta_r + \lambda_t + \varepsilon_{irt} \tag{2.2}$$

Table 2.8 shows that the replacement makes almost no difference with the original results of Table 2.6. Exceptionally, the impact of $\ln(1+\tau_{it})$ on CPI is highly significant and positive, but the magnitude (0.212) becomes far larger than that of original model (0.0599). In fact, this result is predictable because alternative specification (2.2) is no other than a rescaling version of Equation (2.1) in terms of the tariff variable.

2.4 Retail Markup Channel

In the previous section, we find that tariff reductions lower consumer prices in Korea. Nonetheless, the effect does not seem to be sizeable that much. As mentioned before, we show that a 10 percentage point decline in tariffs is associated with 0.60 percent decline in consumer prices (see Table 2.6, column (5)). From 1989 to 2012, (weighted average) tariff rates fall, on average, by 7 percentage points. Our estimate above indicates that the estimated average price decline by 0.42 percent ($= 7 \times 0.06$), which is a very small effect of tariffs on consumer prices. It is a departure from what traditional trade literature considers tariff pass-through as being perfect, meaning that change in consumer prices is proportional to change in tariffs. Section 2.3, however, provides no answers of what the driving forces through which tariff reductions incompletely affect consumer prices are. This section investigates the role of retailers' market powers in determining the imperfect tariff pass-through into consumer prices.

2.4.1 A Simple Theoretical Framework

A traditional set-up in international trade models is that prices are determined by movement of demand and supply at the *manufacturer* level. Yet, recent papers have paid attention to behaviours of *retailers* in price determination. For instance, Nakamura (2008) finds that barcode-level price variations result mainly

from retail- rather than manufacturing-level, suggesting that the retailers' pricing is a crucial factor in determining consumer prices. In this respect, it is worth studying whether tariff pass-through into consumer prices is affected by retail market powers. As a marginal cost shock, decline in tariffs would not fully pass through to consumers if retail firms respond to tariff reductions in a manner that they raise their markups to a significant degree while decreasing their retail prices modestly.

To guide our empirical analysis, we describe how tariff pass-through is deteriorated by retailers' market powers in the simplest partial equilibrium model based on Dornbusch (1987) and Atkeson and Burstein (2008), and an extention of Dornbusch (1987), Hong and Li (2017). Suppose a retailer sells a variety i in an oligopolistic market and it is assumed to take its marginal cost as given. The retailer, thus, sets the price of a variety i as a markup over marginal cost as the following:

$$p_i = \frac{\epsilon_i}{\epsilon_i - 1} w_i \tag{2.3}$$

where ϵ_i is the elasticity of demand of retailer i and w_i is wholesale cost paid by retailer i. We assume that wholesalers import foreign products, they provide only imported goods to retailers, and that they pay tariffs in order to buy foreign goods. Assuming, too, that the wholesale sector is perfectly competitive, they set wholesale price w_i, such that $w_i = c\tau$ where c is the unit price of imported product and τ is a tariff rate for the imported good.[16] Therefore, the retail price can further be expressed by

$$p_i = \frac{\epsilon_i}{\epsilon_i - 1} \cdot c \cdot \tau \tag{2.4}$$

Equation (2.4) indicates that while the process of supplying goods to consumers involves vertically multiple steps, which gives rise to so-called double marginalisations initially analysed by Spengler (1950), this model can avoid the issue by assuming the wholesale sector as being perfectly competitive. The objective of the setting is that we focus on the relationship between tariff pass-through and retailer's market powers. Rewriting Equation (2.4) as a full log differential,

$$d \ln p_i = d \ln \frac{\epsilon_i}{\epsilon_i - 1} + d \ln \tau \equiv d \ln \mu_i + d \ln \tau \tag{2.5}$$

where we assume that c is constant and μ_i denotes retail markups ($=\frac{\epsilon_i}{\epsilon_i - 1}$).

Equation (2.5) reveals how retail markups relate to tariff pass-through. Sup-

[16] In fact, the more precise expression for wholesale price is $w_i = c(1+t)$ where t is ad-valorem tariffs. For the convenience reason, we denote $\tau \equiv 1 + t$ and thus wholesale price can be defined as $w_i = c\tau$ as above, which never change the implications of the model. When we consider specific tariffs rather than ad-valorem tariffs, wholesale price w_i will be $c + \tau$. Though choosing whichever types of tariffs, results are not changed.

pose retail markets are perfectly competitive and μ_i is thus zero, then tariff pass-through will definitely be one ($\frac{d\ln p_i}{d\ln \tau} = 1$), which we call a perfect pass-through case. Consider that retail markup is not variable with respect to tariffs ($\frac{d\ln \mu_i}{d\ln \tau} = 0$). In this case, tariff pass-through will also be zero. Further, tariff pass-through will be more harmed if retailers charge more markups in response of tariff reductions ($\frac{d\ln \mu_i}{d\ln \tau} < 0$).

To make it explicit that variable retail markups affect tariff pass-through from a theoretical perspective, we consider an oligopolistic competition model of variable markups with a nested CES demand system following Atkeson and Burstein (2008).[17] We assume a two-step CES demand with elasticities of substitution ρ between retail varieties within sectors, and $\rho > \eta > 1$ between sectors. In this setting, the demand faced by retailer i producing its own variety is:

$$q_{is} = p_{is}^{-\rho} P_s^{\rho-\eta} Q_s \tag{2.6}$$

where q_{is} is quantity demand of retail variety i, p_{is} is the retailer's price, and Q_s is the sectoral total demand. The sectoral price index is given by $P_s \equiv (\sum_i p_{is}^{1-\rho})^{1/(1-\rho)}$. By definition, market share of firm i is

$$s_{is} \equiv \frac{p_{is}q_{is}}{\sum_{i'} p_{i's}q_{i's}} = (\frac{p_{is}}{P_s})^{1-\rho} \tag{2.7}$$

Under this setting of the model, demand elasticity of retailer i is then derived[18]

$$\epsilon_{is} \equiv -\frac{d\ln q_{is}}{d\ln p_{is}} = \rho(1 - s_{is}) + \eta s_{is} \tag{2.8}$$

where s_{is} is retailer i's market share within sector s.

Equation (2.8) has two important implications as follows: Retail firms with high market shares make their demand less sensitive to their own prices (i.e., low demand elasticities). In addition, high-market-share firms charge higher markups than firms with low market shares. In other words, markups are increasing with market shares, which implies that, unlike Melitz (2003), markup is not constant in this model.

Using (2.4) - (2.8), we can determine how retail markups are associated with tariff pass-through. First, the reduction in tariffs give rises to fall in firm's price by Equation (2.4). When its price decreases, the firm's market share will be higher because the market share is increasing with lower prices, given fixed sectoral price

[17] To describe the model, we follow a treatment presented by Amiti et al. (2014), a paper that explores the role of variable markups and imported inputs in low exchange rate pass-through based on Atkeson and Burstein (2008).

[18] Note that $\partial \ln P_s / \partial \ln p_{is} = s_{is}$.

index, as shown by Equation (2.7). According to Equation (2.8), the rise in market shares leads to fall in firm's elasticity of demand. As a final step, since tariff reductions make the retail firm increase its markups following decrease in demand elasticity, tariff pass-through go downward due to retail markup channel as we explained in (2.5). Therefore, it is clear that retail markups have a negative impact on tariff pass-through.

2.4.2 Retail Market Concentrations and Tariff Pass-Through

Being guided by the simple theoretical framework described by the previous subsection, we start by estimating the relationship between retail market concentrations and tariff pass-through. In highly concentrated retail markets, retail firms, on average, are able to charge higher markups. It is thus predictable that the impact of tariffs on consumer prices is deteriorated in regions where retail market structure is highly concentrated. To test it, we make use of the Herfindahl index for retail markets which varies across provinces and years. The Herfindahl index measures degrees of retail market concentration, where the index in province r and year t (HI_{rt}) is defined as

$$HI_{rt} = \sum_{f} s_{frt}^2$$

where f denotes a retailer and s_{frt}^2 implies the squared market share of retail firm f in province r and year t. Our identification strategy is that, given the tariff pass-through regression specification described in Equation (2.1), we augment it by adding the Herfindahl index and its interaction with a tariff variable on the right-hand side of Equation (2.1). Although tariffs and the Herfindahl index are respectively established at the sector- and at the province level, the impact of the interacted variable is able to be estimated by taking advantage of three-way (i.e., sector, province and year) variations in CPI data. The regression specification to explore the retail market concentration by which tariffs imperfectly pass through is the following:

$$\ln p_{irt} = \alpha HI_{rt} + \beta\tau_{it} + \theta(HI_{rt} \times \tau_{it}) + Z_{it}{}'\gamma + X_{rt}{}'\rho + \delta_i + \eta_r + \lambda_t + \varepsilon_{irt} \quad (2.9)$$

We include sector(δ_i), province(η_r), year(λ_t) fixed effects and sector-year specific control variables(Z_{it}) as in Equation (2.1). In addition to those controls, we add province-year specific observable variables(X_{rt}).[19] The purpose of adding X_{rt} is to alleviate the concern that there may be missing components varying across

[19]The controls that we use are log population density and log gross regional domestic product (GRDP), both of which are varying across provinces and years.

provinces and years. We expect the estimate for θ is negative and statistically significant, because the impact of tariffs on consumer prices is formally expressed as

$$\frac{\partial \ln p_{icrt}}{\partial \tau_{it}} = \hat{\beta} + \hat{\theta} HI_{rt} \tag{2.10}$$

This equation implies that tariff pass-through varies conditional on retail market concentrations (the Herfindahl index). Suppose that province r has a higher market concentration in the retail sector than province $r^{'}$ (i.e., $HI_{rt} > HI_{r^{'}t} > 0$) and that $\hat{\theta}$ is negative. Then, the tariff pass-through rate in r will be smaller than in $r^{'}$ as we hypothesize, because

$$\hat{\beta} + \hat{\theta} HI_{rt} < \hat{\beta} + \hat{\theta} HI_{r^{'}t}$$

In section 2.3.2, sensitivity analysis relying on Oster (2017) revealed that confounding effects, which we consider as sector-level heterogeneities in political powers like lobbying abilities, crucially influence estimates for tariff pass-through. Although it was not as serious as the estimate is thoroughly driven by the political confounders in estimating Equation (2.1), it is uncertain how much degree sectoral political powers would contaminate the estimate for θ in the retail market concentration regression (Equation (2.9)). Moreover, as we mentioned in section 2.3.1, sector-level political confounders are likely to exist as time-variant unobservables. To deal with the potential endogeneity issue, we focus on the change within sector-year clusters by including sector-year fixed effects (ϕ_{it}) in Equation (2.9). We expect ϕ_{it} to control for any kinds of sector-year specific confounders certainly including the political economics factors. Our second regression specification for retail market channel is thus as follows:

$$\ln p_{irt} = \alpha HI_{rt} + \theta(HI_{rt} \times \tau_{it}) + X_{rt}^{'} \rho + \phi_{it} + \eta_r + \varepsilon_{irt} \tag{2.11}$$

Note that, compared to Equation (2.9), τ_{it} and Z_{it} are disappeared in this regression as ϕ_{it} absorbs τ_{it} and Z_{it}.

Table 2.9 illustrates evidence that retail market concentrations matter with tariff pass-through determinations. Columns (1)-(2) reports the regression result of Equation (2.9). The first column shows a estimation result when excluding province-year specific controls(X_{rt}), whereas the second column presents a result with X_{rt} used. The parameter of interest is θ, which identifies the extent to which market concentration level in the retail sector exacerbates tariff pass-through rates. In both (1) and (2), the estimates for θ are negative as expected, but is not statistically significant. Column (3) demonstrates the regression results of Equation (2.11) in which the endogeneity issue embodied in tariffs is completely

controlled by sector-year fixed effects(ϕ_{it}). The estimate for θ is now statistically significant at the 5 percent level, suggesting that a 10 percentage point higher Herfindahl index causes 0.18 percent lower tariff pass-through at the province level. The result contrasts sharply with columns (1)-(2) where the endogeneity is unsolved, in respects of significances as well as magnitudes. Column (1) and (2) have lower $\hat{\theta}$ than column (3) (i.e., $\hat{\theta}$ in columns (1)-(2) has a downward bias), which indicates that political economics factor significantly matters at least to CPI's sensitivity to trade reforms. Most importantly, however, regression results shown by column (3) in Table 2.9 strongly confirm that the hypothesis "the higher retail market concentration is, the lower tariff pass-through to CPI is." is true.

2.4.3 The Role of Retail Markups in Tariff Pass-Through

We have focused on the market structure of the retail sector to explain how retail market concentrations account for the imperfect pass-through of tariff reforms to consumer prices. The rationale behind the argument is that markups charged by retail firms function as market frictions negatively affecting tariff pass-through. The Herfindahl index is, however, not a markup measure but only an indicator of how much degree a market structure is concentrated.[20] To reveal more precisely whether the retail channel acts as a friction when tariff cuts pass through to consumer prices, the focus should be on the impact of *retail markups* on CPI sensitivity to tariffs. The corresponding regression specification is, thus, formulated by

$$\ln p_{irt} = \alpha\mu_{rt} + \theta(\mu_{rt} \times \tau_{it}) + X_{rt}{}'\rho + \phi_{it} + \eta_r + \varepsilon_{irt} \tag{2.12}$$

where μ_{rt} is the markup in province r and year t. As noted in the previous subsection, sector-year fixed effect ϕ_{it} seeks to control for unobserved confounders, which certainly deals with the endogeneity by political economics factor.

There are two challenging issues in examining the role of retail markups. First, it is not straightforward to measure province-level markups for the retail sector. A body of papers adopt a recently developed approach, such as De Loecker and Warzynski (2012) and De Loecker et al. (2016), that empirically estimates markups through the structural model estimations. Despite its elegance in terms of being based on solid theoretical backgrounds, the approach is extremely data demanding that it is hard to employ it to this chapter which is built on the context of Korean retail sector. In particular, the methodology requires us to make use of highly disaggregated firm-level(De Loecker and Warzynski, 2012) or firm-product

[20] In fact, the Herfindahl index is sometimes used as a markup proxy for a particular industry or a geographical market from the literature. In our view, though, it is not because the index is truly a good proxy for markups, but because of lack of proper data.

level(De Loecker et al., 2016) data, because it relies on estimation of output elasticities with respect to variable inputs in the production functions. Unfortunately, there is no data that is such disaggregated and that covers the entire sample period in this study. Second, the endogeneity problem arises once we consider *retail markups* as a mechanism of incomplete pass-through. Previously, we only care about the endogeneity embodied in the tariff variable in Equation (2.9) and (2.11), and the Herfindahl index was treated as an exogenous variable because retail market structure is assumed to be given exogenously as is done in the literature. Yet, it is hard to believe that retail markups are exogenously given. As a result, it is critically important to control for the endogeneity in the markup variable.

Constructing retail markup measures Our markup variable is proxied by price-cost margins constructed by accounting data. While the accounting price-cost margin has often been criticised[21], we consider the measure as a good proxy for retail markups especially under data limitation in Korea. Province-level retail markups are constructed based on KOSTAT(Korea's office of national statistics)'s data for the retail sector. The KOSTAT annually collects information on a sample of Korean retail establishments since 1988. The sample survey encompasses a number of variables, especially those included in income statements. To name a few, it includes annual sales, costs of goods sold, operating costs, and so on. There are two drawbacks in KOSTAT's annual survey. First, it does not contain any information that helps to identify establishments, such as establishment(store) names, firm affiliations and addresses. It is thus impossible to use it as a store-level panel data. However, it provides province codes across establishments belonging to the sample. The province codes enable us to build province-level measures for every variable, all of which are able to be utilised as a panel data. Second, although KOSTAT's retail data is de facto a repeated cross-section data because of no information on store identifications, even the restricted store-level data is only accessible for the post-1996 period. KOSTAT's retail data during the pre-1996 period is provided only as aggregated values, i.e., the sum of values across variables. Figure 2.3 shows that (average) tariff rates decline considerably during the period of 1988 to 1996. To analyse the impact of tariff reductions for years before 1996, we have no choice but to utilise the aggregate values over the whole sample period (1988-2019) so as to establish consistent markup measures over years.

[21]I use Equation 2.13 for measuring price-cost margins. The measure, which approximates price-cost margins using accounting data, may be baised since it typically involves cost measures that do not approximate marginal costs.(Bresnahan, 1989)

In principle, price-cost margins measure the relative importance of markups to prices. To correctly measure, it is required to exploit marginal cost (c_{rt}) information, which is in fact quite difficult to obtain. As in the literature, we use a roundabout method as defined by the following:

$$\mu_{rt} = \frac{p_{rt} - c_{rt}}{p_{rt}} = \frac{(p_{rt} - c_{rt})q_{rt}}{p_{rt}q_{rt}} = \frac{profit_{rt}}{sale_{rt}}$$

where p_{rt}, c_{rt} and q_{rt} are price, marginal cost and sale quantity in province r and year t, respectively. By accounting conventions, profits are calculated as annual sales less costs of goods sold and operating costs. Therefore, price-cost margins can be further expressed as

$$\mu_{rt} = \frac{profit_{rt}}{sale_{rt}} = \frac{(\text{AS}_{rt} - \text{COGS}_{rt} - \text{OC}_{rt})}{\text{AS}_{rt}} \tag{2.13}$$

where AS_{rt}, COGS_{rt} and OC_{rt} denote total annual sales($\equiv \sum_s \text{AS}_{srt}$), total costs of goods sold($\equiv \sum_s \text{COGS}_{srt}$), and total operating costs ($\equiv \sum_s \text{OS}_{srt}$) in province r and year t, respectively. As noted above, all aggregated data necessary to calculate the price-cost margin specified in Equation (2.13) are accessible from KOSTAT's sample survey for the retail sector.

Endogeneity and Instrument Variables In Equation (2.12), retail markups μ_{rt} are endogenously determined as the following reason: Retail firms tend to take consumer prices into account seriously when deciding where to open stores. To be specific, retailers are more likely to open stores in geographical markets where consumer prices are relatively high, because higher consumer prices would contribute to higher profits. Then, it would happen that more retail stores enter intensively into markets in which consumer prices relatively high. In turn, retail market competition would be higher in markets with high level of consumer prices. As a result, average markups charged by retailers would become lower in those markets, which potentially leads to biased estimates for α as well as θ in Equation (2.12).

To address the endogeneity issue, we propose two new instrumental variable strategies. The first instrument variable is the province-level Herfindahl index of the *wholesale* sector. As discussed so far, this chapter pays attention to horizontal market structure of the retail sector. More specifically, we highlight that horizontal market concentrations influence average retail markups, which would eventually affect tariff pass-through. On the other hand, retail firms also have a vertical relationship to wholesale firms. The vertical relationship is clearly documented by the standard industrial classifications, mentioning that establishments

in the wholesale sector are defined as those who are involved in selling products to retailers. In this sense, retailers' pricing behaviours are arguably determined by the vertical relationship with wholesale firms.

In particular, retail markups are likely to be influenced by market concentrations of the wholesale sector. Our data confirms well the relationship between retail markups and wholesale Herfindahl index. We make a plot of (average) retail markups against (average) wholesale Herfindahl index in Figure 2.5. As clearly shown in the figure, two variables are negatively associated with each other. It may indicate that wholesale market concentrations would adversely influence retailer's markup.[22] While the relationship between two variables is intuitively appealing, theoretical works on retail markups and wholesale market concentrations are scarce. A plausible channel by which retail markups are affected by the vertical relationship is, as suggested by Gaudin (2016), the bilateral bargaining between upstream and downstream firms. Specifically, great wholesale market powers increase their bargaining powers as sellers of products to retailers. Then, retailers inevitably take higher price of inputs and thus their markups are negatively affected. Furthermore, wholesale market concentration is unlikely to have a direct impact on CPI determinations, because wholesalers affect consumer markets only *via retail firms*.[23] Guided by this logic, we construct Herfindahl index for the wholesale sector using annual sample survey on Korean wholesalers and use it to instrument for retail markups. The KOSTAT does not only undertake store-level sample survey on retail sector, but it also does establishment-level data on the wholesale sector.

The second variable by which we instrument for the retail markup variable μ_{rt} is ten-year lagged populations. Retailing literature has long emphasised the importance of market sizes as an influential determinant of retail store's location and market entry decisions. Figure 2.6 illustrates location pattern of a particular retail format, large discount stores, as of 2010.[24] Although Figure 2.6 only shows locations of large discount stores, it would be a good example of how retail stores are geographically distributed. Figure 2.6 presents that retail stores are concen-

[22]Instead of analytic Herfindahl index, we use *empirical* index when drawing Figure 2.5. We obtain it by regressing analytic Herfindahl index on province- and year fixed effects. The empirical index aims to construct average Herfindahl index which is to some degrees free from province- and year-level heterogeneities issues. More importantly, since the first stage IV regression model includes the two fixed effect terms, it would be a better way to plot retail markups against the empirical wholesale Herfindahl index rather than the analytic index. Nevertheless, we want to note that using the analytic Herfindahl index makes also very clear negative relationship between two variables.

[23]Our data also ensures that the IV (wholesale Herfindahl index) is uncorrelated with CPI. Regressing (log) CPI on the wholesale Herfindahl index and observable controls(X_{rt}) and fixed effects(δ_i, η_r, λ_t), we find that the coefficient of the index is not statistically significant.

[24]We draw Figure 5 using Korea Chainstore Association (KOCA)'s data on large discount stores.

tratedly spotted on regions with large population, such as Seoul, Kyunggi-do and Pusan. Accordingly, these provinces are likely to have higher competition level, and thus are expected to have lower retail markups.

While populations are often considered as an exogenously given variable, we make use *ten-year lagged* populations, because using long lagged populations would make the variable far more uncorrelated to current consumer prices. To confirm the validity of the proposed instrument, we run a simple regression that relates consumer prices to ten-year lagged population augmented by sector-, province- and year fixed effects plus province-year specific observed controls.[25] Table 2.10 reports the estimation results of the simple experiment. As expected, the parameter of 10-year lagged population is estimated to be statistically insignificant. Hence, we would expect that long lagged populations have no impact on CPI, but indirectly do only through retail markups.

IV results Retail markup μ_{rt} is endogenous, and thus its interaction with tariffs, $\mu_{rt} \times \tau_{it}$, is potentially endogenous as well. To implement the instrument variable approach, at least two instruments are required to be applied as we have two endogenous variables. Of course, we already proposed two instrumental variables - the Herfindahl index of the wholesale sector and ten-year lagged population - for addressing the endogeneity. The two proposed IVs, yet, are only targeting at μ_{rt}. Put differently, the two variables may not be able to be good and valid instruments for the interaction term $\mu_{rt} \times \tau_{it}$. For the reason, on top of two proposed IVs, we utilise their interactions with a tariff variable as additional instrument variables.[26]

Table 2.11 presents the results of Equation (2.12). Column (1) reports OLS estimation results in which IV strategy is not applied. In contrast, column (2) shows the IV estimation results.[27] The estimate of interest in this chapter is that of an interaction between retail markups and tariffs which indicates whether and how retail markups affect CPI's sensitivity to tariffs. The OLS point estimate for θ, a coefficient of $\mu_{rt} \times \tau_{it}$, is of negative sign and is significant at the 5 % level. Even though the OLS result seemingly corresponds well to our hypothesis that retail markups cause fall in tariff pass-through, it should be biased because endogeneity problem is not solved. IV results in column (2) also report that the point estimate

[25]The regression specification is formulated by the following: $\ln p_{irt} = \beta population_{r,t-10} + X_{rt}{}' \gamma + \delta_i + \eta_r + \lambda_t + \varepsilon_{irt}$

[26]One may argue that tariff variable τ_{it} is also endogenous variable. Recall that, however, we showed how sector-year fixed effects (ϕ_{it}) are able to address the endogeneity embodied in τ_{it}. Thus, the endogeneity issue related to the tariff can be controlled for as long as there is sector-year fixed effects(ϕ_{it}) in the IV regression model.

[27]Column (2) as well as (4) are regression results of the second stage IV model. We show the results of the first stage IV estimation in the appendix.

for θ is negative and statistically significant at the 10 % level, but the magnitude (-0.173) becomes much lower than the OLS esimate (-0.0367). The upward bias shown by the OLS result makes sense when recalling the reason why retail markups are endogenous. We already explained that the variable is endogenous because retail markups affects consumer prices and, at the same time, consumer prices do influence retail markups. The endogeneity issue is, thus, distilled as a reverse causality problem. It is no wonder that the IV point estimate for θ is smaller than the OLS results, because sign of the reverse causal effect (i.e., the effect of consumer prices on retail markups) is negative.

As an alternative IV regression model, we augment Equation (2.12) by adding province-year fixed effects. The aim of this strategy is basically similar with when we employ sector-year fixed effects to eliminate political economics factor embodied in tariffs. Put simply, we seek to control for endogeneity in markups through province-year fixed effects, because it will be able to absorb any unobserved endogenous selection of retail markups into CPI. The approach is formulated by the following:

$$\ln p_{irt} = \theta(\mu_{rt} \times \tau_{it}) + \phi_{it} + \pi_{rt} + \varepsilon_{irt} \tag{2.14}$$

where π_{rt} denotes province-year fixed effects. Note that since π_{rt} also absorbs observed variables varying across provinces and years, μ_{rt} and X_{rt} in Equation (2.12) are not appeared in Equation (2.14). We also want to emphasize that the estimate for θ may still be biased. While province-year fixed effects π_{rt} deal with endogeneity of μ_{rt}, it could not address that of $\mu_{rt} \times \tau_{it}$ because the interaction term varies at the sector-by-province-by-year dimension. For the reason, we use interactions between proposed IVs (i.e., 10-year lagged populations, wholesale Herfindahl index) and the tariff variable as instrumental variables for $\mu_{rt} \times \tau_{it}$ in Equation (2.14) as well. Column (3) and (4) in Table 2.11 present estimation results of Equation (2.14).

In terms of both magnitudes and significances, estimation results of column (3) and (4) are very similar to column (1) and (2). This suggests that as long as we instrument μ_{rt} and $\mu_{rt} \times \tau_{it}$, the estimate for θ is robust whether the sector-year fixed effect is used(column (4)) or not(column (2)). We report Kleibergen-Paap rk Wald F statistics for the IV results in column (2) and (4). The statistics aim to reveal results of weak IV tests in case where more than two IVs and clustered standard errors are utilised. The statistics all exceed critical values developed by Stock and Yogo (2005), meaning that there is no concern of weak IV problems.

In conclusion, IV results in column (2) and (4) indicate that the a 10% point increase in retail markups (price-cost margins) causes 1.73% - 1.96% lower tariff pass-through. The regression results so far clearly confirm that retail markups

exacerbate the effect of tariffs on consumer prices.

2.5 Accounting for Heterogeneities and Cross-Section Dependence

2.5.1 Market Heterogeneities in Retail Markups

In this subsection, we analyse the retail markup channel of which regression model is specified by Equation (2.12) and (2.14), focusing on KOSTAT's *census* data for the years of 1996, 2001, 2005 and 2010. In parallel with sample survey on retailers, the KOSTAT has undertaken store-level retail census in those four years. Considering that Korea's tariff cuts have taken place extensively over years, KOSTAT's annual sample data seems to be proper for our study. Nevertheless, it is also worth estimating Equations (2.12) and (2.14) by limiting to the four census years for two reasons.

Firstly, retail census data makes us overcome the limitation of price-cost margin (PCM) measures. Recall that we construct PCM measures using aggregate (i.e., total sum of values) data provided by KOSTAT's annual sample survey. Accordingly, the reliability of PCM measures crucially depends on whether or not KOSTAT's retail sample represents population well. On the other hand, there is no need to concern about the reliability issue when constructing PCM by census data, because the retail census covers entire population of Korea's retail stores. Regarding the number of observations, retail census amounts to 600,000-770,000 per year whereas retail sample data contains information on 20,000-50,000 stores per year. Secondly, retail census data allows us to address the potential heterogeneity issue in provincial markup measures. While retail sample data only contains province codes as a geographical identifiers, retail census data has narrower geographical indicators such as county codes and town codes.[28] This implies that, when constructing provincial PCM measures relying on census data, it is possible to account for county- or town-level market heterogeneities by using corresponding fixed effects.

The province-level PCM measure was defined as

$$\mu_{rt} = \frac{profit_{rt}}{sale_{rt}} = \frac{\sum_s profit_{rst}}{\sum_s sale_{rst}}$$

where r, s and t index provinces, stores and years respectively, and $profit_{rt}$ and $sale_{rt}$ are total sums of store-level values within a province r. Note that the

[28]Even from census data, however, it is inaccessible to store identification information (e.g., store names, firm affiliations, etc.) as in annual sample data.

measure has nothing to do with a statistical mean but it is rather a ratio of aggregate profits to aggregate sales at the province level. Although we use it inevitably for data issue as we mentioned in section 2.4.3, some may argue that the measure is not precise enough to fully reflect province-level markups. Now, exploiting a richness of store-level census data, we construct a new PCM measures at the province-level as follows:

$$\mu_{rt}^{\text{Census1}} = \sum_{s \in R_r} (weight_{rst} \times \mu_{rst}) \tag{2.15}$$

where t is now *census* years. μ_{rst} denotes a PCM of store s in province r and census year t. $weight_{rst}$ is a ratio of store-level sales to total provincial sales, which is formulated by $\frac{sale_{rst}}{\sum_s sale_{rst}}$. The new PCM measure $\mu_{rt}^{\text{Census1}}$ is therefore a weighted average PCM of all stores within a province. It is clear that the new PCM $\mu_{rt}^{\text{Census1}}$ is a better measure than the previous one in the sense that the average statistic is derived from a population of Korea's retail stores.

Even though exploiting store-level variations based on entire population of retail stores, the measure $\mu_{rt}^{\text{Census1}}$ may has another limitation that geographical heterogeneities within a province are not accounted for. Irrespective of locations, the former measures take every store into account identically. In fact, provinces may be too spacious to be a geographical market for a retail store. While we use sales as weights for measuring $\mu_{rt}^{\text{Census1}}$, it may be concerned that market heterogeneities would cause $\mu_{rt}^{\text{Census1}}$ to be incorrectly measured.

Using the census data where detailed geographical codes are available, we further establish a new PCM measure in which county-level market heterogeneity is controlled for. We construct it in three steps. As a first step, we compute weighted average of store-level PCM at the *town* level. The weight is store-level sales proportion as noted before. Thus, the formal specification for weighted average PCM of county j and town w at year t (μ_{rjwt}) is

$$\mu_{rjwt} = \sum_{s \in J_w} \left(\frac{sale_{rjwst}}{\sum_{s \in J_w} sale_{rjwst}} \times \mu_{rjwst} \right)$$

where r indexes provinces, j denotes counties, w means towns, s implies stores and t indicates census-years. J_w is a set of store s belonging to town w. The second step is to generate *empirical* county-level PCMs by running a simple regression as follows:

$$\mu_{rjwt} = \alpha + \sum_{j=1}^{251} \gamma_j county_j + \sum_{t=1}^{4} \delta_t year_t + \varepsilon_{rjt}$$

The equation above is nothing but a regression of μ_{rjwt} on county fixed effects

and census-year fixed effects. The aim of the regression is to derive a fitted value of PCM at county j in census year t, which is to a considerable degree free from heterogeneity issues. Let $\hat{\mu}_{rjt}$ denote fitted PCMs obtained by the second step. Then, the last step is to construct the provincial weighted average across fitted county-level PCMs($\hat{\mu}_{rjt}$) as below:

$$\mu_{rt}^{\texttt{Census2}} = \sum_{j \in R_r} \left(\frac{sale_{rjt}}{\sum_j sale_{rjt}} \times \hat{\mu}_{rjt} \right) \tag{2.16}$$

We re-estimate Equation (2.12) and (2.14), replacing retail markup measures by $\mu_{rt}^{\texttt{Census1}}$ and $\mu_{rt}^{\texttt{Census2}}$, and using data for four census years. The instrument variables used in the new estimations are same as before (i.e., ten-year lagged populations and wholesale Herfindahl index for the four census years). Table 2.12 shows the estimation results of the regressions. Columns (1)-(4) report the results when we use weighted average PCM $\mu_{rt}^{\texttt{Census1}}$ for the markup measure. Among them, column (1) and (2) are estimation results of Equation (2.12), whereas column (3) and (4) are estimation results of Equation (2.14). Regarding the estimated parameter for the interaction between retail markups and tariffs, OLS results in (1) and (3) are not significant unlike Table 2.11, but IV results shown in (2) and (4) are significant at the 10% level and negative. Strikingly, the *negative* impact of retail markups on tariff pass-through estimated by IV approach substantially decline compared to those in Table 2.11. This suggests that accounting for store-level variations in markups by exploiting $\mu_{rt}^{\texttt{Census1}}$ is meaningful. Nonetheless, the IV results clearly indicate that retail markups negatively affect tariff pass-through even after store-level census data is exploited: 10 percentage point increase in retail markups (price-cost margins) causes 0.64-0.69 percent lower tariff pass-through.

Columns (5)-(8) are estimation results when we employ another markup measure $\mu_{rt}^{\texttt{Census2}}$, which is weighted average PCM with county-level market heterogeneity controlled. While being slightly different in magnitudes, the estimates reported in columns (5)-(8) show very similar results as when we use $\mu_{rt}^{\texttt{Census1}}$. The implication of the results is that it is important to account for the market heterogeneity, but once average markups are constructed based on store-level variations, it seems to make little difference. Most importantly, we find that there is a negative causal effect of retail markups on CPI's sensitivity to tariffs, even after controlling for the market heterogeneity in markup measures: 10 percentage point increase in retail markups (price-cost margins) causes 0.85-0.89 percent lower tariff pass-through.[29]

[29]In the Appendix, we report the results when we employ a retail markup measure where, rather than county-specific heterogeneity, town-level market heterogeneity is controlled by the same approach described above. We find that accounting for town-level heterogeneity provides quite robust results.

2.5.2 Cross-Section Dependence in Consumer Prices

This subsection assesses the spatial dependence between consumer prices. One may argue that each provincial CPI is affected by consumer prices of others. From a perspective of econometrics, the cross-section dependence would cause a serious identification problem, which results in biased estimates of coefficients in the regression model. To control for it, we exploit an approach in the field of spatial econometrics to take into account interactions across provincial consumer prices.

We run the retail markup channel model specified in Equation (2.12) and (2.14), by adopting the spatial autoregressive (SAR) model. More specifically, we include a spatial lag term ($W \ln p_{icrt}$) on the right hand side of our regression model, which is given by

$$\ln p_{irt} = \alpha\mu_{rt} + \theta(\mu_{rt} \times \tau_{it}) + X_{rt}{}'\rho + \phi_{it} + \eta_r + \sigma W \ln p_{icrt} + \varepsilon_{irt} \tag{2.17}$$

where W denotes a spatial weight matrix of which each element captures the inverse distance between a pair of provinces. The assumption behind the approach is that province-pairs with short distances are closely interconnected in terms of consumer price determinations, whereas the effect of remote provinces is not big that much.[30]

We construct a block-diagonal matrix for W. Each block, w, represents a $P \times P$ matrix for each sector-year group, where P is the number of provinces. The $P \times P$ matrix is defined as

$$w = \begin{bmatrix} 0 & {d_{12}}^{-1} & {d_{13}}^{-1} & \dots & {d_{1P}}^{-1} \\ {d_{21}}^{-1} & 0 & {d_{23}}^{-1} & \dots & {d_{2P}}^{-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ {d_{P1}}^{-1} & {d_{P2}}^{-1} & {d_{P3}}^{-1} & \dots & 0 \end{bmatrix}$$

where every off-diagonal element is an inverse of the distance from a particular province p to another province $p^{'}$. We normalize the matrix w by min-max normalizing procedure.[31] The full weight matrix W is an $SPT \times SPT$ matrix where S is total number of sectors and T is total number of years. The full weight matrix

[30] Another way to account for interactions across province-level CPI is spatial error model(SEM), where the structure of spatial interaction term is put on the errors. We choose SAR model rather than SEM, because we already make use of standard errors clustered at the sector-year level throughout this chapter. We expect that the clustered standard errors will greatly alleviate the concern about cross dependences across regions in the *error* term.

[31] The min-max normalized matrix is as the following: the $\{p, p^{'}\}$th element of the normalized matrix becomes the value of the original element divided by m, where $m = \min\{$the largest row sum of pre-normalized matrix, the largest column sum of pre-normalized matrix$\}$.

which is a block-diagonal matrix given by

$$W = \begin{bmatrix} w & 0 & 0 & 0 \\ 0 & w & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & w \end{bmatrix}$$

Table 2.13 reports estimation results of Equation (2.17). Even incorporating spatial autoregressive (SAR) model, the results show simliar pattern with those in Table 2.11 where SAR model is not included. In particular, IV results in Table 2.13 suggest that increase in retail markups by 10 percentage points causes 1.10-1.12 percent lower tariff pass-through. The negative effect of retail markups on tariffs slightly declines compared to the regression without SAR model, which indicates that cross-section dependence exists spatially. Nevertheless, the SAR model makes sure that even controlling for the cross-section yields strong evidence that retail markups matter for forming imperfect tariff pass-through.

2.6 Conclusion

We provide evidence on the relationship between tariffs and consumer prices. We show that consumer prices are incompletely responding to tariff reductions. More importantly, We find that the imperfect tariff pass-through is crucially driven by the markups charged by retail intermediaries. While the retail markup channel is investigated by a few exchange rate pass-through literature, it has rarely been studied in the context of trade reforms. According to Bernard et al. (2010), intermediary firms "affect the magnitude and nature of trade frictions and both the pattern of trade and its welfare gains". The findings in this chapter strongly suggest that retail intermediaries play a crucial role in determining how much degree consumers benefit from trade liberalization. Again, this chapter relates to trade induced macroeconomic issues such as welfare analysis, inequality and poverty, which can be explored by future research.

Table 2.1: Average Agricultural Tariff Rates

Year	1989	1990	1992	1995	1996	1997	2000	2005	2010
Tariff (%)	16.71	22.39	23.09	21.98	69.88	73.35	64.72	50.72	53.22

Notes: This table shows average agricultural tariff rates over years. To calculate them, we use UNCTAD-TRAINS and WTO-IDB data. As noted in section 2.2.2, there are missing values in tariff rates for the years of 1991, 1993 and 1994, even in the combination of two sources of data. That is the reason why this table is not including values for the three years.

Table 2.2: Target Average Tariff Rates

Year	1988	1989	1990	1991	1992	1993	1994
Tariff (%)	20.0	15.0	13.0	12.0	11.0	9.0	8.0

Source: Cheung and Hongsik (2010)

Notes: This table presents the average target tariff rates set by the Korean government. Korea carried out tariff reductions in a manner that it followed predetermined reduction plan, and as shown in this table, the target rates are clearly monotonically decreasing trend.

Table 2.3: Descriptive Statistics

Variable Name	N	Mean	St. Dev.	Min	Max
Log CPI ($CPI_{2010} = 100$)	9,750	4.500	0.531	2.670	6.881
Weighted Average Tariff Rates (%)	9,984	7.975	4.1171	0.000	23.201
Herfindahl Index	6,630	0.031	0.028	0.001	0.182
Log World Price ($)	9,152	4.418	2.967	0.000	10.724
Import Penetration Ratio (%)	9,568	0.021	0.038	0.000	0.244
Log imports from China to Korea	9,616	11.322	2.294	1.971	15.597
Log GRDP (Million KRW)	10,114	17.415	0.848	15.050	19.534
Log Population Density (per km^2)	10,140	6.619	1.545	4.489	9.759

Notes: This table shows summary statistics for the main variables used in the empirical analyses of Chapter 1.

Table 2.4: Sector-level Log Consumer Price Indices over Years

	1990		1995		2000		2005		2012	
Sector Name	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Books	3.71	0.03	3.97	0.01	4.28	0.01	4.44	0.00	4.67	0.00
Casual Clothes	3.96	0.08	4.20	0.00	4.27	0.00	4.41	0.00	4.69	0.00
Children's Clothes	4.02	0.16	4.24	0.04	4.37	0.00	4.55	0.00	4.71	0.00
Electric Applicances for Personal Cares	3.59	0.08	4.06	0.09	4.35	0.05	4.45	0.03	4.71	0.02
Equipment for Sound and Pitures	5.95	0.07	5.83	0.07	5.58	0.05	5.31	0.03	4.36	0.02
Fuels and Lubricants for transport Equipment	3.03	0.03	3.44	0.02	4.20	0.02	4.37	0.01	4.76	0.01
Furnitures, Furnishing and Carpets	3.88	0.06	4.10	0.05	4.17	0.03	4.38	0.00	4.68	0.00
Glassware, Tableware and Household utensils	3.97	0.07	4.24	0.06	4.36	0.04	4.46	0.02	4.72	0.02
Household Appliances	4.93	0.07	4.87	0.06	4.73	0.05	4.62	0.03	4.61	0.01
Household Textile	4.24	0.12	4.43	0.07	4.53	0.04	4.43	0.06	4.70	0.01
Information Processing Equipments	6.43	0.12	6.08	0.06	5.65	0.08	5.20	0.02	4.46	0.01
Medical Products, appliances and Equipments	4.20	0.09	4.42	0.05	4.48	0.05	4.57	0.02	4.64	0.01
Men's Clothes	4.17	0.04	4.29	0.02	4.38	0.01	4.42	0.01	4.70	0.01
Non-Durable Houshold Goods	3.94	0.08	4.10	0.06	4.32	0.06	4.44	0.04	4.70	0.03
Other Articles of Clothing and Accessories	4.25	0.06	4.36	0.05	4.49	0.02	4.52	0.02	4.67	0.01
Othe Fuels and Energy	2.90	0.10	3.14	0.07	3.92	0.07	4.37	0.05	4.82	0.04
Othe Major Durables for Recreation and Culture	3.92	0.06	4.20	0.04	4.37	0.04	4.49	0.04	4.65	0.02
Other Recreational Items, Gardens and Pets	4.09	0.06	4.30	0.06	4.45	0.03	4.48	0.02	4.65	0.01
Personal Effects N.E.C.	4.28	0.06	4.23	0.06	4.37	0.04	4.42	0.01	4.64	0.00
Pharmaceutical Products	4.33	0.33	4.50	0.04	4.47	0.02	4.57	0.01	4.61	0.02
Purchase of Vehicles	4.53	0.01	4.55	0.00	4.53	0.00	4.54	0.00	4.61	0.00
Shoes and Other Footwear	3.94	0.05	4.05	0.02	4.28	0.01	4.51	0.01	4.69	0.01
Stationary and Drawing Materials	3.82	0.09	4.15	0.06	4.41	0.04	4.39	0.02	4.72	0.02
Telephone and Telefax Equipments	6.54	0.12	6.50	0.09	6.07	0.07	5.20	0.01	4.52	0.01
Tools and Equipment for House and Garden	4.63	0.16	4.63	0.11	4.53	0.07	4.49	0.05	4.65	0.02
Women's Clothes	4.05	0.02	4.30	0.01	4.36	0.01	4.44	0.01	4.67	0.00

Notes: This table reports sector-level log CPI over Years. A sector-level log CPI is calculated by averaging out all provincial log CPI within each sector. A sector-level standard deviation of log CPI measures the degree to which province-level indices are dispersed within each sector.

Table 2.5: Province-level Log Consumer Price Indices over Years

	1990		1995		2000		2005		2012	
Province Name	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Cheju	4.30	0.88	4.44	0.74	4.54	0.52	4.55	0.27	4.65	0.10
Chungbuk	4.31	0.86	4.45	0.74	4.54	0.50	4.55	0.26	4.65	0.09
Chungnam	4.26	0.85	4.44	0.72	4.55	0.51	4.56	0.26	4.65	0.09
Deagion	4.25	0.88	4.45	0.72	4.56	0.49	4.56	0.26	4.66	0.09
Deagu	4.33	0.92	4.45	0.74	4.54	0.49	4.57	0.27	4.65	0.09
Gangwon	4.28	0.85	4.43	0.72	4.53	0.48	4.55	0.27	4.65	0.09
Gwangju	4.28	0.85	4.41	0.69	4.52	0.45	4.55	0.26	4.65	0.09
Incheon	4.27	0.84	4.43	0.71	4.53	0.48	4.56	0.26	4.65	0.09
Junbuk	4.26	0.88	4.42	0.71	4.53	0.47	4.56	0.26	4.66	0.10
Junnam	4.28	0.86	4.42	0.72	4.54	0.49	4.55	0.26	4.66	0.10
Kyungbuk	4.29	0.85	4.43	0.71	4.53	0.48	4.55	0.26	4.66	0.09
Kyunggi	4.26	0.86	4.42	0.72	4.53	0.49	4.55	0.26	4.65	0.09
Kyungnam	4.28	0.85	4.42	0.72	4.53	0.48	4.54	0.26	4.65	0.10
Pusan	4.26	0.84	4.42	0.71	4.53	0.46	4.55	0.25	4.66	0.09
Seoul	4.30	0.88	4.43	0.74	4.54	0.49	4.54	0.27	4.65	0.09
Woolsan			4.42	0.72	4.54	0.48	4.55	0.26	4.66	0.09

Notes: This table reports province-level log CPI over Years. Each province-level index is the avarage of sectoral log consumer price indices within the province. Standard deviations present degree of variations in log CPI at the province level.

Table 2.6: Tariff Rates and Consumer Price Index

	Dependent variable: Log CPI				
	(1)	(2)	(3)	(4)	(5)
Tariff Rates	-0.000962	-0.000621	-0.000181	0.0656***	0.0599***
	(0.00501)	(0.00782)	(0.00783)	(0.00996)	(0.00965)
Import Penetration Ratio					3.252***
					(0.596)
Log World Price					0.0775***
					(0.0144)
Log Imports from China					-0.153***
					(0.0195)
Sector FE	No	Yes	Yes	Yes	Yes
Province FE	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes
N	9,594	9,594	9,594	9,594	8,716
R^2	0.000	0.612	0.613	0.689	0.745
Clusters	624	624	624	624	551

Notes: This table demonstrates the association between tariffs and consumer prices. *Import Penetration Ratio* measures sector-level degrees of import penetration defined as a ratio between imports and imports plus domestic productions at the sector level. (Log) *World Price* varies across sectors and years and it is proxied by U.S.'s average export prices of products within each sector-year cluster. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the sector-by-year level.

Table 2.7: Assessment of Bias Due to Unobserved Confounding Effects

	Regression Results Without Controls (1)	Regression Results With Observables (2)
Tariff Rates	-0.000962 (β^1)	0.0599 (β^2)
R^2	0.0001 ($R^2{}_1$)	0.7453 ($R^2{}_2$)
Bias due to Unobserved Confouding Effects	-0.0208	

Notes: This table shows a result of sensitivity analysis based on Oster (2017)'s approach. Column (1) reports the regression results when we regress log CPI on tariffs without any control variables. On the other hand, column (2) displays regression results when using observable controls. The observables exploited for column (2) are sector fixed effect, province fixed effect, year fixed effect and three sector-year specific control variables described in section 2.3.1. We present the bias driven by unobserved confounding effects at the bottom of the table. It is calculated by the formula below,

$$\text{Confounding Effect} = \delta \frac{(\beta^1 - \beta^2)(R^2{}_{max} - R^2{}_2)}{(R^2{}_2 - R^2{}_1)}$$

To calculate it, we use regression results reported in columns (1)-(2). As noted in section 2.3.1, δ and $R^2{}_{max}$ are all assumed to be one, because the value is the hardest case to happen.

Table 2.8: Log(1+Tariff) and CPI

	Dependent variable: Log CPI				
	(1)	(2)	(3)	(4)	(5)
Tariff Rates	0.0175	-0.0398	-0.0377	0.217***	0.212***
	(0.0237)	(0.0624)	(0.0625)	(0.0728)	(0.0653)
Import Penetration Ratio					3.101***
					(0.625)
Log World Price					0.0646***
					(0.0134)
Log Imports from China					-0.170***
					(0.0204)
Sector FE	No	Yes	Yes	Yes	Yes
Province FE	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes
N	9,594	9,594	9,594	9,594	8,716
R^2	0.001	0.613	0.614	0.663	0.728
Clusters	624	624	624	624	551

Notes: This table reports the estimation results of Equation (2.2) where $\ln(1+\tau_{it})$ is used as an independent variable instead of τ_{it}. Results are almost identical to those in the original model (Table 2.6). *Import Penetration Ratio* measures sector-level degrees of import penetration defined as ratios between imports and imports plus domestic productions. (Log) *World Price* varies across sectors and years and it is proxied by U.S.'s average export prices across products within each sector-year cluster. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the sector-by-year level.

Table 2.9: Retail Market Concentration and Tariff Pass-Through

	Dependent variable: Log CPI		
	(1)	(2)	(3)
Herfindahl Index	0.454	0.448	0.153*
	(0.601)	(0.601)	(0.0789)
Tariff Rates	0.0521***	0.0521***	
	(0.0167)	(0.0167)	
Herfindahl Index × Tariff Rates	-0.0588	-0.0588	-0.0180**
	(0.0818)	(0.0818)	(0.00715)
Sector-Year Observables	Yes	Yes	No
Province-Year Observables	No	Yes	Yes
Sector FE	Yes	Yes	No
Province FE	Yes	Yes	Yes
Year FE	Yes	Yes	No
Sector-Year FE	No	No	Yes
N	6,390	6,390	6,390
R^2	0.662	0.662	0.993
Clusters	401	401	401

Notes: This table presents that increase in retail market concentration measured by the Herfindahl index lowers tariff pass-through to consumer prices. In particular, the regression result shown by column (3), where endogeneity embodied in tariffs is controlled by *sector-year fixed effects* provide evidence of the negative causal effect of retail market concentrations on the pass-through. Sector-year specific observable controls used for the regressions are import penetration ratio, (log) world prices and (log) imports from China. Province-year observables are (log) population density and (log) gross regional domestic product. * p-value $<$ 0.10, ** p-value $<$ 0.05, *** p-value $<$ 0.01. Standard errors in parentheses are clustered at the sector-by-year level.

Table 2.10: Ten-Year Lagged Populations and Consumer Prices

	Dependent variable: Log CPI	
	(1)	(2)
Population_{t-10}	0.000287	-0.000302
	(0.000954)	(0.00119)
Province-Year Observables	No	Yes
Sector FE	Yes	Yes
Province FE	Yes	Yes
Year FE	Yes	Yes
N	9,750	9,646
R^2	0.653	0.650
Clusters	650	650

Notes: This table demonstrates that ten-year lagged population(Population_{t-10}) is not associated with current consumer prices. The results are robust whether province-year specific observables (log population density and log gross regional domestic product) are used (column (2)) or not (column (1)). * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the sector-by-year level.

Table 2.11: Retail Markup Channel

	Dependent variable: Log CPI			
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Retail Markup	0.161	1.985		
	(0.115)	(1.405)		
Retail Markup × Tariff	-0.0367**	-0.173*	-0.0281**	-0.196**
	(0.0142)	(0.0954)	(0.0122)	(0.0966)
Province-Year Observables	Yes	Yes	No	No
Sector-Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	No	No
Province-Year FE	No	No	Yes	Yes
N	5,550	4,800	5,500	4,800
R^2	0.992	0.992	0.992	0.992
Clusters	350	300	350	300
Kleibergen-Paap rk Wald F		41.685		313.304

Notes: This table illustrates that increase in provincial retail markups lead to fall in tariff pass-through to consumer prices. Column (1) and (3) are OLS results without IVs, whereas Column (2) and (4) shows estimation results of (the second stage) IV regressions. Kleibergen-Paap rk Wald F statistics present the results of weak IV test in the case where more than two IVs and clustered standard errors are exploited. In column (2) and (4), the statistics all exceed critical values proposed by Stock and Yogo (2005), which implies that there is no weak IV problem in the IV results. Province-year observables are (log) population density and (log) gross regional domestic product. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the sector-by-year level.

Table 2.12: Retail Markup Channel, *Analysed by Census Data*

	Weighted Average PCM				*Weighted Average PCM +Controlling for County Heterogeneity*			
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)
Retail Markup	0.136	0.553			0.0855	0.728*		
	(0.131)	(0.360)			(0.0521)	(0.390)		
Retail Markup × Tariff	-0.0171	-0.0643*	-0.0175	-0.0693*	-0.0062	-0.0853*	-0.0052	-0.0892*
	(0.0134)	(0.0363)	(0.0135)	(0.0365)	(0.0062)	(0.0474)	(0.0060)	(0.0478)
Province-Year Observables	Yes	Yes	No	No	Yes	Yes	No	No
Sector-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	No	No	Yes	Yes	No	No
Province-Year FE	No	No	Yes	Yes	No	No	Yes	Yes
N	1,590	1,590	1,590	1,590	1,590	1,590	1,590	1,590
R^2	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994
Clusters	101	101	101	101	101	101	101	101
Kleibergen-Paap rk Wald F		50.698		490.822		284.164		3012.51

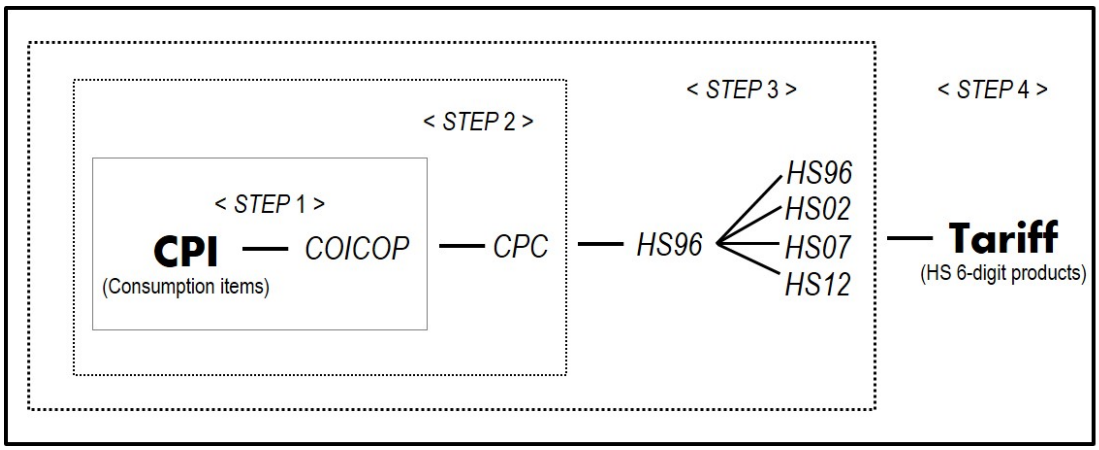
Notes: This table illustrates the IV results which show that increase in provincial retail markups lead to fall in tariff pass-through to consumer prices using the retail census data. Columns (1)-(4) show estimation results when we use weightd average of store-level PCMs ($\mu_{rt}^{\texttt{Census1}}$). On the other hand, columns (5)-(8) report estimation results when, controlling for county-level market heterogeneity, we use weighted average of store-level PCMs ($\mu_{rt}^{\texttt{Census2}}$). Column (1), (3), (5) and (7) are OLS results without IVs, whereas Column (2), (4), (6) and (8) show estimation results of (the second stage) IV regressions. Kleibergen-Paap rk Wald F statistics presents the results of weak IV test in the case where more than two IVs and clustered standard errors are exploited. In column (2), (4), (6) and (8), the statistics all exceed critical values proposed by Stock and Yogo (2005), which implies that there is no weak IV problem in the IV results. Province-year observables are (log) population density and (log) gross regional domestic product. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the sector-by-year level.

Table 2.13: Retail Markup Channel, *Spatial Autoregressive (SAR) Model*

	Dependent variable: Log CPI			
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Retail Markup	0.124	0.773***		
	(0.0938)	(0.180)		
Retail Markup × Tariff	-0.0238**	-0.112***	-0.0202*	-0.110***
	(0.0114)	(0.0218)	(0.0114)	(0.0216)
Spatial Lag Term	Yes	Yes	Yes	Yes
Province-Year Observables	Yes	Yes	No	No
Sector-Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	No	No
Province-Year FE	No	No	Yes	Yes
N	4,992	4,992	4,992	4,992

Notes: This table evaluates whether accounting for spatially cross-section dependence would lead to different estimation results of retail markup channel regression model (Equation (2.12) and (2.14)). We find that all the results are similar to Table 2.11, indicating that even controlling for the cross section dependence yields clear evidence that retail markups play a role in decrease in tariff pass-through. The technique by which the issue of cross-section dependence is dealt with is Spatial Autoregressive (SAR) model. Column (1) and (3) are OLS results without IVs, whereas Column (2) and (4) show estimation results of (the second stage) IV regressions. Province-year observables are (log) population density and (log) gross regional domestic product. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the sector-by-year level.

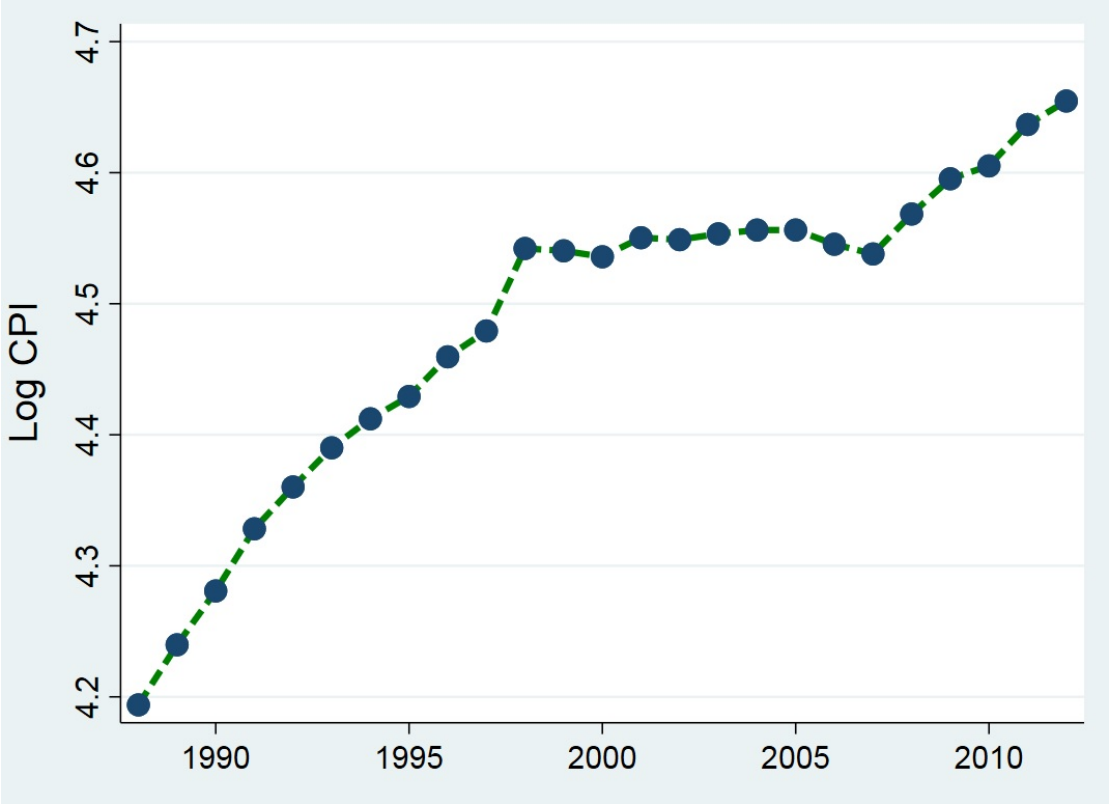
Figure 2.1: Matching procedure between CPI and Tariffs



Note A: This figure illustrates how we link consumption-item level CPI data to product-level tariff data, which consists of four steps. First, we relate CPI to COICOP (step 1). Then, we merge the matched data with CPC via UNSD's concordance (step 2). We further merge the matched CPC from step 2 to HS codes using both the concordance between CPC and HS nomenclature 1996 and the concordance among different versions of HS codes (step 3). Finally, we construct weighted average tariff rates of all matched products within each CPI sector.

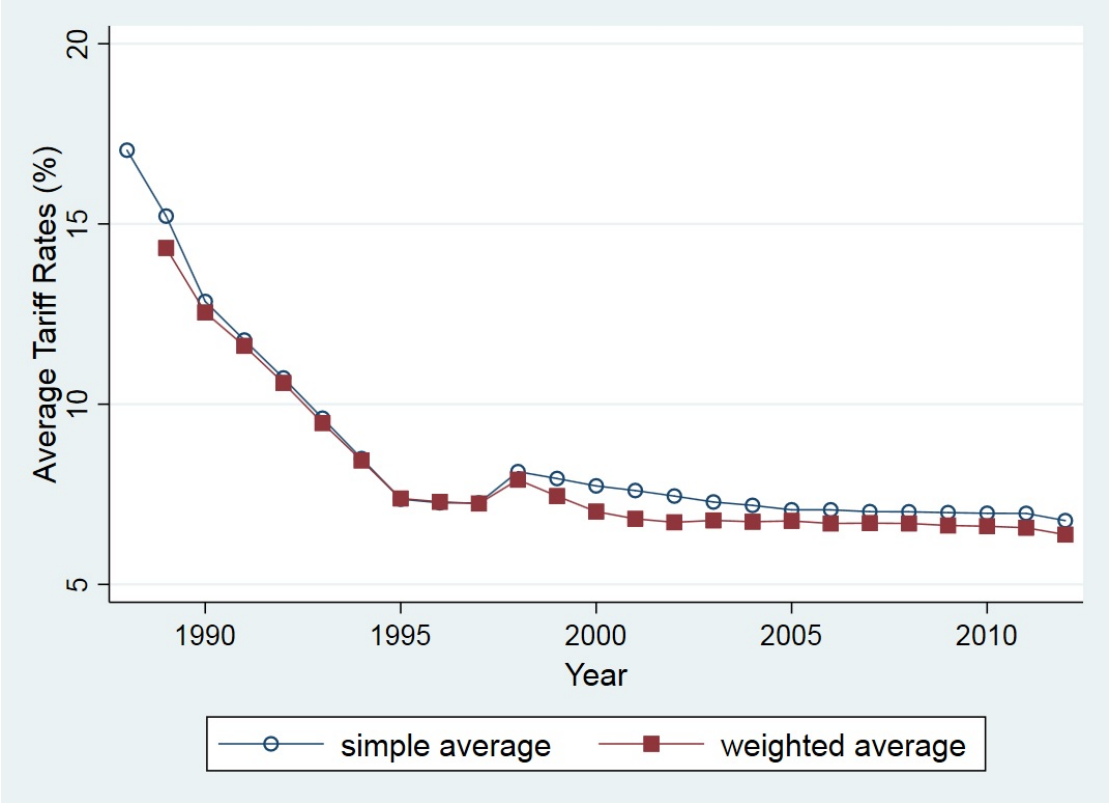
Note B: COICOP is a reference classification for consumption items. CPC is a product classification based on physical characteristics. All the concordances can be found in UNSD(United Nation Statistics Division) classification registry (http://unstats.un.org/unsd/cr/registry).

Figure 2.2: Time Trend in Log Consumer Price Index



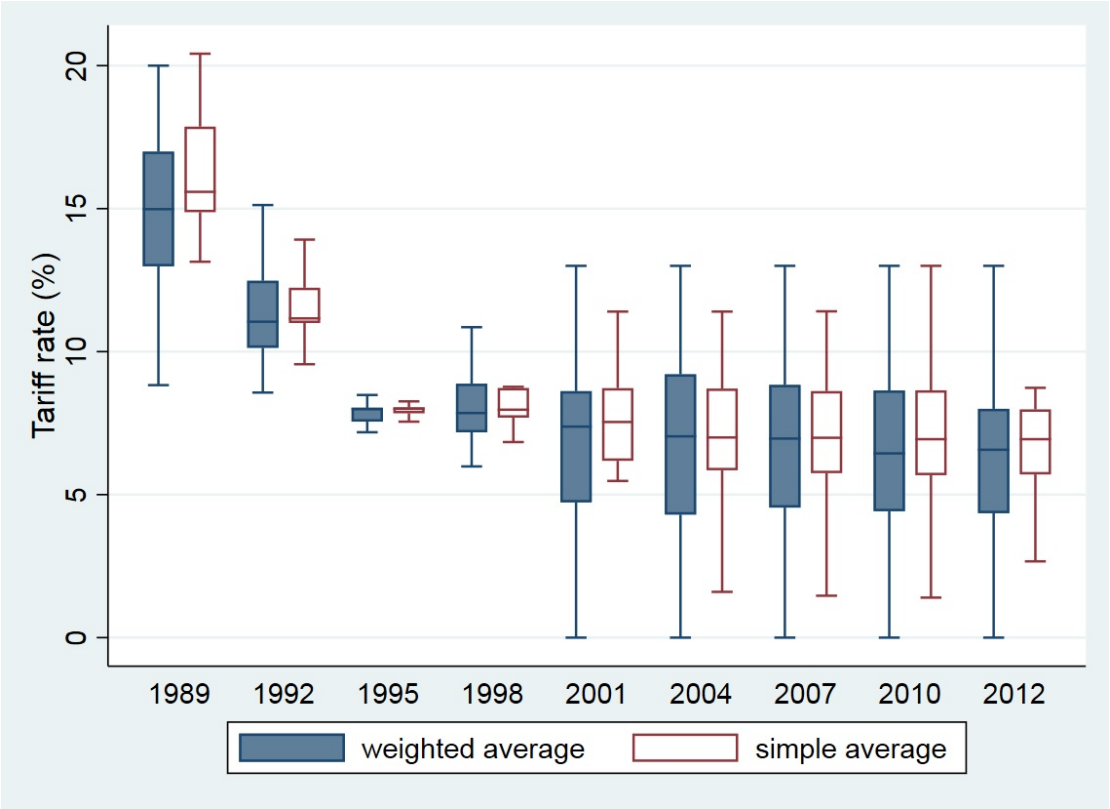
Notes: This figure presents a time trend in the logarithm of average consumer price index with a base year 2010.

Figure 2.3: Time trend in Average Tariff Rates



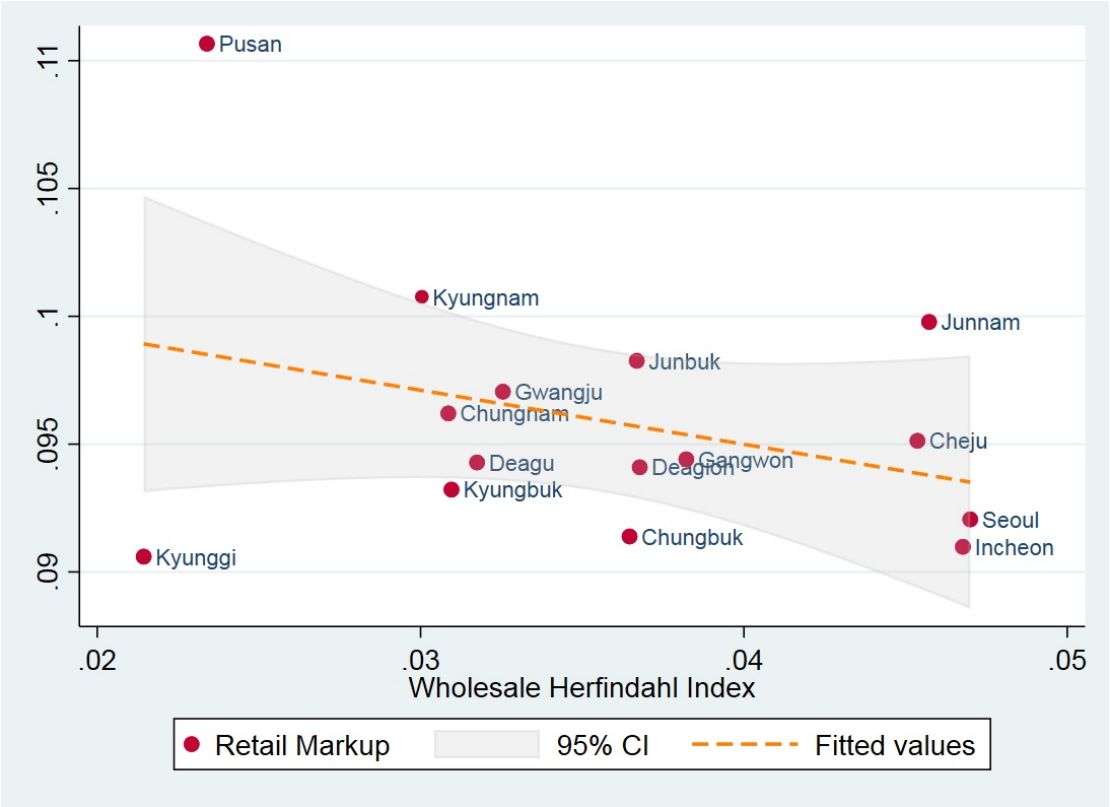
Notes: This figure illustrates the evolutions of Korea's average tariff rates from 1998 to 2012. *Simple average* shows a trajectory of mean values of sectoral (simple) average tariff rates. *Weighted average* describes a trajectory of mean values of sectoral weighted averages. Weights used to measure sectoral weighted averages are one-year lagged import shares in total values.

Figure 2.4: Tariff Dispersions across Sectors



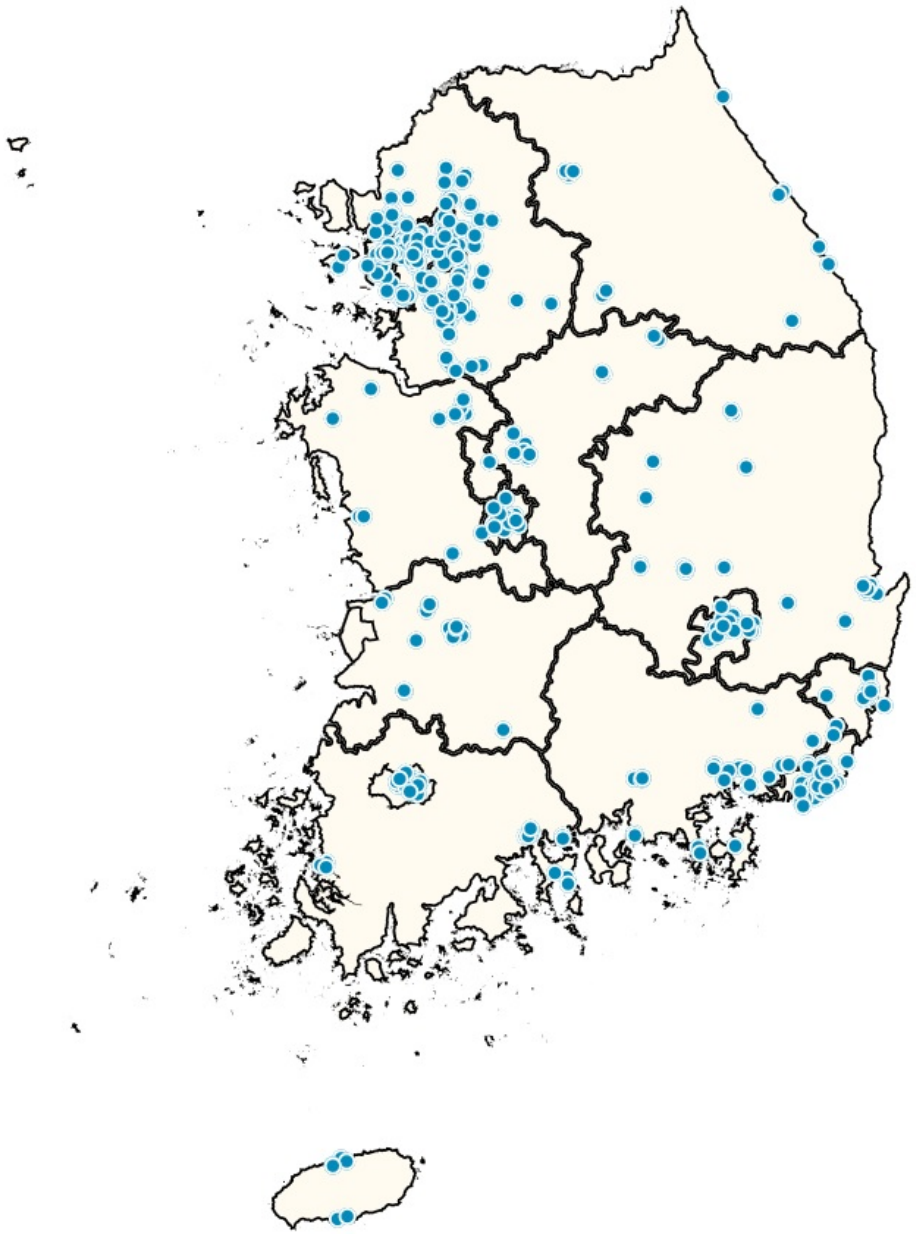
Notes: This figure describes how sectoral dispersions in tariff rates evolve during the period of 1989 to 2012.

Figure 2.5: Association Between Retail Markups and Wholesale Herfindahl Index



Notes: This figure shows the correlation between the average retail markup and the average Herfindahl index of the wholesale sector across years. It demonstrates that the two variables are negatively associated. With regard to the Wholesale Herfindahl index, we use the empirical index obtained by regressing the analytic Herfindahl index on province- and year fixed effects. It aims both to account for heterogeneities and to reflect the modified index which eventually fits for the first stage IV regression model including the fixed effect terms.

Figure 2.6: Large Discount Stores in Korea (2010)



Notes: This figure shows the spatial distribution of large discount stores in Korea. Large discount stores are a fairly new retail store format that has proliferated since the mid-1990s in Korea. The spatial distribution of the particular store format illustrates well how locations of retail stores relate to populations.

Chapter 1 Appendix Tables

Table A1. Retail Markup Channel, *First Stage IV results*

Dependent Variable:	*Retail Markup* (1)	*Retail Markup* × *Tariff* (2)
Population_{t-10}	0.000257	0.00430
	(0.000398)	(0.00293)
Wholesale Herhindahl Index	-0.0284***	0.141***
	(0.00486)	(0.038)
$\text{Population}_{t-10} \times$ Tariff	0.00000718	-0.000244***
	(0.0000108)	(0.0000931)
Wholesale Herhindahl Index × Tariff	-0.000269	-0.0518***
	(0.000542)	(0.00411)
N	4,800	4,800
R^2	0.670	0.972
Clusters	300	300
F-statistics	267.25	239.08

Notes: This table presents estimation results of the first stage IV regression of Equation (2.12). As mentioned in the main text, there are two endogenous variables: retail markup (μ_{rt}) and retail markup × Tariff ($\mu_{rt} \times \tau_{it}$). Each column shows the result with respect to each endogenous variable. Population_{t-10} is ten-year lagged provincial populations. Although we do not report here, the fist stage IV regressions include sector-year fixed effect(ϕ_{it}) and province-year specific observed controls ((log) population density and (log) gross regional domestic product). * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the sector-by-year level.

Table A2. Retail Markup Channel, *Analysed by Census Data*

	Weighted Average PCM + Controlling for Town Heterogeneity			
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Retail Markup	0.166	6.133*		
	(0.192)	(3.503)		
Retail Markup × Tariff	-0.00090	-0.113*	-0.0052	-0.0892*
	(0.0051)	(0.0680)	(0.0060)	(0.0478)
Province-Year Observables	Yes	Yes	No	No
Sector-Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	No	No
Province-Year FE	No	No	Yes	Yes
N	1,590	1,590	1,590	1,590
R^2	0.994	0.994	0.994	0.994
Clusters	101	101	101	101
Kleibergen-Paap rk Wald F		3.416		3012.51

Notes: This table suggests that retail market powers decrease tariff pass-through when we use retail markup measure constructed by census data. The measure is weighted average price-cost margins, and we construct them with *town* level market heterogeneity controlled for. More specifically, we apply the same method to establish $\mu_{rt}^{\texttt{Census}2}$ as described in section 2.5.1, but the only difference is that we use town fixed effect to account for town-level heterogeneity. Column (1) and (3) are OLS results without IVs, whereas column (2) and (4) show estimation results of (the second stage) IV regressions. Kleibergen-Paap rk Wald F statistics presents the results of weak IV test in the case where more than two IVs and clustered standard errors are exploited. While in column (4) the statistics exceed critical values proposed by Stock and Yogo (2005), in column (2) the F statistics do not pass. Thus, unlike column (4), the IV results reported in column (2) have potentially the weak instrument variable problem. Province-year observables are (log) population density and (log) gross regional domestic product. * p-value $<$ 0.10, ** p-value $<$ 0.05, *** p-value $<$ 0.01. Standard errors in parentheses are clustered at the sector-by-year level.

Chapter 3

Foreign Retail Entry Under Domestic Rivals

3.1 Introduction

Foreign direct investment (FDI) literature has rarely paid attention to strategic aspects of market entry. In particular, the actions of *domestic* rivals within host-country markets are completely neglected. The disregard for domestic rivalry stems from a pervasive assumption on host-country market structures. In standard FDI theories, industry within which FDI occurs is assumed to be inherently perfectly competitive. Yet, we have seen various oligopolistic industries where multinational enterprises are major players all over the world.[1] Under such imperfectly competitive markets, even multinational enterprises cannot act in a vacuum. They may deeply care about entry choices of domestic competitors, when they decide whether or not to enter such host-country markets. This chapter aims to study the impact of domestic rivalry on FDI's entry.

This study relies on an empirical context of Korea, a country where market structure of the discount retailing industry is highly concentrated. Specifically, there are two dominant retail firms in Korea. They account for 73.1 % of total sales in the industry. More importantly, one is a Korean domestic retailer, Emart, and the other is a multinational retailer, Tesco. As the first and the second biggest retail firms, they compete with each other by operating multiple large discount stores.[2] In Korea, Tesco is the de facto only nation-wide foreign retail chains. Therefore, we consider Tesco's large discount stores as the only foreign owned stores and regard stores operated by Emart as Tesco's domestic competitors. We use the context of Korean retail sector to shed light on strategic market entries between two retailers.

[1]For instance, a Japan based multinational car maker Toyota is the third largest firm (market share(2014): 14%) in the U.S. car industry. In the Indian Mobile Network industry, the second largest firm is a U.K. based company Vodafone (market share(2016): 23%).

[2]Large discount stores are a fairly new store format which has dominated not only the discount retailing industry but also the Korea's entire retail sector since the mid-1990s. We elaborate on what large discount store are in section 3.2.1.

The analysis exploits data on large discount stores in Korea, obtained from the Korea Chain-store Association (KOCA). From KOCA data, we extract information about a status of foreign (Tesco) retail entry and its domestic rival's (Emart) entry across counties. In addition, this chapter makes use of county-level market characteristics. Although we use demographical data (e.g. population, average ages, etc.) as covariates in the regression, it is still doubtful that they are not sufficient enough to control for unobserved heterogeneity across markets. To lessen the concern, this chapter takes advantage of establishment-level retail census data, and constructs a variety of county-specific variables to capture retail sector characteristics (e.g., price-cost margin, labour productivity, net entry rate, etc.).

The analysis proceeds in two parts. In the first part, we estimate reduced-form effects of domestic competitor's presence on foreign retailer's market entry. We do not only adopt the linear probability model but also apply binary choice model. A central econometric concern when studying the impact of the domestic presence on foreign entry is that domestic retailer's market presence is an endogenous variable for two reasons. First, it is highly likely that there are unobserved market characteristics that affect both the entry of foreign stores and that of domestic stores. Although we make use of fixed effects (linear probability model) and random effects (binary choice model), there may still be unobserved time-varying confounding factors. Taking advantage of a richness of Korea's retail census data, we use retail market characteristics across counties as control variables to mitigate the simultaneity issue.

Despite a variety of covariates, the adverse causality problem would not be able to be fully taken away. In particular, if two retail firms make strategic entry choices as we hypothesise, the reverse causality renders estimation results incorrect. To avoid the issue, we employ the instrumental variable (IV) approach. The instrument is distance from each Emart store to the closest Emart's distribution centre. Retailers who operate large discount stores take into account the location of distribution centres to take advantage of economies of density in distributions.(Holmes, 2011) In this regard, the IV would affect Emart store's market entry. The validity of the IV relies on the disconnection between the proposed instrument and the error term in the regression. The IV is not correlated with any factors, other than the endogenous variable, which would determine entry of foreign stores. It is because, considering that profit-making activities are different between stores and distribution centres (i.e., sales vs. distribution), location of Emart's distribution centres should be determined independently of entry determinants of Emart's stores. Furthermore, a number of market-specific controls contribute to facilitating the IV's disconnection to error terms.

In principle, the instrumental variable strategy rests on the linearity of model specification. For the reason, the literature adopts alternative ways (e.g., Two-step GMM) in binary choice models with endogenous variables. Even the alternative methods are, though, unsuitable for the binary choice model with *dummy* endogenous variables (Greene, 2007). Instead, we apply the bivariate probit model to the regression. It has been exploited from the literature to find consistent estimates in cases where the dependent variable and endogenous variables are all binary variables. Using the IV strategy, we find that the presence of domestic competitor, Emart, has strong and significant negative effects on the entry decision of foreign retailer, Tesco. Based on the results of the bivariate probit model, the average predicted probability of Tesco's entry is lower by 10.4% in counties with domestic rival's stores.

At the second part of the analysis lies the structural model estimation. Reduced-form regression clearly provides evidence that there is a negative effect of domestic rival's presence on foreign retail entry. However, it does not by itself take into account firm's strategic interactions incorporated in their entry choice model. To shed light on the strategic aspects of market entry, we turn to investigating the question through the lens of empirical game-theoretical models. We construct a static entry game model between foreign and domestic retailers, following the framework developed by Ciliberto and Tamer (2009) and Ellickson and Misra (2011). The model in the analysis is a game of incomplete information. It aims to reflect the wisdom of FDI literature: asymmetric information between foreign and domestic firms. Multinational enterprises (MNEs) tend to have superior technology and managerial know-how to domestic firms, whereas domestic firms are able to obtain local demand information better than MNEs. We incorporate this aspect of asymmetric information into the model by building an entry game model of incomplete information. Accordingly, the equilibrium concept is Bayesian Nash equilibrium. That is, in equilibrium, each retailer chooses an action that maximises expected profits given the belief about its competitor's action.

For the model estimation, we make use of two-step estimation approach which began with Hotz and Miller (1993). Using a non-parametric regression, we firstly estimate each firm's belief or ex-ante probability about competitor's action. Then we plug the estimated beliefs into the equilibrium conditions to obtain the fitted equilibrium probabilities. Building the log-likelihood function based on the fitted probability, we identify the structural parameters on the static entry game model. The structural model estimation clearly shows that Tesco and Emart make entry choices *strategically*. Specifically, the estimates for structural parameters on the rival's entry are negative and highly significant for both Tesco and Emart.

Furthermore, we estimate a game-theoretical model with different informa-

tional assumption. While the setting of the model is exactly same as before, the only difference is that, for now, the model is a game of *complete* information. Applying Berry's (1992) approach, we identify the structural parameters in the model. We find that a parameter indicating firms' strategic entry becomes either weakened or even disappeared, compared to incomplete information games. Also, given that the IV results may convey the true causality between Tesco's and Emart's market entries, what is corresponding to the reduced-form evidence is results of incomplete information game estimation. Consequently, asymmetric information across foreign and domestic firms plays a role to a considerable degree in forming retail FDI's strategic market entry into Korean market.

This chapter relates and contributes to the literature on empirical games of market entry. While models of market entry or location of retail stores originate from Hotelling (1929), structural model estimation began with the seminal work by Bresnahan and Reiss (1990). In particular, our study is in line with papers applying static game estimations with multi-store retailers such as Jia (2008), Holmes (2011) and Ellickson et al. (2013). This chapter differs from them in that we build an entry game-theoretical model in the context of competition between foreign and domestic firms. Based on that context, we apply a game model of incomplete information, which is also well explained by FDI literature. To our understanding, this is the first study to estimate a market entry game model under the competition between FDI and its domestic rivals.

On top of that, what makes this chapter differentiate from the existing researches is that we present estimation results of not only the structural model but also reduced-form regressions. Previously, literature on the retail market entry is likely to focus on either reduced-form approach (e.g. Basker and Noel (2009)) or structural estimation. Supported by evidence by two approaches, we further show the role of asymmetric information in determination of FDI's market entry. This is also a novel point of this study rarely explored in the literature.

This chapter is related to the international trade literature on entry or location decisions of multinational enterprises. Focusing on the manufacturing sector, the literature generally views motivations of MNEs as both serving the host-country markets and exporting to other markets outside of markets where they locate. (e.g. Markusen (1984), Helpman (1984), Ekholm et al. (2007), Ramondo and Rodríguez-Clare (2013), Tintelnot (2017), etc.) This chapter differs with the existing literature in three regards. First, while most of existing research has a strong market structure assumptions such as perfect or monopolistic competitions, we allow for strategic interactions between firms under oligopoly, which has impacts on entry/location decisions. Second, this chapter focuses on the service FDI. Despite the fact that half value of global green-field investments are ascribed to service

industries(UNCTAD, 2015), studies on service FDI is scanty. The analysis in this chapter contributes to the literature by exploring what factors affect service FDI's entry decisions. Third, existing papers analyse country-level entry/location decisions, but we study within-country regional factors affecting FDI's decisions.[3]

There are a very few number of papers which investigate strategic interaction of FDI in oligopolistic industries.(e.g. Knickerbocker (1973), Head et al. (2002), Hoenen and Hansen (2009)) Their view is that, in globally oligopolistic industries, firms tend to follow rivals into foreign markets as a result of strategic behaviours. Our work is related to the papers in that we study FDI's strategic market entry in imperfectly competitive markets. However, we consider oligopolistic regional markets within a country, whereas they study globally oligopolistic industries.

The remainder of this chapter proceeds as follows. In section 3.2, we describes the background of Korean retail sector. Especially, section 3.2.2 illustrates salient characteristics shown in spatial competition between retail FDI and its domestic competitor. Section 3.3 explains market definitions and data sets we make use of in this chapter. Section 3.4 empirically unveil the causal relationship between retail FDI's entry and domestic rival's presence. Section 3.5 outlines empirical model of market entry game and presents the structural estimation results. Section 3.6 concludes.

3.2 Background

3.2.1 Large Discount Store Market in Korea

Korea opened retail market to foreign firms in 1996. Since then, four multinational retailers – Wal-Mart, Carrefour, Tesco and Costco – have entered into the Korean market. The retail FDI liberalisation resulted in a striking phenomenon that has long dominated the Korean retail sector: the prevalence of large discount stores. The large discount store is a fairly new retail store format introduced in Korea in the mid-1990s. A typical store belonging to this format has 40,000 square meters of floor spaces, sells 50,000-100,000 product items, and hires 100-150 employees. The large discount store is analogous to a hypermarket or a superstore that sells general merchandises at low prices (Cho et al., 2015).

There are two reasons why FDI liberalisation could serve as a catalyst for the expansion of large discount stores in Korea. First, every foreign retailer that entered Korea has opened and operated only this particular retail store format. Taking advantage of economies of scale by multinational retailers, the retail format

[3]There are a few papers to study within-country regional factors affecting service FDI's entry such as Goldberg and Grosse (1994) and Jones and Wren (2016).

have already influenced on shaping the landscape of the retail sector across many countries. Naturally, large discount stores were transplanted into the Korean market as a business model by the multinationals. Second, the FDI liberalisation acted as a stimulus to Korean domestic retail firms. Before the mid-1990s, Korean retail sector had been dominated by small shops and local supermarkets (Cho et al., 2015). Being forced to face a new circumstance by FDI liberalisation, a few Korean conglomerates ('Chaebol'), which have run their businesses in the consumer goods sector or department stores, had to prepare for the competition with foreign retail firms. To gain a competitive edge, domestic incumbents actively introduced the new retail store format on their own and rapidly expanded large discount stores.

Table 3.1 displays the evolution of total sales in large discount stores. According to the Korea Fair Trade Commission(KFTC, 2009), sales revenue for large discount stores sharply increases from 5.0 trillion Korean won (KRW) in 1998 to 31.9 trillion KRW in 2008. In growth rate term, the average growth rate per annum (CAGR) is 18.3%, which is far faster than that of the entire retail sector, 8.5%. The share of large discount stores in total sales of the entire retail sector is also shown to be an increasing trend. In 1998, large discount stores account for only 5.7% whereas the share increases by three times (14.7%) in 2008. That is to say, large discount stores dominate the other sub-sectors (e.g., local supermarkets, convenience stores, department stores, etc.) in the retail sector. According to KOSTAT(Korea's office of national statistics), while large discount stores were ranked as the fourth largest sub-sector in 1998, they have ranked as the first since 2001.

As described so far, large discount stores are the largest and the fastest growing industry in the Korean retail sector. Another important feature in discount retailing industry is that there are two dominant firms in the industry. A Korean retail firm, Emart, is the biggest retailer who operate large discount stores in Korea. In 2010, Emart operates a total of 130 large discount stores. In terms of sales revenue, Emart's sales amount is 12.4 trillion KRW in 2008, which accounts for approximately 39% in total sales of large discount stores.[4] The second largest retailer who operates large discount stores is a UK based firm, Tesco. In 2010, Tesco operates 121 large discount stores in Korea. The sales amount is 10.9 trillion KRW in 2008, and thus the market share of Tesco is 34.2%. Two facts are noteworthy. First, it is clear that Emart and Tesco are in an intense competition to each other.

[4]To compute market shares, we use total sales of all Korean large discount stores (Table 3.1) reported by KFTC (2009), but we inevitably use each firm's sales revenue in 2008 reported in its income statement due to unavailability of proper data. Because of differences in data sources and years, there may be a slight error in calculating firm-level market shares. Nevertheless, it is certain that large discount retailing industry is concentrated by Emart and Tesco.

The little gap in both store numbers and sales amount mirrors this perspective. Second, large discount store market is to a high degree concentrated by two firms. The sum of market shares for them is 73.1%. Hence, it can be said that market structure of the sector of large discount stores is almost an oligopolistic market.

Among the four foreign firms that have entered Korea, Carrefour and Wal-Mart exited from Korea in 2006. The number of large discount stores was 30 for Carrefour and 14 for Wal-Mart at the end of 2005. Interestingly, Carrefour stores were all sold by Tesco and Wal-Mart stores were merged by Emart. That is, M&A relating to Carrefour and Wal-Mart contributed to increase in market concentrations towards Emart and Tesco. Unlike Carrefour and Wal-Mart, Costco is still operating 13 large discount stores in Korea. Since an absolute majority of Costco stores are located on a specific region, Seoul metropolitan area, Costco cannot be regarded as national chains. In this sense, Tesco is de facto the only nation-wide foreign retail chain in Korea. Therefore, it is clear that the market for large discount stores is dominated by one Korean firm, Emart, and one foreign firm, Tesco.

3.2.2 Illustrative Analyses on Retail Spatial Competition

According to Coe and Lee (2013), Tesco firstly prioritised opening stores in Pusan, Kyungnam and Kyungbuk provinces which are all located in the southeast of the Korean peninsula. Coe and Lee (2013) point out that this initial opening strategy of Tesco noticeably differs from that of Emart that focused on Seoul metropolitan Area in the early 2000s. In Figure 3.1, where we display spatial expansions of large discount stores by two firms, this pattern is well illustrated in Panel (a) and (b). In 2002, Tesco opened 47.6% of stores in Pusan, Kyungnam and Kyungbuk provinces, whereas Emart opened 48.0% of stores in Seoul metropolitan Area. In short, the early diffusion pattern of stores by both firms indicates that Tesco's market entry decision seems to be made by taking into account the entry made by Emart.

The strategic entry of Tesco against Emart continues over time as illustrated by Panels (c), (d), (e) and (f) in Figure 3.1. Since the mid of 2000s, Tesco started to increase stores in the Seoul metropolitan area. Yet, Tesco still avoids opening stores at counties within Seoul metropolitan area where Emart's stores are present. This county-level entry pattern is more evident when having a closer look at markets in a particular province, Seoul. In Panel (a) and (b) in Figure 3.2, there are three Tesco large discount stores at three counties in Seoul in 2004. Interestingly, those three counties are regions that Emart has never opened stores in before 2004. In other words, Tesco chose to open stores in counties without

Emart stores. The pattern is also found, to a certain extent, in the late 2000s as depicted by Panel (c), (d), (e) and (f) in Figure 3.2. Figure 3.3 shows the expansion of stores by two firms focusing on the Kyunggi province. Although many stores are spotted around the central region of Kyunggi province, there seems to be a pattern that Tesco tries to open stores in non-overlapping counties with Emart. For instance, among nine counties in the Kyunggi province where Tesco has entered up to 2006 (Panel (c) and (d) in Figure 3.3), five counties are non-overlapping markets with Emart stores.

In Figure 3.4, we display the evolution of Tesco's county-level market entries sorted by Emart's pre-occupations. Blue bars show how many Tesco's entries are made in counties that Emart stores have not entered in. Meanwhile, red bars present the number of Tesco's entries occured in markets that has experienced Emart's pre-occupations. As shown in the figure, the blue bars are mostly higher than red bars over years, indicating that Tesco opens stores taking care of its competitor's entry. Put differently, Figure 3.4 suggests that Tesco and Emart make a strategic entry in Korea's market. However, it is observed that the strategic entry becomes weaken at the end of 2000s. From 2009 to 2010, most Tesco's entries are appeared in markets in which Emart stores exist. Arguably, the weakening of Tesco's strategic market entry may be affected by the market saturation of the large discount store industry. Suh and Hahn (2004) estimate the acceptable maximum number of large discount stores in Korea, and show that the large discount store market would be saturated when the number of stores reaches to 400-450. Accordingly, it would be difficult to find non-overlapping markets as the number of large discount stores increases. The pattern in 2009 and 2010 may reflect this point of view.

The illustrative analyses in this subsection strongly indicate that a foreign retailer, Tesco, makes strategic entry decisions in Korea, deeply caring about the entry decisions of its competitor indigenous to the host country, Emart. In this chapter, we study the strategic aspect of FDI's entry using the context of Korea's discount retailing industry.

3.3 Data

A market definition crucially affects the modelling approach used for this study. In this chapter, a market is defined as a county. There are 251 counties in Korea. Table 2 shows a distribution of counties according to market entries by two firms. As shown in Table 2, a total of 140 counties are occupied by either Tesco's or Emart's large discount store in 2010. Neither Tesco nor Emart enter into 111 counties. The number of counties which both firms enter in is 44. On the other

hand, 43 and 53 counties are markets where only Tesco or only Emart store is present, respectively.

Table 2 has two noteworthy implications. First, both firms consistently expand large discount stores over years. As a result, more than half of counties are occupied by Tesco or Emart stores in 2010 as we noted above. Second, the distribution of store presences by two firms does not look skewed in favour of a particular firm. It indicates that two firms are clearly in a spatial competition in terms of market entry. Interestingly, there are few counties (23) which are occupied by more than two stores with the same firm affiliation. This fact suggests that a county is likely to a geographical unit on which two retail firms make store opening decision. In this chapter, we exploit county-level spatial variations in exposure to large discount stores by Tesco and Emart to study the impact of domestic competitor's entry on foreign retailer's market entry.

The primary data source is a dataset on large discount stores collected by the Korea Chainstore Association (KOCA). KOCA has annually published Yearbook of Retail Industry that includes detailed information on every large discount store. KOCA data encompasses store-level information such as names, firm affiliations, addresses and opening dates. Using KOCA data, we construct a county-level dummy variable to indicate whether each county is occupied by FDI (Tesco) stores. We also build a dummy variable to show whether its domestic competitor's (Emart's) store exists at each county.

This chapter also makes use of data on county-level market characteristics. We utilise demographical data obtained from Korea's population census. Some may argue that only demographics are not sufficient enough to control for market heterogeneity. Nevertheless, previous literature on empirical models of market entry generally uses only demographics due to lack of available data. This chapter, however, takes advantage of a new data source which has scarcely been used, to our knowledge, in the previous literature. We utilise store-level census data collected by KOSTAT, and using census data we construct retail market characteristics at the county level.

KOSTAT undertakes retail census every five year.[5] Retail census covers a population of entire Korean retail stores. With regards to the number of observations, it amounts to 600,000-770,000 per each census year. In fact, store-level census data had long been publicly inaccessible. KOSTAT had rather released province-level aggregated data for some selected variables. In 2016, KOSTAT firstly released store-level micro data for four previous census years(1996, 2001, 2005 and 2010).

[5]The first (available) retail census was carried out in 1996. The next one was undertaken in 2001. The third census was, though, surveyed in *2005*, *four* years after the previous census. Then, in 2010 the latest (available) census was conducted.

The establishment of retail market characteristics is attributable to the recent policy change from non-disclosure to public offering of micro-data by KOSTAT.

Nonetheless, KOSTAT still conceals some variables that allow for store identifications (e.g. store names, store identifiers, etc.) because of the regulation on private data protection. It is, thus, impossible to use retail census as a store-level panel data. Retail census, however, contains store-level geographical indicators such as province codes and county codes. This means that it is possible to construct a county-level panel data by aggregating store-level data at the county-level. This chapter exploits the aggregation strategy to generate retail market characteristics.

The county-level market characteristics used in this study are as follows: First, we utilise three demographical data, populations, average ages and per-capita retail sales. The three demographics are all that we get access to at the county level. Second, we make use of census-based eight market characteristics: price-cost margin, labour productivity, employment, wage, net entry rate, store floor space, rent cost, tax cost. Definitions of the eight variables are as below:[6]

- *Price-cost margin* is defined as the relative importance of markup (= price - marginal cost) to price. We measure it by ratio between profit and sale. We use average price-cost margin across stores within a county.
- *Labour productivity* is defined as sales amount divided by employment. We use mean value across within-county stores.
- *Employment* is the total number of employees who work at retail stores within a county.
- *Wage* is an average value across within-county retail stores, and store-level wage is proxied by labour cost divided by employment.
- *Net entry rate* is defined as net entry(= gross entry - gross exits) divided by the total number of within-county retail stores.
- *Store floor space* is used as the total value across within-county stores.
- *Rent cost* and *tax cost* are also total values of every store in each county.

In sum, this chapter utilises a county-level panel dataset on market entry status of foreign and domestic stores supplemented with a variety of county-specific market characteristics for the years of 2001, 2005 and 2016.[7]

[6]When calculating mean values, we use weighted average across within-county stores, and the weight is store-specific sales share in total value. Labour productivity, wage, rent cost and tax cost are real values deflated by consumer price index.

[7]In fact, the first retail census year that we access to as a micro data is 1996. We do not use 1996 retail census data because most key variables are missing in the dataset.

3.4 Reduced-Form Evidence

3.4.1 Baseline Regression Specification

To estimate the impact of domestic rival's presence on the probability of foreign retail entry, we firstly estimate the following linear probability model:

$$FDIstore_{mt} = \beta Homestore_{mt} + \gamma' X_{mt} + \mu_m + \eta_t + \varepsilon_{mt} \tag{3.1}$$

where m indexes counties, which we defined as a unit of markets, and t indexes census years. $FDIstore_{mt}$ is an entry indicator of a foreign retail store in market m and census year t. $FDIstore_{mt}$ thus takes the value of 1 if a foreign store enters in market m and census year t. $Homestore_{mt}$ is a dummy variable to indicate whether a domestic store that competes against the foreign store enters the market m in census year t. β, a coefficient of interest, captures what extent the probability that a foreign retail store enters a market is affected by domestic rival store existing in that market. To account for the unobservable market heterogeneities, we use county fixed effects (μ_m). We also add census-year fixed effects (η_t) to control for unobserved macroeconomic shocks across census years. ε_{mt} is a *i.i.d.* disturbance term.

Although we include county fixed effects, there could be a concern that time-varying components influencing foreign store entry may be missing. To mitigate the concern, we additionally include a number of time-variant market characteristics, X_{mt}, in the regression model. As mentioned in section 3.3, X_{mt} is a combined set of demographics and market characteristics established by retail census. When multinational retailers choose whether or not to enter a local market, they should not only pay attention to demographics, but also should take retail market characteristics into consideration. Taking advantage of the richness of Korea's retail census data, we add various retail market indicators at the county level.

It is well known that the linear probability model has a flaw. By definition, probabilities must have values in a range between zero and one. Linear probability model, however, cannot guarantee that the estimated predicted values range from zero to one. For this reason, we further estimate a binary choice model to explore whether the presence of Korean (Emart) stores affects the probability of Retail FDI (Tesco) entry. The regression specification of binary choice model is as follows:

$$FDIstore_{mt} = \begin{cases} 1 & \text{if} \quad \alpha + \beta Homestore_{mt} + \gamma' X_{mt} + \upsilon_{mt} > 0. \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

The key issue is the assumption of υ_{mt}. In the linear probability model of Equa-

tion (3.1), υ_{mt} was specified by $\mu_m + \eta_t + \varepsilon_{mt}$. As noted before, we estimate the linear probability model by treating μ_m as the fixed effect, which implies that μ_m is assumed to be correlated to $Homestore_{mt}$. However, it is infeasible to estimate Equation (3.2) under the same assumption about μ_m as in the linear model. This is because, when the length of time periods is short as in the current study (Note that our data set has only four periods of time at most.), so-called *incidental parameters problems* will make the binary choice model with fixed effects get inconsistent estimates (Neyman and Scott, 1948). For the reason, we estimate Equation (3.2) by the probit model with random effects. While the approach relies on a seemingly strong assumption that $f(\mu_m|X_{mt})$ is not dependent on X_{mt}, it certainly makes sure that the predicted probabilities are in a normal range. Consequently, since the linear probability model and the binary choice model complement each other, we analyse the effect of $Homestore_{mt}$ on $FDIstore_{mt}$ by estimating both models.

3.4.2 Baseline Results

Table 3.3 presents baseline estimation results of linear probability model (LPM) specified by Equation 3.1. In Table 3.3, there are three vertical blocks which divided by what market characteristics are included in the regression equation. Each block is further divided depending on what fixed effects are used. Columns (1)-(3) are estimation results when no market characteristics are included. Whether we put no fixed effects (Column (1)), county fixed effects (Column (2)), or both county and census-year fixed effects (Column (3)), the impact of domestic competitor's presence on foreign retail entry is estimated to be significant and positive.

However, it has nothing to do with causal effects particularly if there are time-variant unobserved factors that influence both foreign (Tesco) and domestic (Emart) retail entry. The second block (Columns (4)-(6)) reports results when we add three demographical variables as controls. While columns (4)-(6) still present insignificant estimate for β, a parameter of $Homestore_{mt}$, the magnitudes become smaller compared to those in the first block. In results of the last block (columns (7)-(9)) where county-specific retail market characteristics are additionally augmented in the regression, the magnitudes of estimates for β far more decline than columns (4)-(6). Therefore, Table 3.3 indicates that time-variant market characteristics play an important role in controlling for unobserved confounders.

It is difficult to believe that Table 3.3 shows trustworthy results because the LPM model is unable to assure that fitted probabilities are required to value in normal bounds. To complement the issue, we estimate the probit model of Equation (3.2). Table 3.4 reports results of the probit estimation. Once again, we sort estimation results by control groups and fixed effects in Table 3.4. The same pat-

tern appeared in the LPM is repeated in Table 3.4. That is, when further adding market characteristics, the impact of $Homestore_{mt}$ on $FDIstore_{mt}$ is estimated to become lower. Unlike the estimates in the LPM, however, results with a full of market characteristics (Columns (7)-(9)) show that the effects of domestic presences on foreign entry probabilities are statistically significant and *negative* at the 5%-10% level. Based on column (9), the probit baseline model tells us that the average predicted probability of Tesco's market entry is lower by 6.8% points in counties where Emart stores exist, compared to counties without Emart stores.[8]

3.4.3 Endogeneity and IV Estimation

Endogeneity The main concern for the causal identification is that $Homestore_{mt}$ is endogenous. Basically, it results from two sources of endogeneity. First, endogeneity problem definitely arises if there are unobservable local market factors that affect both firms' entry choices.While we use fixed effects, demographics, and a number of controls exploiting a richness of census data, one may argue that those variables are not sufficient enough to eliminate the concern. In particular, random effect applied to the probit model depends on a strong independence assumption between $f(\mu_m|X_{mt})$ and X_{mt}. If the assumption does not hold, it could be argued that the endogeneity concern is still alive at least in the probit model.

The second source of endogeneity is reverse causality problems. As discussed in Section 3.2, Korea's discount retailing industry is an oligopolistic market and that the two dominant retailers Tesco and Emart compete to each other by opening their stores across Korea's local markets. Accordingly, it is highly likely that firms would make strategic entry decisions by considering their rivals' market entry as shown by illustrative analyses in Section 3.2.2. In other words, $Homestore_{mt}$ does not only affect $FDIstore_{mt}$ but $Homestore_{mt}$ is also affected by $FDIstore_{mt}$ in Equation 3.1 and 3.2. To entirely escape from two sources of endogeneity, we develop an instrument variable (IV) approach.

Instrument variable The IV strategy in this chapter takes inspiration from the literature on retailers' market entry such as Jia (2008) and Ellickson and Misra (2011). They point out that retail firms who operate large discount stores take into account the location of their own distribution centres when they open their stores. Among them, Holmes (2011) is worth mentioning. Holmes (2011) empirically shows that location of Wal-mart stores follows a clear pattern that it tends to open stores in regions that are close to Wal-mart's distribution centres. According to Holmes (2011), this pattern emerges from *economies of density* in

[8]The estimate -6.8% is calculated based on the average marginal effects.

distributions, meaning that retail firms can save their distribution costs by having an agglomeration pattern in terms of opening stores.

Based on the intuition, we construct an instrument variable using information on Emart distribution centres. We measure distances from each store to its closest Emart distribution centre.[9] [10] This measure aims to capture the idea that the probability of Emart entry is highly affected by the distance from Emart's distribution center. Note that each Emart store has started its business at a different time. Similarly, Emart's distribution centres have differently established in terms of opening time. This means, for a certain county, the closest distribution center can be varied across census years. We account for this aspect when constructing IV, and thus the IV varies not only at the spatial dimension but also at the time dimension.

The validity of the proposed IV relies on its uncorrelation with the error terms in the regression of Equation (3.1) and Equation (3.2). That is, the instrument variable needs to be uncorrelated by unobservable factors, other than $Homestore_{mt}$, which would determine entries of retail FDI. To facilitate the disconnection between IV and the errors, we apply the IV approach along with fixed(random) effects and a variety of controls constructed by retail census and population census.

Indeed, the disconnection between distribution centres and the error term seems to make theoretical sense. Intuitively, there are scarcely common demand-side factors to affect entries of both large discount stores and distribution centres, because it is clear that profit-making activities are different from each other: sales versus distribution. For instance, market size is a critical factor for retail store's market entry, but not for distribution centres. From a cost perspective, it seems that the situation is similar to the demand side. That is, it is hard to find components commonly determining store entry decision and distribution centre's entry decision. One may argue that labour market variables could be exceptions in the sense that labour input demands could overlap over both parties. We expect that observable controls such as (log) employments and (log) wages would address the concern. Finally, we have to emphasise that retail FDI (Tesco) has no consideration for locations of its domestic competitor's (Emart's) distribution centers. This is due to the fact that Emart's distribution centres serve only for Emart stores. Therefore, we consider with confidence that the proposed instrument variable is

[9]We calculate euclidean distances from each county to the closest Emart distribution centre by using calculation performed by Google Map.

[10]For counties in which Emart stores are not present, we measure distances from the central point of the county to the closest Emart distribution centre. We did so because those counties are not be chosen by Emart due to their long distances to Emart distribution centres relative to counties with Emart stores.

valid.

Traditionally, the IV approach rests on the linearity of model specification. For the reason, instrument variable methodology can be undoubtedly applied to the *linear probability model*, whereas the binary choice model based on the nonlinear specification results in the difficulty of applying standard IV approach. Even though either two-step GMM estimator or minimum chi-square estimator (Newey, 1987) can be used in binary choice models with endogenous variables (Greene, 2007), these models are in fact practical for the case in which the endogenous variable is continuous. Therefore, neither two-step GMM estimator nor minimum chi-square estimator offer the solution for estimating the causal effect of $Homestore_{mt}$ on $FDIstore_{mt}$ in the Equation (3.2).

Bivariate probit model is an appropriate approach to estimating binary choice model with a dummy endogenous variable as in this chapter. The model in this case becomes,

$$\begin{aligned} FDIstore_{mt} &= 1(\alpha_1 + \beta Homestore_{mt} + \gamma_1{}' X_{mt} + \varepsilon_{1mt} \geq 0) \\ Homestore_{mt} &= 1(\alpha_2 + \delta Z_{mt} + \gamma_2{}' X_{mt} + \varepsilon_{2mt} \geq 0) \end{aligned} \tag{3.3}$$

where Z_{mt} is an exclusion restriction (instrument) variable. The bivariate probit model illustrated in Equation (3.3) is the case where the dependent variable in the second equation, $Homestore_{mt}$, appears in the right-hand side of the first equation. This model is a recursive and simultaneous-equation model, rather than a Seemingly Unrelated Regression(SUR) model. As explained in Greene (2007), even in the simultaneous-equation structure like (3.3), the system of log-likelihood equations in maximum likelihood framework will be transformed into a SUR-structure bivariate probit model, leading to constructing consistent estimators. (Heckman (1977), Maddala (1986)) Relying on the property, we obtain a consistent estimate for β using bivariate probit model specified by (3.3). As Han and Vytlacil (2016) points out, it is important to exploit a valid exclusion restriction in the bivariate probit model with an endogenous dummy variable. Since distances from Emart's closest distribution centre are a valid instrument variable as we explained above, we impose the proposed instrument as the exclusion restriction in this model.

3.4.4 IV Results

Table 3.5 presents IV regression results when linear probability model (LPM) specified by Equation (3.1) is employed. Recall that, in the baseline model, the impact of $Homestore_{mt}$ on $FDIstore_{mt}$ was positive and small magnitude even after using a full set of controls (See column (9) in Table 3.3). In contrast, the IV point

estimates for β in the LPM model are all negative and relatively increased in an absolute values compared to the baseline model, which is shown in the first row of Table 3.5. With a full set of market characteristics used, the IV estimate for β without any fixed effects is -0.935 and significant at the 10% level reported in column (1). If using fixed effects, however, the IV results become insignificant as shown in column (2) and (3).

Column (3), where IV results with a full set of controls, indicates that there is no *causal* effects of domestic rival store's presence on foreign retailer's market entry. As we illustrated in section 3.2.2, graphical analyses show that Tesco hardly avoids opening its stores in markets where Emart is present at the end of 2000s.[11] Based on the analytic finding, we run the same IV regression by taking a sample in which 2010 census data is excluded.[12] In other words, we test whether $Homestore_{mt}$ affects $FDIstore_{mt}$ during the period of 1996 to 2005. We present the results in columns (4)-(6) in Table 3.5.

When we take the pre-2010 sample, IV point estimates of β presented in columns (4)-(6) are all significant, negative and far smaller than when we take the entire sample period. Moreover, the results are robust whether we use fixed effects or not. This experiment clearly suggests that there is a negative *causal* effect of domestic rival's presence on foreign retail entry during the pre-2010 period. Based on column (6) that presents IV result with a full set of controls, the instrumental variable strategy under the LPM model reveals that the entry probability of foreign retailer is lower by 33.5% points in markets where Emart operates its stores, compared to those without Emart stores. In short, the negative causal impact only appears during the period form 1996 up to 2005.

Table 3.5 reports the first stage F-statistics across specifications. The statistics are generally used to determine whether the proposed IV is weak or not. In every column, first stage F-statistics are all below ten, which means they do not exceed a rule-of-thumb criteria for ensuring that the weak IV problem is not significant. The IV estimates, thus, can be thought to be seemingly unreliable results in terms of weak IV problem. While it is possible for our IV results to be plagued by weak IV issue, it can also be argued that the first-stage F-statistics are not reliable for themselves as they are constructed from the linear probability model. As discussed before, LPM estimates are unlikely to be unbiased and consistent because the predicted probabilities could be outside the normal probability bound (i.e., $\hat{FDIstore}_{mt} \in (0, 1)$). Interestingly, most of other variables than $Homestore_{mt}$

[11] We suggest that the weakening may be probably due to the increase in market saturations. Though why this happens is also an interesting topic, we do not explore it because it is beyond this chapter's scope.

[12] Recall that Korea's retail census is undertaken every five or four year. Throughout the chapter, we originally use census data for the years of 1996, 2001, 2005 and 2010.

have insignificant estimated parameters, and it is difficult to believe as the truth.

Therefore, it is crucial to test a bivariate probit regression model to confirm the causal relationship between foreign and domestic entry. Table 3.6 shows that results of bivariate probit model estimation also shows a strong, highly significant and negative effect of domestic competitor's presence on FDI retail entry. Column (1) and (2) are the results of bivariate probit model regression, using the entire period of sample. On the other hand, column (4) and (5) present the estimation results when pre-2010 sample is exploited. Unlike linear probability model, the negative causal effects are appeared all across columns in Table 3.6.

To control for unobserved market heterogeneities, we utilise a bunch of market characteristics data when running the bivariate probit regression as well. One may argue that there can still be unobserved confounders even if controlling them by observed market characteristics. A complete way for dealing with time-invariant unobservables is to use county fixed effects (or adding county dummy variables) as is done in the linear probability models. However, the approach should cause the incidental parameters problem as explained in section 3.4.1. For the reason, we add province specific dummy variables as well as census-year dummies in the regression. While this approach may be insufficient to controls for unobserved market heterogeneities completely, we would expect that it would mitigate the concern to a considerable degree. With regard to the estimation results, province- and year fixed effects make a little difference in terms of magnitude but still yield highly significant and negative estimates for β as shown in column (2) and (4). As a final remark, many of estimated parameters for market characteristics are now estimated to be significant in Table 3.6 (e.g., Log mean age, per-capita retail sales, price-cost margins, Log employment, etc.), which is in stark contrast with LPM estimation results. In conclusion, the average predicted probability of Tesco's market entry is smaller by 10.37% points in counties that experience Emart's entry, rather than counties without Emart stores.

3.5 Structural Estimation

3.5.1 An Empirical Model for Market Entry

In this section, we examine a market entry model between two retail firms. In particular, we develop an entry game theoretical model between a foreign retailer and its domestic competitor so that we can take into account firms' strategic interactions incorporated in their market entry decisions. The model follows the industrial organization(IO) literature on the empirical models of market entry, initiating from Bresnahan and Reiss (1990). Basically, most studies involved in

the literature build on the principle of revealed preferences as standard discrete choice models do. The principle of revealed preference implies that if we observe a firm enters a market, it is because the choice of market entry should have yielded higher profits than the reservation value or the value achieved when market entry were not been chosen. Yet, there is a crucial difference between empirical models of market entry games and discrete choice models. Entry game estimations differs from the discrete choice model in the sense that profit functions must intrinsically have entry choices of *rival firms* as a conditioning argument.

In principle, profit function can be obtained from primitive assumptions on demand and supply, and derive it from the equilibrium conditions. Although this approach would clearly yield structural parameters, all of which have their own structural interpretations, it has a serious limitation that, without appropriate price and quantity data, it would lead model estimations to being intractable. Alternatively, entry game estimation literature generally follows a convention that profit functions are assumed to take a linear form which depends on observable market characteristics. This so-called reduced-form specification of profits has been a de facto standard since Berry (1992). (Ellickson and Misra, 2011) In this chapter, we also follow payoff structure of the previous literature.

With regards to a time horizon of the game, this model is an one-shot game to concentrate on a static and simultaneous moving setting. As such, we consider the *entry* as being active in the market, rather than a transition from being inactive to being active or vice versa (Aguirregabiria, 2014). Taking a look at large discount stores which have entered the Korean market, we can hardly observe that large discount stores by Tesco and Emart exit out of markets once they enter the market. That is, it is clear that large discount stores are continuously active in the market if they once make entry choices in Korea. Therefore, the static game seems to reflect well the reality of the Korean retail sector.

One may argue that estimation of dynamic games is more appropriate for this study in that dynamic games explain well the entry pattern of large discount stores. As discussed in Ackerberg et al. (2007), the main difficulty for estimating dynamic structural models is computational complexity.[13] In particular, dynamic application is much more difficult in this study, because the number of time periods in our dataset is not sufficient enough to estimate structural parameters of dynamic games. There are so many parameters needed to be estimated on which dynamic equilibrium conditions are imposed. In this sense, it is very important to exploit data where periods of time are long. For instance, papers on retailers' dynamic entry games such as Holmes (2011) and Aguirregabiria and Vicentini (2016) use

[13]According to Ackerberg et al. (2007), equilibrium exists as the continuation values in dynamic structural models, which would make computations extremely burdensome.

data sets where the number of time periods are 44 and 56, respectively. On the other hand, the number in our data set is at most four.

The model is based on Berry (1992), followed by Ciliberto and Tamer (2009) and Ellickson and Misra (2011).[14] We assume that there are two retail firms, a foreign retailer and a domestic retailer, denoted by F and H respectively. In an economy where there are M local markets, two firms are competing with each other in every local market by taking an action. The firms' actions are two: either entering the market or not entering the market. We also assume that each retailer can open a single store per market and that firms' actions have no spillover effects across other markets. This implies that two firms compete only within a local market. As a result, the model is assumed to be independent replications of 2×2 discrete game.

Let the profit function of retail firm $i = \{F, K\}$ in market m be $\pi_{im}(\theta; y_{-im})$ where y_{-im} implies the rival firm's action in market m. θ is a finite number of parameters of interest shaping the profit function. The profit function of firm i in market m depends on both market specific and firm-market specific characteristics. Let X_m be a vector of market specific variables common to both firms. $Z_{\mathrm{F}m}$ and $Z_{\mathrm{H}m}$ denote F's and H's firm-market specific variables respectively. The profit function in market m is given by

$$\pi_{im} = X_m^{'}\alpha + Z_{im}^{'}\beta_i + \delta_i y_{-im} + \varepsilon_{im} \tag{3.4}$$

where ε_{im} is a component of profits which is unobservable to the econometricians.

This model is a game of incomplete information. We assume that foreign retailer and domestic retailer cannot observe everything about each other's profit. In this chapter, We pay attention to two aspects as the sources of incomplete information. First, foreign direct investment(FDI) theories emphasise that multinational enterprises(MNEs) have superior technology and knowledge over domestic firms such as brand names, better managerial ability and so on. The advantages manifest themselves as higher productivity, resulting in offering them opportunities to obtain market shares in host economies. In this model, we assume that domestic firm cannot observe foreign retailer's superior technology or knowledge to make FDI more profitable. The other source of asymmetric information is a difference in ability to obtain local demand information between foreign and domestic retailers. In this case, the domestic firm has a superior knowledge on local demand over the foreign firm. We assume that this local demand advantage is incorporated only in domestic firm's profit.

[14] In particular, Ellickson and Misra (2011) displays entry game model, estimation methodology and its application to the retail sector. The structural estimation in this chapter borrows from many of Ellickson and Misra (2011)'s approaches.

To implement incomplete information in the model, we assume that firm i observes its own $\varepsilon_i = \{\varepsilon_{i1}, \varepsilon_{i2}, .., \varepsilon_{iM}\}$, but cannot observe its competitor's ε_{-i}. Though firm i cannot observe individual draws of ε_{-i}, we also assume that firm i knows the distribution of ε_{-i}, denoted by $F(\varepsilon_{-i})$, and that it forms beliefs over its rival's action through $F(\varepsilon_{-i})$. Each firm chooses an action which maximises expected profits given those beliefs. Further, we assume that econometricians, who cannot observe individual draws of ε_i and ε_{-i}, know the distributions of them. As noted in Ellickson and Misra (2011), this will make firms stand on an equal footing with the econometricians and will be useful when constructing an estimator. If firms make choices that will maximise expected profits, given the beliefs over the competitors' actions(y_{-im}), then

$$\begin{aligned} y_{\mathrm{F}m} &= 1(X_m^{'}\alpha + Z_{\mathrm{F}m}^{'}\beta_{\mathrm{F}} + \delta_{\mathrm{F}} E\big[y_{\mathrm{H}m}(X_m, Z_{\mathrm{H}m}, \varepsilon_{\mathrm{H}m})|X_m, Z_{\mathrm{H}m}\big] + \varepsilon_{\mathrm{F}m} \geq 0) \\ y_{\mathrm{H}m} &= 1(X_m^{'}\alpha + Z_{\mathrm{H}m}^{'}\beta_{\mathrm{H}} + \delta_{\mathrm{H}} E\big[y_{\mathrm{F}m}(X_m, Z_{\mathrm{F}m}, \varepsilon_{\mathrm{F}m})|X_m, Z_{\mathrm{F}m}\big] + \varepsilon_{\mathrm{H}m} \geq 0) \end{aligned} \tag{3.5}$$

Let $E\big[y_{im}(X_m, \varepsilon_{im})|X_m\big] \equiv p_{im}$. Then, Bayesian Nash equilibrium should satisfy the system of equations below.

$$\begin{aligned} p_{\mathrm{F}m}^{*} &= \psi(X_m^{'}\alpha + Z_{\mathrm{F}m}^{'}\beta_{\mathrm{F}} + \delta_{\mathrm{F}} p_{\mathrm{H}m}^{*}) \\ p_{\mathrm{H}m}^{*} &= \psi(X_m^{'}\alpha + Z_{\mathrm{H}m}^{'}\beta_{\mathrm{H}} + \delta_{\mathrm{H}} p_{\mathrm{F}m}^{*}) \end{aligned} \tag{3.6}$$

where the non-linear function $\psi(\cdot)$ depends on the specification of $F(\cdot)$. Equations (3.6) represents best response probabilities of two firms, wherein expected profits are mapped to ex-ante choice probabilities of firms.

3.5.2 Structural Estimation

If $F(\cdot)$ is an absolutely continuous function, Bayesian Nash equilibrium in Equations (3.6) is guaranteed to exist by Brouwer's fixed point theorem.[15] Since the system of equilibrium Equations in (3.6) is highly non-linear, it is numerically difficult to derive a closed form solution. To find estimates of structural parameters in the model, one can rely on successive approximation approaches to obtain an equilibrium.[16] A serious limitation in the approach based on successive approximations, however, is that it assumes uniqueness of equilibrium. In principle,

[15]Brouwer's fixed point theorem states that for any continuous function f which maps a compact set into itself, there should be a point x_0 such that $f(x_0) = x_0$

[16]This approach, developed by Rust (1987), is based on the Brouwer's fixed point theorem and rests on the uniqueness assumption of the equilibrium. In particular, once finding the solution by using successive approximation methods, it is possible to nest the approximated solutions into likelihood functions and then we can estimate structural parameters in Equations (3.6) by maximum likelihood estimator. This methodology is called the nested fixed point (NFXP) approach by its identification strategy.

games may admit multiple equilibria. This multiple equilibria situation can also be illustrated in this model. Suppose δ_{F}, $\delta_{\mathrm{H}} < 0$ in Equations (3.5). It is clear that multiple equilibria happen when

$$-X_m^{'}\alpha_i - Z_{im}^{'}\beta \leq \varepsilon_{im} \leq -X_m^{'}\alpha - Z_{im}^{'}\beta_i - \delta_i \tag{3.7}$$

for $\forall m$, $i = \mathrm{F},\ \mathrm{M}$. More specifically, when $\varepsilon_{\mathrm{F}m}$ has a value such that Equation (3.7) satisfies, both {F enters, H does not enter} and {F does not enter, H enters} are equilibrium pairs of firms' market entry choices. In this case, the relationship from (X, Z, ε) to y is not one-to-one mapping, resulting in an *imcomplete econometric model* (Tamer, 2003).

In this chapter, we make use of the two-step estimation approach which began with Hotz and Miller (1993). The benefit of the method is that it can be exploited to complete the empirical model in games of incomplete information.[17] As illustrated earlier, market entry model with incomplete information setting entails the (ex-ante) choice probabilities for the equilibrium (i.e., $p_{im} \equiv E\big[y_{im}(X_m, \varepsilon_{im})|X_m\big]$). The two-step approach is firstly to estimate the conditional choice probabilities (CCP) by regressing firms' entry choices on observables, and then to plug the estimates for CCP into the right hand side of Equations (3.6).

For the first stage, we construct consistent estimators of the equilibrium conditional choice probabilities. As discussed in Ellickson and Misra (2011), the CCP should be ideally estimated *non-parametically*, because the CCPs are not derived from any economic primitives. Any parametric model is thus highly likely to be mis-specified and inconsistent. For the reason, we employ the kernel regression method to estimate the CCPs. Let $\mathrm{w}_m^i \equiv (X_m, Z_{im})$. The first-step Kernel estimator is

$$\hat{p}_{im}(\mathrm{w}_m^i) = \frac{\sum_{s=1}^{M} y_{is}\, \mathrm{K}(\frac{\mathrm{w}_m^i - \mathrm{w}_s^i}{b})}{\sum_{s=1}^{M} \mathrm{K}(\frac{\mathrm{w}_m^i - \mathrm{w}_s^i}{b})} \tag{3.8}$$

where $\mathrm{K}(\cdot)$ is a kernel function, b is the bandwidth, and $i = \{F, K\}$

At the second stage, we plug the estimated CCPs obtained from the first stage, $\{\hat{p}_{\mathrm{F}m}, \hat{p}_{\mathrm{H}m}\}$, into the right hand side of Equation (3.6) like,

$$p_{im}^{*} = \phi(X_m^{'}\alpha + Z_{im}^{'}\beta_i + \delta_i \hat{p}_{im}) \tag{3.9}$$

where $\phi(\cdot)$ is assumed to be the standard normal density. Once the fitted probabilities(p_{im}^{*}) are obtained from Equation (3.9), log-likelihood is a funtion nested by the fitted

[17] As a matter of fact, Hotz and Miller (1993) originally proposed the approach to estimate dynamic discrete game of complete information. Recall that equilibrium of dynamic games exists as continuation values, which leads to the computational complexity. In this regard, the two-step approach is also able to be applied to estimations of incomplete informational static games where equilibrium is also expressed as continuation values specified by Equation (3.6).

CCPs as the following:

$$\ln L = \sum_{m=1}^{M} \sum_{i \in \{F,K\}} y_{im} \ln (p^*_{im}) + (1 - y_{im}) \ln (1 - p^*_{im}) \tag{3.10}$$

Two-step approach is evaluated as an estimation method to be able to deal with multiple equilibria issue. In particular, this approach addresses the multiplicity by considering that only one equilibrium is played in data, so that the fitted estimates for CCPs (p^*_{im}) will effectively condition on the equilibrium that was played in data (Ellickson and Misra, 2011). This makes two-step approach differ from other models, like the nested fixed point (NFXP) approach that unique equilibrium is assumed to exist in the *model* from the beginning.

Unlike reduced-form estimations, we make use of a cross-section data for a particular year in order to identify the structural parameters on the entry game model. As emphasised earlier, the structural model is static in nature. That is, we suppose that all Tesco and Emart stores are opened in one period. We begin with estimating the entry game using data in 2010. As two retail firms have continually expanded their large discount stores and have never divested their stores during our sample period, taking the latest cross-section data would enable me to have as large variations in retail entry as possible. Then, we do the structural model estimation through 2005 data as well. In section 3.2.2, we show that strategic market entry looks weakened at the end of 2000s via graphical analyses. In this sense, on top of data for 2010, it is worth further estimating the structural entry game model, using 2005 data. When we described model, we assume that each firm's profit function consists of market characteristics and firm-market specific components. Market characteristics used in the model estimation are same as in the reduced-form regression, and each store's distance from the closest distribution center are used for the firm-market specific variable.

Table 3.7 lists the parameter estimates for the entry game model between a foreign and a Korean retailer. According to Angrist and Pischke (2008), bootstrapped standard errors are desirable particularly when estimation involves a number of steps. Since we exploit two-step approach, we apply bootstrapping method to calculate estimates of standard errors of parameters. Columns (1) is the result of structural estimation when 2010 data is exploited. The estimate for the impact of a domestic competitor Emart's entry on a foreign retailer Tesco's entry is negative(-0.811) and is statistically significant at the 1% level. The highly significant and negative estimate is in line with the result of IV estimation (bivariate probit model) illustrated in Table 3.6 (-1.400; see the column (2)). This means that the adverse effect of domestic rival's stores on foreign retailer's entry, which

has been estimated in the reduced-form regression, is confirmed by the entry game model estimation as well. In column 2, we report the results of entry game estimation where a cross-sectional data in 2005 is utilised. The structural parameter that shows the effect of Emart's entry on Tesco's entry is estimated to be -0.603, which is statistically significant at the 10% level. Column (1) and (2) suggest that, between Tesco and Emart, strategic market entry really appears in Korean retail market whether we take 2005 data or 2010 data. Interestingly, structural estimation results correspond to IV results via bivariate probit model, but are not matched to linear probability model with IV strategy. Considering the limitation of linear probability model explained earlier, both structural model estimations and bivariate probit model regression jointly indicate that foreign retailer makes a strategic entry decision by taking into account its competitor's entry choice in home country from 1996 to 2010.

So far, we have shown that retail multinationals are unlikely to make entry choices into Korean markets without considering choices by host economies' influential competitors. It is confirmed both by reduced-form estimations with IV strategy, and by structural game-theoretical model estimation in which strategic actions between firms are fully incorporated. *Why* does FDI strategically enter host-country markets strategically? The answer may be related to the question of *when* or under what circumstances FDI makes strategic entries. We already explained that Korean retail sector has a marked feature that it is highly concentrated on two big firms, Tesco and Emart. In such an oligopolistic market, even multinationals who have technological edges have no choice but to take into account actions of host countries' rivals. As a result, it would seem that market structures are profoundly associated to FDI's strategic entry into host-country markets.

3.5.3 Incomplete Information versus Complete Information

Empirical studies of entry game models rest on the informational assumptions between players. Our structural estimations identify structural parameters under entry games of *incomplete information* between a foreign and a Korean retailer. A worthwhile question is what happens to structural parameter estimates when the structural model is a game of *complete information*. In this subsection, we estimate the entry game outlined in section 3.5.1 by imposing a different assumption on information sets.

As in a game of incomplete information, the profit functions are given by

$$\pi_{im} = X_m^{'}\alpha + Z_{im}^{'}\beta_i + \delta_i y_{-im} + \varepsilon_{im} \tag{3.11}$$

Though the shape of the function is identical to Equation (3.4), the assumption on ε_{im} becomes different from the incomplete information model. ε_{im}, a component of profits which are unobserved to econometricians, is common knowledge to both firms. In other words, they perfectly know each other's profit function. Applying a simultaneous moving again, the equilibrium is Nash equilibrium determined by the following:

$$\begin{aligned} y_{\mathrm{F}m}^{*} &= 1(X_m^{'}\alpha + Z_{\mathrm{F}m}^{'}\beta_{\mathrm{F}} + \delta_{\mathrm{F}} y_{\mathrm{H}m}^{*} + \varepsilon_{\mathrm{F}m} \geq 0) \\ y_{\mathrm{H}m}^{*} &= 1(X_m^{'}\alpha + Z_{\mathrm{H}m}^{'}\beta_{\mathrm{H}} + \delta_{\mathrm{H}} y_{\mathrm{F}m}^{*} + \varepsilon_{\mathrm{H}m} \geq 0) \end{aligned} \tag{3.12}$$

where an equilibrium, $\{\ y_{\mathrm{F}m}^{*},\ y_{\mathrm{H}m}^{*}\ \}$ satisfies two nonlinear equations above simultaneously.[18]

To identify the game of *complete* information above, we apply Berry's (1992) approach to find structural parameter estimates. The approach has a merit that it can solve the multiple equilibria issue for complete information games. Specifically, Berry's approach addresses the issue in a way that he completes the empirical model by assuming a sequential entry rule among multiple equilibria outcomes. Suppose that both $\{1, 0\}$ and $\{0, 1\}$ are simultaneously equilibrium outcomes in the model where the first element of the bracket indicates the entry status of Tesco (entry $\equiv$ 1), and the second one does Emart's. In this case, since the number of firms in that market can be either one or two, it is impossible to complete the empirical model, thus disabling the model to be estimated. However, if a particular entry order is given, a prediction of the number of firms in the market is uniquely determined. For instance, if Emart always enters first, then $\{1, 0\}$ will automatically be removed out of two equilibria outcomes $\{1, 0\}$ and $\{0, 1\}$. In this way, it becomes possible to complete econometric model through giving the entry order. Like the illustrative example above, we assume that, when both firms' expected profits are all positive (i.e., For both firms, best response strategy is to entry into market.), domestic retailer (Emart) always enters first. The entry rule is grounded in fact to a certain degree. In our data set, there are much more cases that Emart stores enter earlier than Tesco.

Given the profit functions and the entry order assumption, we construct prob-

[18]Note that, in the game model of incomplete information (section 3.5.1), the equilibrium is Bayesian Nash equilibrium, which is characterised by best response probabilities.

abilities for four cases as the following:

$$
\begin{aligned}
&\text{(C1)}\quad p(y_{\mathrm{F}m}=0, y_{\mathrm{H}m}=0) \\
&\qquad = \phi(X_m^{'}\alpha + Z_{\mathrm{F}m}^{'}\beta_{\mathrm{F}} + \varepsilon_{\mathrm{F}m} < 0) \times \phi(X_m^{'}\alpha + Z_{\mathrm{H}m}^{'}\beta_{\mathrm{H}} + \varepsilon_{\mathrm{H}m} < 0) \\
&\text{(C2)}\quad p(y_{\mathrm{F}m}=1, y_{\mathrm{H}m}=1) \\
&\qquad = \phi(X_m^{'}\alpha + Z_{\mathrm{F}m}^{'}\beta_{\mathrm{F}} + \delta_{\mathrm{F}} + \varepsilon_{\mathrm{F}m} \geq 0) \times \phi(X_m^{'}\alpha + Z_{\mathrm{H}m}^{'}\beta_{\mathrm{H}} + \delta_{\mathrm{H}} + \varepsilon_{\mathrm{H}m} \geq 0) \\
&\text{(C3)}\quad p(y_{\mathrm{F}m}=0, y_{\mathrm{H}m}=1) \\
&\qquad = \phi(X_m^{'}\alpha + Z_{\mathrm{F}m}^{'}\beta_{\mathrm{F}} + \delta_{\mathrm{F}} + \varepsilon_{\mathrm{F}m} < 0) \times \phi(X_m^{'}\alpha + Z_{\mathrm{H}m}^{'}\beta_{\mathrm{H}} + \varepsilon_{\mathrm{H}m} \geq 0) \\
&\text{(C4)}\quad p(y_{\mathrm{F}m}=1, y_{\mathrm{H}m}=0) \\
&\qquad = 1 - p(y_{\mathrm{F}m}=0, y_{\mathrm{H}m}=0) - p(y_{\mathrm{F}m}=1, y_{\mathrm{H}m}=1) - p(y_{\mathrm{F}m}=0, y_{\mathrm{H}m}=1)
\end{aligned}
$$

where $\phi(\cdot)$ is assumed to be the standard normal density. Then, we formulate a log likelihood function by using the constructed probabilities above. The maximum likelihood estimator is defined as

$$
\begin{aligned}
\hat{\theta}_{Berry} = \arg\max_{\theta} \sum\nolimits_{m=1}^{M} \Big[\; & \mathbf{1}\{y_{\mathrm{F}m}=0, y_{\mathrm{H}m}=0\} \cdot \ln p(y_{\mathrm{F}m}=0, y_{\mathrm{H}m}=0) + \\
& \mathbf{1}\{y_{\mathrm{F}m}=1, y_{\mathrm{H}m}=1\} \cdot \ln p(y_{\mathrm{F}m}=1, y_{\mathrm{H}m}=1) + \\
& \mathbf{1}\{y_{\mathrm{F}m}=0, y_{\mathrm{H}m}=1\} \cdot \ln p(y_{\mathrm{F}m}=0, y_{\mathrm{H}m}=1) + \\
& \mathbf{1}\{y_{\mathrm{F}m}=1, y_{\mathrm{H}m}=0\} \cdot \ln p(y_{\mathrm{F}m}=1, y_{\mathrm{H}m}=0) \; \Big]
\end{aligned}
$$

where $\mathbf{1}\{y_{\mathrm{F}m}=i, y_{\mathrm{H}m}=j\}$ for $i, j = 1, 0$ is a dummy indicator which equals to one if the expression inside bracket holds, and zero otherwise.

Table 3.8 presents results of structural estimation of the complete information game, which is conducted by Berry's (1992) approach. Column (1), a result for 2010 data, shows that the impact of domestic competitor's entry on foreign retail entry is negative(-0.523) and significant at the 10% level, suggesting that a multinational retailer, Tesco, makes a strategic entry choices, caring about the actions of Korean retailer, Emart. On the other hand, the effect of FDI entry on choices by indigenous firm is statistically insignificant, which indicates that Emart enters Korean local markets independent on Tesco's entry choices. In column (2), we report the estimation results of complete informational entry game, using data for the year of 2005. With regards to competition effects between two firms, column (2) demonstrates the opposite effects to column (1). Tesco is now unlikely to make strategic actions whereas Emart's entry behaviour looks dependent on its competitor, Tesco.

In contrast, the estimates for the competition effects under complete information in Table 3.8 show a different pattern compared to model estimations of incomplete information game. Recall that Table 3.7 clearly shows that both retail

firms make strategic interactions to each other in terms of market entry, whether looking at 2005 data or 2010 data. In other words, there is no doubt that Tesco and Emart enter the market strategically or caring about each other's choices under the assumption of asymmetric information. Turning to a game of complete information documented in Table 3.8, however, we do not find strong evidence that strategic entry matters for Tesco and Emart.

Given that two models deliver contrary results, a pivotal question is the following: Which model makes more sense? According to (Ellickson and Misra, 2011), the answer would rest on the specifics to which we want to apply game model estimations. We consider that it is more appropriate to assume that payoff functions are not common knowledge in both foreign and domestic firms for two reasons. First, incomplete information makes sence from a theoretical perspective. Standard theories of multinational enterprises or FDI typically indicate that there is asymmetric information between foreign firms and indigenous firms in the host countries. In particular, it is well known that MNEs are more likely to have superior information about technology and managerial know-how, whereas domestic firms may have advantage in obtaining local market information.

Furthermore, in Section 3.4, we showed that there is a reduced-form effect of domestic store presence on foreign entry. Using bivariate probit models based on instrumental variable strategy, we found that domestic store's market presence has a negative caual impact on foreign retail entry. Moreover, the negative caual impact appears regardless of whether we chose all census years or pre-2010 census years as the sample period (Table 3.6). Note that this patter is more consistent with the structural estimation results of incomplete information game model. In particular, when we apply incomplete informational game, the structural parameters that indicates the impact of domestic entry on foreign entry are clearly significant and negative not only in 2005 but also in 2010 (Table 3.7). On the other hand, when estimating the complete information game model, the impact of domestic entry on foreign entry is estimated to be negative and significant only for the 2010 sample whereas it is not for the 2005 sample (Table 3.8). Consequently, the reduced-form evidence suggests that incomplete information game model corresponds to the observed data better than complete informational game model.

There is an important implication in the contrast of two structural models. The contrasting estimation results of two informational game models would indicate that asymmetric information between foreign and Korean retailer plays a certain role in forming strategic entries. When we take into account asymmetric information between FDI and domestic, the model estimations are likely to fit for the real world reflected by the observed data. In other words, when two firms

are ready to or have intentions to do strategical behaviours in oligopolistic markets, if information is perfectly known to each other, strategic interactions do not eventually happen at Korea's retail industry.

3.6 Conclusion

We empirically showed the existence of retail FDI's strategic market entry into host economies by reduced-form evidence. Additionally, we build a model in which the strategic interactions are fully incorporated, and estimate the game-theoretical model. The existence of strategic interactions are clearly confirmed by structural estimation results as well. Finally, we compare two structural model estimation results between incomplete information game and complete information game. We find that asymmetric information plays a crucial role in the strategic interactions incorporated in retail FDI's market entry into Korean market.

Table 3.1: Total Sales in Korean Large Discount Stores from 1988 to 2008

Year	1998	2001	2005	2008
Total Sales of Large Discount Stores	5.0	17.0	23.7	31.9
Share in the Retail Sector (%)	5.7	13.9	16.2	14.7
Total Sales of the Retail Sector	88.4	122.4	146.3	217.0

Source: Korea Fair Trade Commission(KFTC, 2009)

Notes: This table shows how much Large Discount Stores accounts for in the Korean retail sector over years. All the sales amounts in this table are reported by a unit of Korean Trillion Won.

Table 3.2: Distribution of Counties by Tesco's and Emart's Entry

Year	1996	2001	2005	2010	Total
Foreign (Tesco's) Store Presence	2	18	25	43	88
Domestic (Emart's) Store Presence	5	27	50	53	135
Both Foreign and Domestic Presence	0	7	25	44	76
No Store	244	199	151	111	705

Source: Korea Chainstore Association

Notes: This table presents distributions of counties according to county-level market entry by Tesco and Emart.

Table 3.3: Baseline Results: Linear Probability Model

	No Market Characteristics			*Demographics*			*Demog. + Retail Market Char.*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Homestore	0.241***	0.286***	0.128***	-0.00393	0.0538	0.0413	-0.0626	0.00650	0.001
	(0.0444)	(0.0418)	(0.0446)	(0.0526)	(0.0598)	(0.0608)	(0.0504)	(0.0604)	(0.0598)
Log Population				0.239***	0.591***	0.548***	0.0441	0.522**	0.501**
				(0.0326)	(0.159)	(0.165)	(0.0744)	(0.205)	(0.226)
Log Mean Age				0.422**	1.342***	0.0664	-0.0418	0.536	0.0998
				(0.181)	(0.262)	(0.940)	(0.193)	(0.435)	(0.907)
Per-Capita Retail Sale				-0.00268	0.00605	-0.00282	-0.0148**	-0.00247	-0.00440
				(0.00249)	(0.00751)	(0.00520)	(0.00619)	(0.00396)	(0.00397)
Price-Cost Margin							-1.205***	-0.160	0.0619
							(0.440)	(0.419)	(0.437)
Log Labour Productivity							0.0120	0.000573	-0.0195
							(0.0335)	(0.0326)	(0.0390)
Log Employment							0.185*	0.0705	0.104
							(0.0967)	(0.149)	(0.183)
Log Wage							-0.00614	0.0197	0.0395
							(0.0404)	(0.0430)	(0.0485)
Net Entry Rate							0.436***	0.428***	0.377***
							(0.106)	(0.118)	(0.127)
Log Store Floor Space							0.0337	-0.00648	-0.0258
							(0.0784)	(0.0600)	(0.0606)
Log Rent Cost							-0.110***	-0.0275	-0.0409
							(0.0416)	(0.0504)	(0.0502)
Log Tax Cost							0.0626***	0.0443*	0.0415*
							(0.0251)	(0.0234)	(0.0233)
County Fixed Effect	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Census-Year Fixed Effect	No	No	Yes	No	No	Yes	No	No	Yes
N	951	951	951	673	673	673	655	655	655
R^2	0.070	0.576	0.631	0.213	0.732	0.737	0.297	0.750	0.751
Clusters	251	251	251	251	251	251	244	244	244

Notes: This table illustrates baseline results of *linear probability model* to relate the entry dummy of foreign stores to the presence indicator of domestic competitor's store. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the county level.

Table 3.4: Baseline Results: Binary Choice (Probit) Model

	No Market Characteristics			*Demographics*			*Demog. + Retail Market Char.*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Homestore	0.822***	1.549***	0.938***	-0.0979	0.0675	-0.704*	-0.496***	-0.985**	-1.110**
	(0.133)	(0.228)	(0.462)	(0.171)	(0.340)	(0.400)	(0.190)	(0.478)	(0.490)
Log Population				1.168***	3.372***	2.194***	-1.022*	-2.696*	-2.595*
				(0.176)	(0.807)	(0.543)	(0.613)	(1.580)	(1.477)
Log Mean Age				1.741*	9.858***	-6.720**	-1.300	-2.884	-8.172**
				(1.022)	(3.117)	(3.281)	(1.154)	(3.018)	(4.106)
Per-Capita Retail Sale				-0.00779	0.0360	-0.0552	-0.323***	-0.798***	-0.679***
				(0.0159)	(0.0375)	(0.0462)	(0.0789)	(0.257)	(0.227)
Price-Cost Margin							-13.01***	-23.13***	-17.35**
							(2.849)	(7.364)	(7.044)
Log Labour Productivity							0.0281	0.279	0.213
							(0.189)	(0.533)	(0.584)
Log Employment							2.072***	4.123**	5.153***
							(0.561)	(1.633)	(1.804)
Log Wage							0.462*	1.206*	1.355*
							(0.268)	(0.711)	(0.759)
Net Entry Rate							2.086***	6.492***	4.199***
							(0.566)	(1.735)	(1.595)
Log Store Floor Space							0.595	1.830*	1.525
							(0.375)	(0.957)	(0.944)
Log Rent Cost							-0.600*	-1.003	-1.754**
							(0.323)	(0.817)	(0.851)
Log Tax Cost							0.385	1.351**	1.056*
							(0.252)	(0.660)	(0.663)
County Random Effect	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Census-Year Dummy	No	No	Yes	No	No	Yes	No	No	Yes
N	951	951	951	673	673	673	655	655	655
R^2	0.070	0.576	0.631	0.213	0.732	0.737	0.297	0.750	0.751
Clusters	251	251	251	251	251	251	244	244	244

Notes: This table illustrates baseline results of *probit model* to estimate the effect of domestic competitor store's presence on the foreign store entry. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the county level.

Table 3.5: IV Results: Linear Probability Model

	All Census Years			*Excluding 2010 Census*		
	(1)	(2)	(3)	(4)	(5)	(6)
Homestore	-0.935*	-0.0184	-0.00855	-0.826*	-0.333***	-0.335***
	(0.501)	(0.142)	(0.157)	(0.461)	(0.122)	(0.120)
Log Population	0.153	0.544**	0.507*	0.0618	1.004***	1.919***
	(0.131)	(0.232)	(0.239)	(0.104)	(0.294)	(0.404)
Log Mean Age	0.0490	0.567	0.0975	0.205	2.186***	3.747**
	(0.258)	(0.466)	(0.911)	(0.284)	(0.714)	(1.894)
Per-Capita Retail Sale	-0.0163*	-0.00203	-0.00428	-0.0205**	-0.0141	-0.0211
	(0.00868)	(0.00479)	(0.00454)	(0.00939)	(0.0270)	(0.0266)
Price-Cost Margin	-0.766	-0.144	0.0717	-0.801	0.374	0.323
	(0.626)	(0.431)	(0.479)	(0.577)	(0.485)	(0.493)
Log Labour Productivity	0.00257	-0.00150	-0.0201	0.0801	0.0545	0.0515
	(0.0437)	(0.0377)	(0.0425)	(0.0676)	(0.0475)	(0.0479)
Log Employment	0.138	0.0645	0.104	-0.0648	0.215	0.117
	(0.151)	(0.153)	(0.184)	(0.163)	(0.215)	(0.249)
Log Wage	0.0808	0.0226	0.0404	0.0654	0.0108	0.0181
	(0.0805)	(0.0481)	(0.0525)	(0.0935)	(0.0537)	(0.0535)
Net Entry Rate	0.778***	0.433***	0.376***	0.648***	0.168	0.203
	(0.262)	(0.118)	(0.128)	(0.232)	(0.147)	(0.152)
Log Store Floor Space	0.146	-0.00422	-0.0252	0.365*	0.122	0.132
	(0.134)	(0.0632)	(0.0628)	(0.212)	(0.134)	(0.141)
Log Rent Cost	-0.142**	-0.0278	-0.0413	-0.113**	0.0201	0.0332
	(0.0644)	(0.0499)	(0.0486)	(0.0545)	(0.104)	(0.102)
Log Tax Cost	0.121**	0.0454**	0.0418*	0.0676**	0.0423**	0.0439**
	(0.0440)	(0.0225)	(0.0223)	(0.0331)	(0.0195)	(0.0190)
County Fixed Effect	No	Yes	Yes	No	Yes	Yes
Census-Year Fixed Effect	No	No	Yes	No	No	Yes
N	655	625	625	417	404	404
R^2	0.322	0.750	0.751	0.396	0.804	0.805
1st Stage F-Stat.	5.08	8.80	7.08	4.50	8.08	5.92
Clusters	244	214	214	215	202	202

Notes: This table shows instrumental variable estimation results of *linear Probability Model* to assess the causal impact of domestic competitor store's existence on the foreign store entry. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the county level.

Table 3.6: IV Results: Binary Choice (Probit) Model

	All Census Years		*Excluding 2010 Census*	
	(1)	(2)	(3)	(4)
Homestore	-1.714***	-1.400**	-1.747***	-1.443***
	(0.176)	(0.711)	(0.156)	(0.538)
Log Population	-0.511	-1.333	0.254	-0.608
	(0.626)	(0.823)	(0.530)	(0.832)
Log Mean Age	-0.919	-4.422**	0.905	-3.598
	(0.936)	(1.854)	(1.159)	(3.048)
Per-Capita Retail Sale	0.228**	-0.296***	-0.175***	-0.260**
	(0.106)	(0.0981)	(0.0628)	(0.103)
Price-Cost Margin	-8.533**	-8.066**	-8.404***	-10.670***
	(3.376)	(3.324)	(2.483)	(3.674)
Log Labour Productivity	-0.0224	0.0461	0.284	0.878*
	(0.158)	(0.235)	(0.248)	(0.375)
Log Employment	1.414**	2.295***	0.577	1.794**
	(0.645)	(0.864)	(0.595)	(0.885)
Log Wage	0.592**	0.562*	0.366	-0.446
	(0.240)	(0.307)	(0.305)	(0.497)
Net Entry Rate	2.320***	1.671**	2.546***	2.368**
	(0.563)	(0.702)	(0.722)	(1.012)
Log Store Floor Space	0.829**	1.251***	1.466***	2.352**
	(0.336)	(0.397)	(0.501)	(0.713)
Log Rent Cost	-0.506*	-1.024**	-0.621**	-1.606***
	(0.267)	(0.419)	(0.265)	(0.503)
Log Tax Cost	0.356	0.396	0.143	0.221
	(0.222)	(0.244)	(0.217)	(0.330)
Province Dummy	No	Yes	No	Yes
Census-Year Dummy	No	Yes	No	Yes
N	655	655	417	417
Log Pseudo Likelihood	-438.0191	-400.0115	-221.5791	-197.628

Notes: This table demonstrates instrumental variable estimation results of *binary choice(probit) model* to estimate the causal impact of domestic retail competitor's market presence on foreign retail entry. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01.

Table 3.7: Structural Parameter Estimates: Incomplete Information Game

	2010 Census (1)	*2005 Census* (2)
The Effect of Domestic Entry on Foreign Entry	-0.811***	-0.603*
	(0.251)	(0.347)
The Effect of Foreign Entry on Domestic Entry	-0.939***	-1.258**
	(0.288)	(0.538)
Log Population	-0.324	-0.382
	(0.595)	(0.588)
Log Mean Age	-2.741**	-3.792**
	(1.200)	(1.715)
Per-Capita Retail Sale	-0.0645	-0.182***
	(0.0729)	(0.0642)
Price-Cost Margin	-2.812	-4.097
	(3.764)	(2.673)
Log Labour Productivity	-0.240	-0.746
	(0.274)	(0.554)
Log Employment	1.264**	0.865
	(0.590)	(0.940)
Log Wage	0.143	-0.267
	(0.412)	(0.638)
Net Entry Rate	1.092	0.761
	(0.921)	(1.208)
Log Store Floor Space	0.109	1.259*
	(0.335)	(0.663)
Log Rent Cost	-0.541	-0.890**
	(0.340)	(0.407)
Log Tax Cost	0.750***	0.663**
	(0.244)	(0.272)
Distance From Foreign (Tesco) Distribution Centre	-0.000624	-0.00351
	(0.00239)	(0.00349)
Distance From Foreign (Tesco) Distribution Centre	0.00123	0.00126
	(0.00197)	(0.00199)
N	238	215
Log Likelihood	203.509	130.037

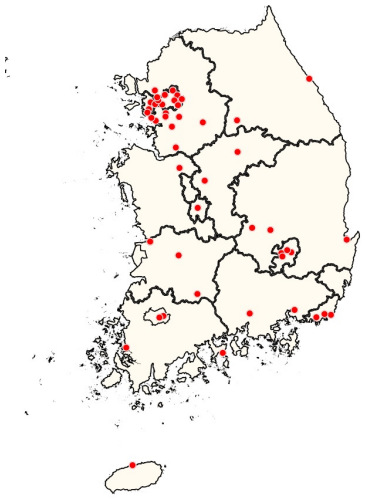
Notes: This table shows estimates for structural parameters of entry game-theoretical model between a foreign retail store and a domestic retail store under *incomplete* information assumption. For parameter identifications, we use the two-step approach initially developed by Hotz and Miller (1993). * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors of parameter estimates are bootstrapped stardard errors.

Table 3.8: Structural Parameter Estimates: Complete Information Game

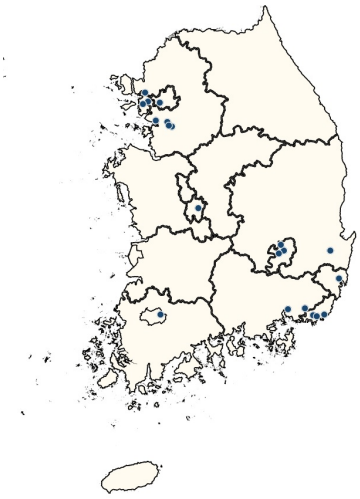
	2010 Census (1)	*2005 Census* (2)
The Effect of Domestic Entry on Foreign Entry	-0.523*	5.732
	(0.295)	(12.909)
The Effect of Foreign Entry on Domestic Entry	-0.272	-1.377***
	(0.332)	(0.236)
Log Population	-0.274	-0.590
	(0.407)	(0.599)
Log Mean Age	-2.267**	-3.900**
	(1.273)	(1.823)
Per-Capita Retail Sale	-0.0524**	-0.218***
	(0.0265)	(0.0815)
Price-Cost Margin	-2.294	-5.576*
	(4.096)	(2.873)
Log Labour Productivity	-0.173	0.936*
	(0.219)	(0.516)
Log Employment	1.065*	1.120
	(0.609)	(0.873)
Log Wage	0.071	-0.509
	(0.408)	(0.596)
Net Entry Rate	0.918	1.103
	(0.818)	(1.237)
Log Store Floor Space	0.126	1.164**
	(0.318)	(0.568)
Log Rent Cost	-0.489	-0.942**
	(0.382)	(0.383)
Log Tax Cost	0.677*	0.719**
	(0.335)	(0.325)
Distance From Foreign (Tesco) Distribution Centre	-0.00105	-0.00575*
	(0.00198)	(0.00314)
Distance From Foreign (Tesco) Distribution Centre	0.00144	0.000402
	(0.00194)	(0.00186)
N	238	215
Log Likelihood	214.470	130.924

Notes: This table presents estimates for structural parameters of entry game-theoretical model between a foreign retail store and a domestic retail store under *complete* information assumption. We apply Berry (1992)'s identification approach to estimating the game of complete information. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors of parameter estimates are bootstrapped standard errors.

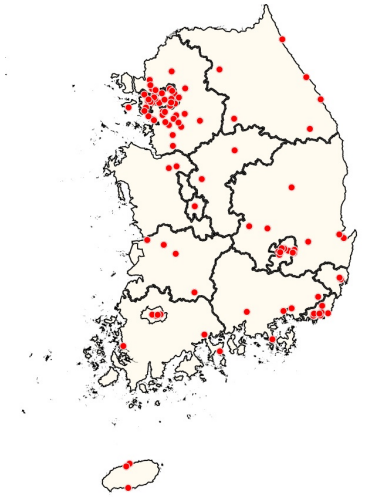
Figure 3.1: Spatial Expansions of Large Discount Stores



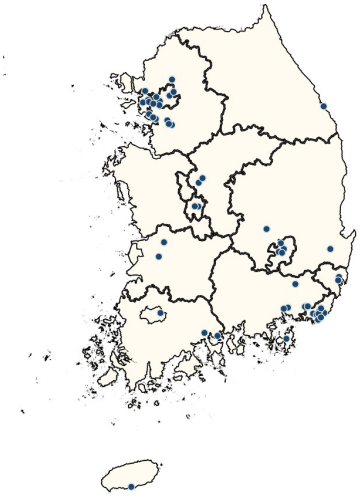
(a) Emart, 2002



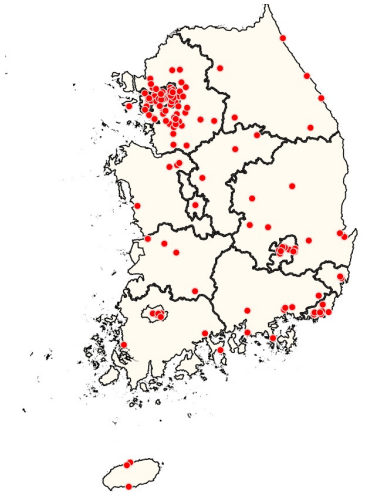
(b) Tesco, 2002



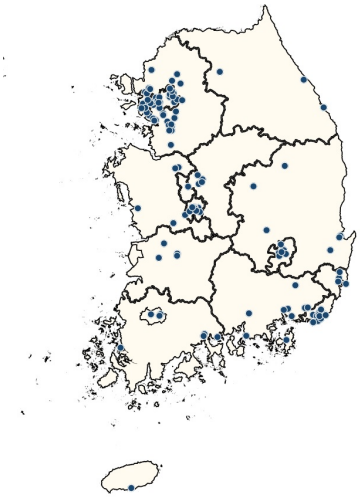
(c) Emart, 2006



(d) Tesco, 2006

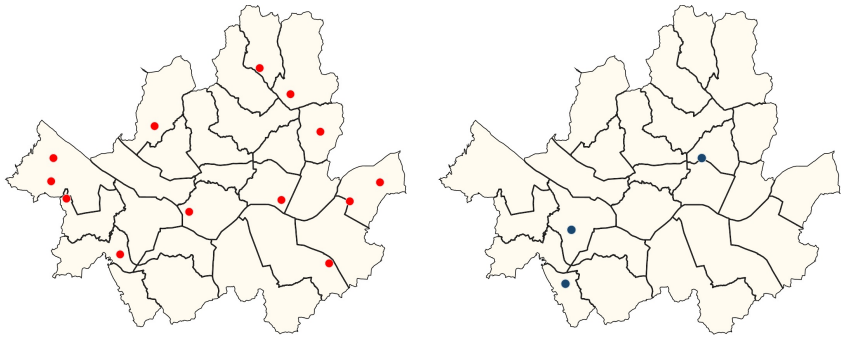


(e) Emart, 2010

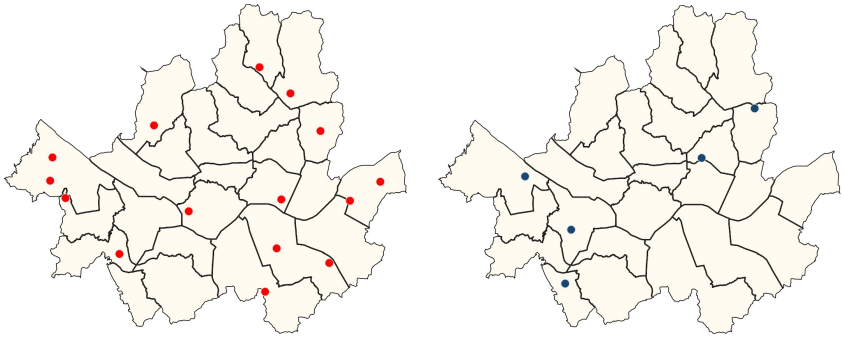


(f) Tesco, 2010

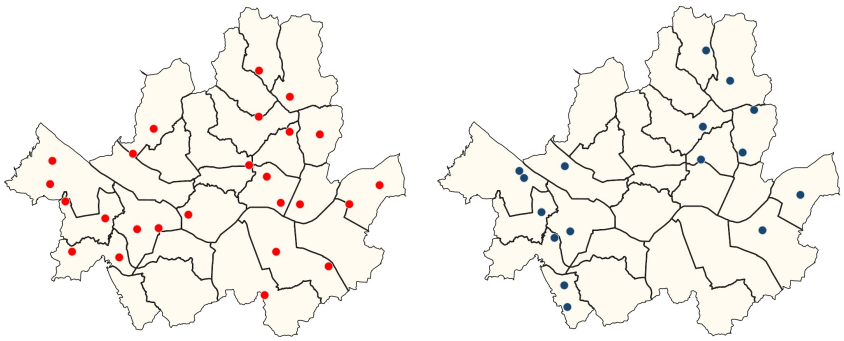
Figure 3.2: Spatial Expansions of Large Discount Stores in Seoul



(a) Emart, 2004 (b) Tesco, 2004

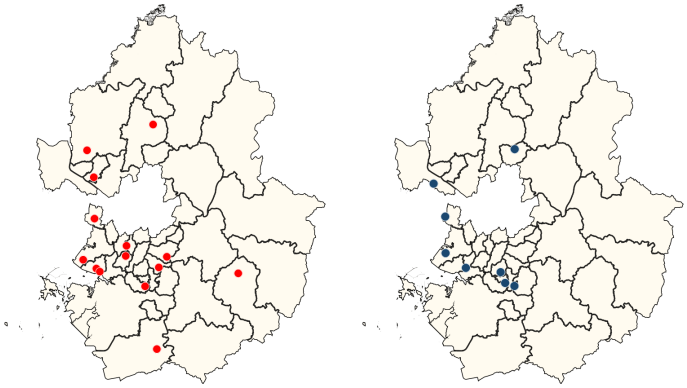


(c) Emart, 2006 (d) Tesco, 2006

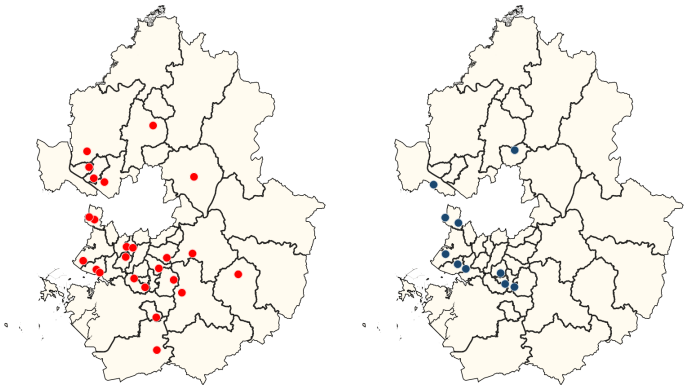


(e) Emart, 2010 (f) Tesco, 2010

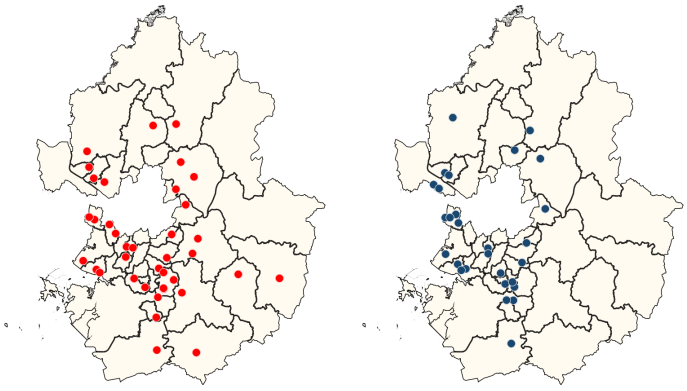
Figure 3.3: Spatial Expansions of Large Discount Stores in Kyunggi



(a) Emart, 2004 (b) Tesco, 2004

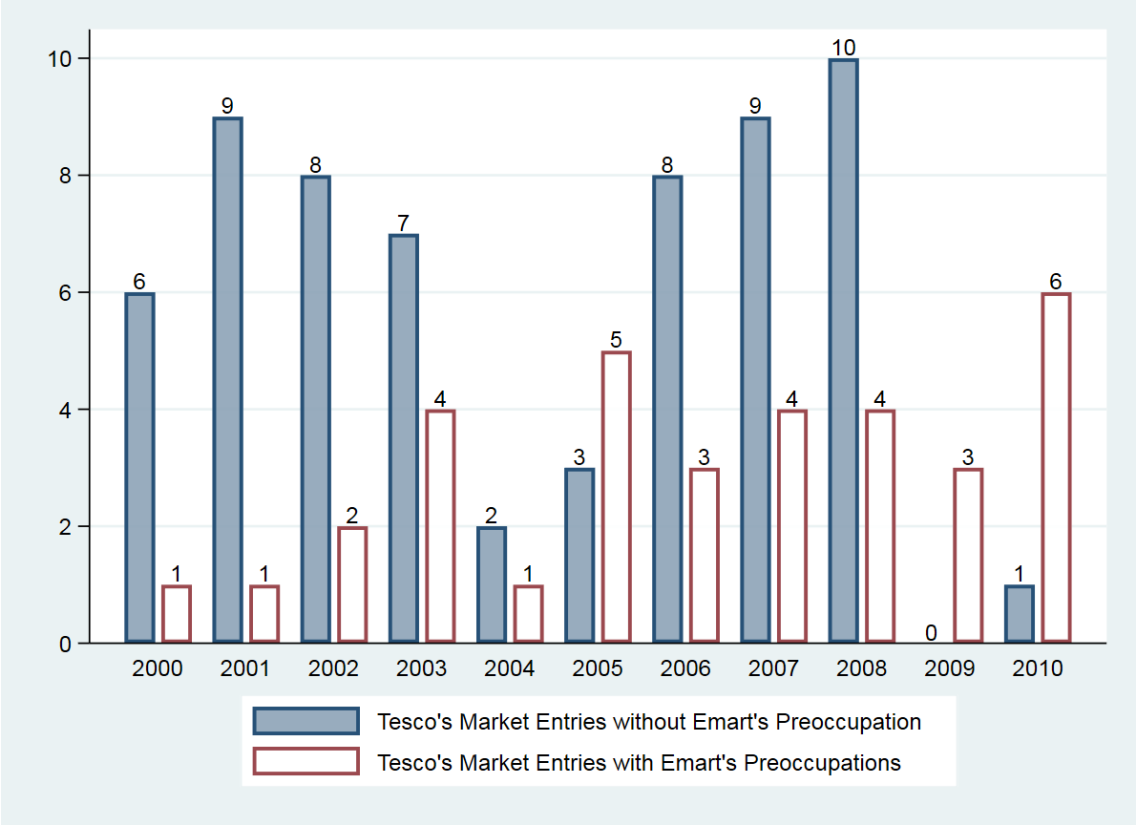


(c) Emart, 2006 (d) Tesco, 2006



(e) Emart, 2010 (f) Tesco, 2010

Figure 3.4: Tesco's Market Entry by Emart's Pre-occupation



Notes: This figure shows the evolution of Tesco's county-level market entries sorted by Emart's pre-occupations. Blue bars display the numbers of Tesco's entries in counties where Emart has not entered. Red bars present the number of Tesco's entries made in (county-level) markets that has not experienced Emart's pre-occupations.

Chapter 4

Do Multinationals Reduce Domestic Markups?

4.1 Introduction

A recent body of papers develop trade models with endogenous markups where trade-induced pro-competition effects are introduced, and empirically show that trade liberalisation leads to reduction in average markup. (e.g., Melitz and Ottaviano (2008), Atkin and Donaldson (2015), Feenstra and Weinstein (2016)) With regard to foreign direct investment (FDI) liberalisation, a similar conjecture would be able to be made: increase in foreign presence enhances the intensity of competition in host economies. Such a competition effect would reduce average markups in host-country markets. While it is intuitively appealing, the impact of FDI liberalisation on markups has rarely been paid attention to both theoretically and empirically.

FDI literature that studies the impact of foreign entry on host-country markets has mainly focused on productivity spillovers. In contrast, the impact of FDI on markups has little been studied. Although a very few number of papers investigate the link between FDI and markups, almost all of them focus on foreign firms' markup premium over firms native to host-market countries (e.g., Muraközy and Russ (2015)). Evidence of the pro-competitive effect of FDI on local firms are scanty. This chapter aims to fill the gap. We assess the impact of inward FDI on average markups of local firms in host-country markets in the context of the Korean retail sector.

FDI in Korea's retail sector was completely liberalised in 1996. By descriptive analyses, we find there has been structural changes in the Korean retail sector since 1996. First, large discount stores operated by multinational retailers were rapidly expanded. The number of foreign stores increases from 7 in 2000 to 121 in 2009. Second, a massive firm churning within the retail sector took place in Korea. Many of small retailers that have lower productivities exited from the market. The number of small retailers was 751 thousands in 1996, but it substantially declines

to 579 thousands in 2010. Third, average price-cost margins fall from 0.199 to 0.102 during the period of 1996 to 2010. While small retailers, many of whom lost market shares or had exited from the market, contributed to reduction in price-cost margins (0.248 in 1996 $\rightarrow$ 0.179 in 2000), we find that price-cost margins for relatively large-scaled retail firms also decline to a considerable degree. Between 1996 to 2010, corporation-type retailers, on average, decreased price-cost margins from 0.115 to 0.051. Lastly, there are significant differences in average markups between markets with and without FDI. A town-level panel regression of average domestic markups, which are proxied by average price-cost margins of local firms, on a retail FDI dummy augmented by town and year fixed effects shows a negative and significant estimated parameter for the FDI dummy. Therefore, these four facets, put together, indicate that arrivals of multinational retailers followed by FDI liberalisation in 1996 intensified competition levels within the sector to a high degree.

Despite the descriptive evidence showing a negative association between retail FDI and markups, identification of the *causal* effect of retail FDI on average markups of local retailers is not straightforward because of difficulty in addressing the selection issue. The selection problem arises if markets receiving retail FDI have lower markups to begin with. Specifically, the markup differentials found by a standard linear regression mentioned above cannot be interpreted as causal consequences of retail FDI, because the approach would not be able to address selection bias.

Our empirical strategy to assess the causal impact of retail FDI on Korean local firms' markups relies on microeconometric evaluation techniques. In framework of Rubin Causal model, selection problem stems from being unable to control for missing counterfactuals. That is, average local markups for treated towns if they had not been treated are unobservable and the missing data results in biased estimates for treatment effects. In this chapter, we apply the inverse propensity-score weighting (IPW) model to construct missing counterfactuals. Once constructing a pseudo-randomised sample by IPW technique, we employ difference-in-difference (DD) regression model on the weighted sample. Since DD regression will eliminate time-invariant unobserved confounders, IPW regression with DD model provides a correct estimate for the treatment effect of retail FDI.

Using retail census data for the years of 2001, 2005 and 2010, we find that on average, retail FDI causes decrease in local retailers' markups. If considering towns receiving FDI from 2002 to 2005, multinational retailers contribute to reduction in domestic mean markups by -0.008% points over nine years (i.e., 2001-2009). Through analyses for shorter time horizons, we also find that the retail FDI intensively affect local markups over the period of 2001 to 2005 (-0.012% points),

whereas the impact seems to be insignificant during the second half of 2000s in Korea. Combining two separate time-horizons into a single panel data, we assess the overall treatment effect of multinational retailers on domestic average markups, and we find that retail FDI causes decline in domestic firm's average markups by -0.007% points.

An important issue for our findings above to be more reliable is that the effect of retail FDI on markups must have nothing to do with FDI's productivity spillovers. Through the lens of combination of IPW and DD model, we evaluate whether retail FDI boosts local retailers' productivity. We find no productivity spillover effects in Korea's retail markets, indicating that the impact of FDI on local markups is purely driven by competition effect. In this study, we define a market for foreign stores as a county following the definition convention in the related literature. Thus, an underlying assumption of town-level analyses is that treatments are spatially homogeneous across towns so long as the towns belong to counties foreign stores enter in. In other words, it means that treatments do not decay with distance within a county receiving retail FDI. Since our results crucially rely on the assumption, we attempt two experiments, via the same method as before, to test if treatment homogeneity assumption holds by data. First, we do the same estimation using data constructed at the county level. As expected, the results are almost identical to the town-level results. Second, we estimate retail FDI's treatment effect on markups by modified data in which we drop towns directly receiving FDI within their own borders from original data. The results of the second experiment are also almost same as the regression with original data.

This chapter is related and contributes to the literature that empirically investigates the pro-competitive effects of globalisation. There are a large body of empirical papers on trade reforms and markups. (e.g., Levinsohn (1993), Harrison (1994), Görg et al. (2003), De Loecker and Warzynski (2012)) On the other hand, there are almost no empirical studies on FDI liberalisations and domestic markups. Sembenelli and Siotis (2008) is an exception. Sembenelli and Siotis (2008) estimate the effect of manufacturing FDI on domestic firms' price-cost margins using a dynamic panel data model. Their results show that while FDI negatively affects markups in the short run, the effect is eventually disappeared in the long run. We add new evidence on pro-competitive effects of FDI in this chapter in the context of the retail sector.

This chapter is closely related to the empirical papers on FDI and its economic consequences on host countries. While the literature studies the impact of FDI on economic outcomes such as wages(Girma and Görg, 2007), employment(Girma, 2005), much more attentions have been given to host country productivity. (e.g., Haskel et al. (2007), Keller and Yeaple (2009), Arnold and Javorcik (2009), Girma

et al. (2015)) Our study differs from the existing papers in two regards. First, we employ methods from microeconometric program evaluation to estimate the indirect effect of multinationals. The methodology is used exclusively to estimate direct impacts of foreign ownerships, whereas it has not been applied when studying indirect effects of FDI. This chapter, however, evaluates FDI's effect on average domestic markups (i.e., indirect effect of FDI) through propensity-score based methods, which is a novel point in this chapter. Second, our study focuses on a new economic outcome variable, *markup*, and is studied under the background of *retail* sector which has rarely been uncovered. To our knowledge, this is the first research that relates retail FDI to domestic average markups.

This chapter is also related to a short strand of works on influences of multinational retail stores over host countries. While previous papers focus on productivity spillovers of retail FDI on local suppliers (e.g., Iacovone et al. (2015), Javorcik and Li (2013)), this chapter studies the effects of multinationals on domestic firms *within* the retail sector. As far as we know, this chapter is also the first study to investigate markup spillovers of multinational retailers towards domestic retailers.

We structure the remainder of this chapter as the following. Section 4.2 describes the background of this research. Section 4.3 explains data sources and descriptive evidence. Section 4.4 presents the framework of our empirical model and the identification strategy. Section 4.5 presents results and section 4.6 concludes.

4.2 Background

Foreign direct investment (FDI) in the Korean retail sector was liberalised in 1996. The retail liberalisation was a result of the Uruguay Round negotiations and Korea's accession to the Organization for Economic Cooperation and Development (OECD) in 1996 (Morrissey and te Velde, 2003). Accordingly, foreign retailers began to enter Korean market since 1996.

While FDI in the retail sector was completely liberalised in 1996, the liberalisation had actually phased over the period from 1993 to 1996. In 1993, the Korean government first allowed multinational retail firms to operate a maximum of 10 stores of 1,000 square metres or less for each foreign-invested company (the Korean Ministry of Trade, Industry and Energy). Although it was the first step towards retail market liberalisation, the incomplete liberalisation gave multinational retailers no incentive to enter the Korean market due to inability to exploit economies of scale and scope.(Kim and Kim, 2003).[1] It was not until 1996 that

[1]On the other hand, zoning regulations, which had long been acted as the entry barrier preventing retail firms from establishing large-scaled stores, was removed only for Korean retail

all of store- and floor space limits which had been applied to FDI were completely eliminated. Since 1996, large discount stores by multinational retailers have been expanded rapidly.

Table 4.1 shows evolutions of the numbers of stores owned by foreign retailers in Korea. In 1996, Carrefour and Makro first entered Korean retail sector. Next, Wal-Mart entered into the Korean retail market by acquiring two Macro's stores in 1998. Tesco and Costco began to operate their stores since 1997 and 1998, respectively. Carrefour and Wal-Mart operated and expanded retail stores up to the mid-2000s, but they divested all their stores in 2006. Carrefour's 30 stores were bought by a Korean retailer, Homever, and Homever sold them on to Tesco in 2008. Wal-Mart's 14 stores were merged by a Korean domestic firm, Emart. As a result of two M&As, there have been in fact only two foreign retailers, Tesco and Costco since 2006. Especially, a U.K. based multinational retailer Tesco has consistently expanded their stores all over the Korean market. Although surviving multinationals decline over years, the total number of foreign retail stores has consistently increased since 1996, as apparently shown by Table 4.1.

The entry of foreign retail stores, which is driven by FDI liberalisation, led to rise in level of competition across firms. Korea's retail sector has experienced high degree of firm churning since 1996. During the period from 1996 to 2001, the total number of entries in retail stores was 89 thousands, whereas the total number of exits is 231 thousands. The gross exits in this period are 2.6 times as many as the gross entrants. This pattern lasted for a long time. The gross entrants and the gross exits of retail establishments from 2001 to 2005 were 106 thousands and 150 thousands, respectively. Though being slightly modest, exits were far more than entrants during this period as well. Most of the negative net entrants resulted from the decline in small-sized retailers. In 1996, the number of retail establishments owned by sole traders (i.e., self- and family-owned retailers) was 751 thousands, whereas it decreased to 579 thousands in 2010. In terms of sales amount, they accounted for 62% of total retail sales in 1996, but their market share fell to 39% in 2010.

The divestment of world's two largest multinational retailers may mirror the high degree of competition in Korea. A study conducted by Suh and Hahn (2004) shows that the maximum number of stores at which Korea's large discount store

firms in 1993. Like liberalisations in other service sectors, the Korean government implemented domestic deregulations and FDI liberalisation at the same time. Domestic deregulation aimed to meet requirements of joining OECD or to prepare for the Uruguay round. In addition, the deregulation implicitly intended to help Korean domestic retailers cope with the retail FDI liberalisation by building market position prior to FDI's entrance. Specifically, in 1993, the semi-agricultural and forest areas were redefined to allow Korean (not foreign) domestic retailers to build less than 30,000 square metres. Due to the deregulation, the first large discount store was opened in Chang-dong, Seoul by a Korean retailer, Emart, in 1993.

market is saturated is estimated to be 400-450. In 2006 when Wal-mart and Carrefour withdrew from Korea, the total number of stores by top five retail firms in Korea already reached to 247. It is just what the discounting retail industry had accomplished only within ten years after the FDI liberalisation. As a result, the rapid expansion and the high competition in the discounting retail industry brought about the 'shake-out' amongst the biggest retail firms, leading to the retreats of Carrefour and Wal-mart from the Korean market in 2006. (Coe and Lee, 2013)[2]

The fiercer competition after FDI liberalisation negatively affect average markups. In other words, it is likely that foreign retail entry brings about pro-competitive effects. Using price-cost margins as a proxy for markups, we find that the average price-cost margin in the Korean retail sector was 0.199 in 2001, but it decreases to 0.102 in 2010. We pointed out that there has been significant firm churning mainly done by massive exits of small-sized retailers. The average price-cost margin across sole traders was 0.248 in 1996, whereas it falls to 0.179 in 2010. Yet, it is unlikely that the decline in market average price-cost margins is solely driven by small retailers. When calculating the average price-cost margin of corporation-type retail stores, the mean price-cost margin is also highly decreasing. In particular, it was 0.115 in 1996, and it falls to 0.051 in 2010. The significant decrease of price-cost margins in corporation-type establishments indicates that they also experienced high-level of competitions by the retail FDI (Kim and Kim, 2003).

4.3 Data and Preliminary analysis

This chapter aims to estimate the causal effect of retail FDI entry on average markups of local firms in host-country markets. To this end, we draw on two sources of data. The first source is a data set on large discount stores collected by the Korea Chainstore Association (KOCA). Since KOCA data contains opening dates and addresses across all large discount store, we establish a town-level treatment dummy variable indicating whether each town receives retail FDI. The second source is Korea's retail census data collected by KOSTAT (Korea's office of national statistics). We construct town-level market characteristics on which we regress a probit model of the retail FDI treatment variable. We create them by aggregating store-level data at the town level. We give full details of the data

[2]A few papers in the business literature have endeavoured to analyse the causes of Walmart's and Carrefour's exits from Korea (e.g., Coe and Lee (2013)). According to those studies, the inability to understand the shopping preferences of Korean consumers and to adjust firm organisations to the prevailing domestic culture are regarded as main reasons. On the other hand, They point out that Tesco who has successfully done so-called strategic localisations has been able to secure the sustained growth in the Korean market.

set in Chapter 3, because Chapter 3 makes use of the same data sources.

Another key variable for this chapter is town-level average markups. As we mentioned in details in Chapter 2, it is quite hard to measure average markup. While new approaches to empirically estimate firm-level markups were recently proposed (e.g., De Loecker and Warzynski (2012)), the method is unsuited to our work as it demands firm-level (balance-sheet) panel data over a relatively long period of time, which is inaccessible in Korea. Rather, we make use of price-cost margins to proxy markups, not only because we can get access to all necessary data from retail census, but also because, as shown by Siotis (2003), the measure is similar to econometrically estimated markups at least at the aggregate level.

In this chapter, we define price-cost margin of store s in town i and census year t as

$$y_{sit} = \frac{p_{sit} - c_{sit}}{p_{sit}} = \frac{(p_{sit} - c_{sit})q_{sit}}{p_{sit}q_{sit}} = \frac{profit_{sit}}{sale_{sit}}$$

where p_{sit}, c_{sit} and q_{sit} are price, marginal cost and sale quantity of store s in town i and census year t, respectively. $profit_{sit}$ and $sale_{sit}$ are store-level profits and sales. To construct town-level markups, we calculate weighted average markups across store-level price-cost margins and we use shares of store-level sales amount in total town sales as weight. It is formulated by

$$y_{it} = \sum\nolimits_s \left(\frac{sale_{sit}}{\sum_s sale_{sit}} \times y_{sit}\right)$$

Figure 4.1 illustrates average price-cost margins in the Korean retail sector over years. As well illustrated in the figure, average price-cost margins that proxy average markups have continuously decreased over years. It is in stark contrast to the fact that the number of foreign stores has highly increased over the same period. Thus, Figure 4.1 may descriptively present that there might have been a pro-competitive effect of FDI on Korean retail markets since the FDI liberalisation in 1996.

Also, Figure 4.2 may potentially suggest the existence of pro-competitive effects in a different angle. In particular, Figure 4.2 illustrates the average markups in the counties where foreign stores are present versus the average markups in the counties without foreign presences. Two facts are noteworthy. First, markups has been reduced irrespective of whether counties receive FDI or not. Second, although every county tends to experience decline in markups, counties with FDI have much lower average markups than counties without FDI. Thus, this figure could indirectly show that retail FDI might be associated with the decrease in average markups with FDI received.

To evaluate the association between retail FDI and markups, we estimate a

simple regression with town-level panel data as follows:

$$y_{it}^{-FDI} = \alpha + \beta D_{it} + \delta_i + \delta_t + \varepsilon_{it} \tag{4.1}$$

where i indexes towns and t implies census years. y_{it}^{-FDI} is town-year specific weighted average markups (i.e., price-cost margins) of all retail stores *other than foreign stores (FDI)*. Recall that the primary goal of this chapter is to investigate whether retail FDI has an impact on markups by host-country's local firms. It is, thus, necessary to analyse *pure* indirect effects of FDI, not to assess the impact of FDI on total effects which include the influence on FDI itself. For the reason, we exclude foreign stores when we calculate market average values, which is indexed by y_{it}^{-FDI}. D_{it} is a dummy variable to indicate whether each town is treated by foreign retail stores. Importantly, the treated towns for this analysis are defined as towns belonging to counties foreign stores enter in. Put differently, our treatment definition implies that we do not restrict treatments to the towns that receive retail FDI within their own borders. In the literature on large discount stores or so-called the *Wal-mart effect* literature, markets for large discount stores tend to be defined as counties. This is because counties are appropriate geographical regions in which large discount stores have market powers. Moreover, firms operating large discount stores make an entry decision on the basis of a county-level unit as explained in Chapter 2. As a result, D_{it} is defined as one if town i is subordinated to the counties in which foreign stores exist at census year t. We will explain detailed discussions around the definition of treatments in section 4.5.3.

Table 4.2 presents the results of the town-level panel regression of Equation (4.1). The first column shows a pooling OLS estimation result where no fixed effects are applied. The estimated coefficient on D_{it}, the retail FDI treatment, is -0.056, of which sign is in line with Figure 4.2. To capture the unobserved town-level heterogeneities, we add the town fixed effect to the OLS regression and column (2) is a estimation result. The estimated effect of retail FDI is also negative and significant as in column (1) even after controlling for the town-level fixed effects. Column (3) show a result only with census-year fixed effect. Still, the estimated coefficient for FDI is negative and significant. Finally, column (4) reports a result where town and census-year fixed effects are all added in the regression. The parameter estimate for retail FDI is significant and negative. Also notable is the fact that the negative effect of foreign stores on markups is the smallest in column (4) as we control for unobserved town- and year-specific factors.

Table 4.2 clearly shows that there is a significant difference in the average markups between towns with and without FDI. In particular, Table 4.2 indicates that markets foreign retailers enter in exhibit, on average, lower markups by 0.005

% points than markets only composed of Korean retailers (Column (4)). Yet, the result cannot be evidence that retail FDI *causes* a reduction in mean markups if markets undergoing foreign entry have lower markups *to begin with*. In other words, markup differentials between markets with and without multinationals are difficult to be interpreted as causal evidence when foreign retailers tend to enter markets in which average markups are originally low.

For instance, as documented in recent models of international trade and economic geography, it is predicted that market size is (negatively) associated with markups due to toughness of competitions (e.g., Melitz and Ottaviano (2008), Behrens and Robert-Nicoud (2014)). At the same time, retailing literature has emphasised that market size is one of the most important determinants on where retail firms open stores. Put together, markups may matter for multinationals' entry decision, thus not enabling *selection into* (foreign retail entry) *treatment* to be random. In conclusion, it is critically important for this study to control for the selection bias.

4.4 Identification

4.4.1 Empirical Strategy

A fundamental problem when estimating causal effects of retail FDI on average markups of local stores relates to the selection issue, as mentioned in section 4.3. Our empirical strategy to assess causal impacts of the entry of multinationals on domestic retail firms relies on microeconometric evaluation techniques. In terms of a potential outcome framework, selection problems stem from the fact that it is impossible to observe so-called *counterfactuals*. In context of our study, average markups for treated towns if they had not been treated are unobservable. Failure to control for the missing data results in biased estimates for treatment effects. Treatment effects evaluation literature has suggested regression techniques by which missing counterfactuals are constructed and correct treatment effects are estimated. Among them, this chapter makes use of the inverse propensity-score weighting (IPW) model, one of propensity-score based methods in order to control for the selection bias.

In fact, propensity-score based methods are able to address the selection issue relying on the conditional independence assumption (CIA), i.e., treatment is randomly assigned conditional on observable confounders. On the other hand, they are unable to deal with selection on unobservables. As Heckman et al. (1997) point out, combining propensity-score based methods with difference-in-difference regression techniques can not only control for the selection bias but also remove

time-invariant unobservable confounders. We apply the approach in this chapter, focusing on the difference in mean markups after the state of towns is transformed from no foreign-treated markets to foreign-treated markets and comparing the differenced performance with towns that do not experience foreign retail entry.

4.4.2 Diff-in-Diff Inverse Propensity-Score Weighting estimator

As before, let treatment variable, D_{it}, be a dummy variable that equals one if foreign retail stores enter a town i at census year t, and zero otherwise. Now, let $y_{it}(1)$ denote weighted average markup of domestic stores when town i is treated at census year t.[3] Also, $y_{it}(0)$ indexes the weighted average markup if town i had not been treated. For each town, the causal impact of multinational retailers' entry is defined as the difference between two potential outcomes like

$$y_{it}(1) - y_{it}(0)$$

and the average treatment effect on the treated towns (ATT) is, by definition,

$$E\big[y_{it}(1) - y_{it}(0)|D_{it} = 1\big] = E\big[y_{it}(1)|D_{it} = 1\big] - E\big[y_{it}(0)|D_{it} = 1\big] \tag{4.2}$$

The main problem when estimating ATT is that we cannot observe $y_{it}(1)$ and $y_{it}(0)$ for each town at the same time. Likewise, it is impossible to directly calculate the right hand side of Equation (4.2), because $E\big[y_{it}(0)|D_{it} = 1\big]$ is unobservable. One may think that $E\big[y_{it}(1)|D_{it} = 1\big] - E\big[y_{it}(0)|D_{it} = 0\big]$ is used for estimating ATT instead of Equation (4.2). The 'naïve' difference-in-mean estimator, however, leads to biased results, because it is decomposed into ATT and selection bias as the following:

$$\begin{aligned} & E\big[y_{it}(1)|D_{it} = 1\big] - E\big[y_{it}(0)|D_{it} = 0\big] = \\ \text{(ATT)} \quad & E\big[y_{it}(1)|D_{it} = 1\big] - E\big[y_{it}(0)|D_{it} = 1\big] + \\ \text{(Selection bias)} \quad & E\big[y_{it}(0)|D_{it} = 1\big] - E\big[y_{it}(0)|D_{it} = 0\big] \end{aligned} \tag{4.3}$$

and, as discussed above, the selection bias should not be ignored in the context of foreign retail entry.

We exploit the inverse propensity-score weighting regression (IPW) model to address the selection issue. The bottom line of IPW estimator is that selection bias in Equation (4.3) is eliminated by constructing an artificial or pseudo-randomised

[3]Previously we termed weighted average markup of domestic stores as $y_{it}^{(\text{-}FDI)}$. For the notational convenience, we will, throughout the chapter, drop (-FDI) from now on.

population in which outcomes are independent of treatment assignments (Horvitz and Thompson (1952), Robins et al. (1994)). In principle, IPW model weights each observation by imposing the inverse of its propensity score, through which selection into treatment disappears in the weighted data.

Let X_{it} denote observed covariates which collects all information on controls explaining the treatment variable D_{it}. Then, the propensity score, the probability of being treated conditional on X_{it} is defined as

$$p_{it} = P(D_{it} = 1|X_{it})$$

Rosenbaum and Rubin (1983) show that if potential outcomes are independent of treatment conditional on covariates (i.e., CIA assumption holds) they are independent of treatment conditional on the *propensity score*. This can be formulated by

$$y_{it}(1),\, y_{it}(0) \perp D_{it}|X_{it} \Leftrightarrow y_{it}(1),\, y_{it}(0) \perp D_{it}|\, p_{it}(X_{it}) \tag{4.4}$$

where X_{it} is a vector whereas $p_{it}(X_{it})$ is a scalar, implying that the Rosenbaum and Rubin (1983)'s property deals with the dimensionality problem.

The average treatment effect on the treated towns (ATT) is, by definition,

$$\begin{aligned}\lambda &= E\big[y_{it}(1) - y_{it}(0)|D_{it} = 1\big] \\ &= E_{X_{it}}\big[E\big[y_{it}(1) - y_{it}(0)|D_{it} = 1, X_{it}\big]\big] \\ &= \underbrace{E_{X_{it}}\big[E\big[y_{it}(1)|D_{it} = 1, X_{it}\big]\big]}_{(A)} - \underbrace{E_{X_{it}}\big[E\big[y_{it}(0)|D_{it} = 1, X_{it}\big]\big]}_{(B)}\end{aligned} \tag{4.5}$$

Note that Equation (4.5) is derived by the law of iterated expectation. Expression (A) above can be transformed as the following

$$E_{X_{it}}\big[E\big[y_{it}(1)|D_{it} = 1, X_{it}\big] = E_{X_{it}}\big[E\big[y_{it} \cdot \mathbf{1}\{D_{it} = 1\}|X_{it}\big] = E\big[y_{it} \cdot \mathbf{1}\{D_{it} = 1\}\big] \tag{4.6}$$

where $\mathbf{1}\{D_{it} = j\}$ is a dummy indicator which equals to one if $D_{it} = j$ for $j = 1, 0$, and zero otherwise. The last term of Equation (4.6) is obtained by integrating over X_{it}. Using the CIA and the property by Rosenbaum and Rubin (1983), Expression

(B) in Equation (4.5) is solved as

$$\begin{aligned} E_{X_{it}}\big[E\big[y_{it}(0)|D_{it}=1,X_{it}\big]\big] &= E_{X_{it}}\left[\frac{E\big[y_{it}\cdot\mathbf{1}\{D_{it}=0\}|D_{it}=1,X_{it}\big]}{1-p_{it}(X_{it})}\right] \\ &= E\left[\frac{\big[y_{it}\cdot\mathbf{1}\{D_{it}=0\}|D_{it}=1\big]}{1-p_{it}(X_{it})}\right] \\ &= E\left[\big[y_{it}\cdot\mathbf{1}\{D_{it}=0\}\big]\frac{p_{it}(X_{it})}{1-p_{it}(X_{it})}\right] \end{aligned} \tag{4.7}$$

where $1-p_{it}(X_{it}) = P(D_{it}=0|X_{it})$ which means the conditional probability of being untreated. Finally, the ATT in Equation (4.5) can be calculated as

$$\lambda = E\Big[y_{it}\cdot\mathbf{1}\{D_{it}=1\}\Big] - E\left[\big[y_{it}\cdot\mathbf{1}\{D_{it}=0\}\big]\frac{p_{it}(X_{it})}{1-p_{it}(X_{it})}\right] \tag{4.8}$$

Relying on ATT expression in (4.8), the IPW estimator of ATT can be given by (Busso et al., 2014)

$$\hat{\lambda}_{IPW} = \frac{1}{n}\sum_i\left[D_{it}y_{it}\right] - \frac{1}{n}\sum_i\left[(1-D_{it})y_{it}\frac{\hat{p}_{it}}{1-\hat{p}_{it}}\right] \tag{4.9}$$

Clearly as shown above, the IPW estimator of ATT is the difference between the sample means of the treated and the untreated groups in the weighted data. In practice, the IPW ATT estimator can be regarded as a method-of-moments estimator based on ATT $= E(w_{it}y_{it})$, where the weight, w_{it} is as below

$$w_{it} = D_{it} + (1-D_{it})\frac{\hat{p}_{it}}{1-\hat{p}_{it}}$$

Expression (4.9) can be improved by reweighting process. (e.g. Imbens (2004), Busso et al. (2014)) Reweighting estimator suggests renormalizing the weights so that the sum of the weights can be one. Therefore, $\hat{\lambda}_{IPW}$ is transformed into

$$\hat{\lambda}_{IPW} = \frac{1}{n_1}\sum_i\left[D_{it}y_{it}\right] - \frac{1}{n_2}\sum_i\left[(1-D_{it})y_{it}\frac{\hat{p}_{it}}{1-\hat{p}_{it}}\right] \tag{4.10}$$

where

$$n_1 = \sum_i D_{it} \qquad n_2 = \sum_i (1-D_{it})\frac{\hat{p}_{it}}{1-\hat{p}_{it}}$$

Although the ATT estimator expressed by Equation (4.10) is expected to eliminate the selection bias and to provide unbiased results, it is only valid if there are no unobservable confounders. In this sense, we combine a difference-in-difference

method with the IPW regression model. As discussed earlier, this strategy leads to a correct result by addressing unobserved confounders as well as selection problems. The difference-in-difference (DD) IPW estimator of the ATT is defined as follows:

$$\hat{\lambda}_{DDIPW} = \frac{1}{n_1}\sum_i \left[D_{it}\Delta y_{it} \right] - \frac{1}{n_2}\sum_i \left[(1 - D_{it})\Delta y_{it}\frac{\hat{p}_{it}}{1-\hat{p}_{it}} \right] \tag{4.11}$$

where Δy_{it} is the first difference between the pre-treatment and the post-treatment weighted average markups. The only change from the original IPW estimator is that the outcome variable is a differenced term now.

Let $\Delta {y_{it}}^t$ denote the first difference of a outcome variable between before and after treatment in the *treated* group. On the other hand, $\Delta {y_{it}}^c$ implies the first difference of y_{it} in the *untreated* or *control* group. Then, the double difference between two terms is, by definition, a DD estimator.

$$\hat{\delta}_{DD} = \Delta {y_{it}}^t - \Delta {y_{it}}^c$$

In terms of a regression specification, the DD estimator can be defined as a regression result of the following model

$$\Delta y_{it} = \alpha + \hat{\delta}_{DD} D_{it} + \varepsilon_{it} \quad \text{for } j = 1, 0 \tag{4.12}$$

Therefore, the DDIPW estimator is obtained by DD estimation with the *weighted* data. To put it another way, using the difference of outcome variables rather than level values when doing the IWP estimation will lead to a result of the DDIPW estimator as formulated by Expression (4.11).

4.4.3 Constructing Pseudo-Randomised Sample

The DDIPW (as well as IPW) estimator weights observations on the outcome variable to account for the missing selection process. The weighting scheme intends to generate a *pseudo-randomised* sample in which the treatment and the control groups are ex-ante equally likely to be treated, meaning that the treatment becomes random. In the context of this chapter, towns with and without foreign entry are, on average, ex-ante equally likely to be chosen as the local market where multinational retailers enter. Critically, it is important to estimate propensity scores rightly, because the weights used in the DDIPW estimator are built on propensity scores as specified in section 4.4.2.

To obtain propensity scores, we estimate a probit model of a binary variable which takes on the value of one when town i receives foreign retail entry at time

t. On the other hand, the outcome variable equals to zero when there has not appeared a foreign retail store in town i at time t. The explanatory variables are exogenous market characteristics that we construct from the Korean retail census data. To eliminate any possibility that treatment itself influences market characteristics, we use one-period lagged $(t-1)$ market characteristics as explanatory variables of the probit model following the treatment effects estimation literature.

Recall that the census has been undertaken every five (or four) year since 1996. The available retail census data is that of years 1996, 2001, 2005 and 2010. It is also worth considering that many of critical variables are not surveyed in the 1996 retail census. Due to the structure and the limitation of data, we define town i as being treated if foreign stores exist in the 2005 data but not in the 2001 data. To be more specific, we define the treatment as a town in which multinational retailers open stores between 2002 and 2005. Naturally, the pre-treatment (census) year, which we denote as $t-1$ above, is 2001 in this analysis.

Our town specific characteristics include PCM, PCM growth, labour productivity (sales divided by employment), labour productivity growth, employment(total number of workers in the retail sector), employment growth, wage(proxied by labour cost divided by employment), wage growth, net entry rate(net entry(= gross entry - gross exit) divided by total store number), store floor space, rent cost, tax cost and an indicator whether Korean large discount stores (previously) exist. As noted above, all of the level variables are one period lagged values $(t-1)$, and growth terms aiming to capture pretreatment trends are measured by the growth rates between $t-2$ and $t-1$. If appropriate, level variables are taken a logarithm and are measured in constant Korean wons with base year 2010. The results of probit model are presented in Table 4.3.[4] We compute a town's propensity score using the probit estimates.

Once propensity scores are obtained, it is very important to ensure that the treatment is random in the weighted data, which is called balancing test in the literature. Balancing test aims to check if the weighted distribution of covariates are approximately equal between the treatment and the control groups. Table 5 reports standardised differences in weighted means and standardised differences

[4]Probit results report that the entry of foreign stores is unlikely to be affected by previous market averages of PCM, productivity, wage. Previous growth rates of those variables do not appear to be significant either. Results show that past growth of employment is associated with foreign store's entry, while previous employment level is not likely to affect foreign entry. Market characteristics relating to tax and rent costs seem to be determinants of multinational retailers' entry in Korea. While the effect of tax costs is estimated to be negative, rent fee's effect is not. Lastly, probit model reports that foreign retailers are likely to be affected by previous history of Korean large discount stores' entry. Note that the dependent variable is a dummy variable indicating whether foreign stores exist in 2005 and independent variables are the ones four years past. The weak associations in some of explanatory variable may be influenced by the relatively long time discrepancy.

in weighted ratio of variances adjusted by the IPW model. The second column presents the difference in means between the treatment and the control groups when sample is unweighted. The differences in *weighted* means with the weighted data, which is presented in the third column, is much smaller than those in unweighted data across all variables and even seem to be negligible. Variance ratios in the weighted data in column 5 are all near one, which is in a stark comparison with unweighted results in column 4 as well. To our understanding, there is no statistical criteria to inform the size of standardised differences and variance ratios by which one can judge whether balance tests are passed or not. Nonetheless, Table 5 apparently shows that the difference dramatically decreases and that ratio of variances becomes close to one in a weighted data.[5]

4.5 Results

4.5.1 Treatment Effects on Average Domestic Markups

To estimate Difference-in-Difference model, it is required to have information on at least three periods of time. As described before, Korean retail census data contains information at the five (or four) year intervals (i.e., census years: 1996, 2001, 2005, 2010) and the 1996 census data has no information on some key variables. Inevitably, we construct DDIPW estimator relying on 2001, 2005 and 2010 census data.

Table 4.5 informs how three periods of time serve for estimating ATT and how we define the treatment and the control groups using retail census data for the three periods of time. First, the treatment is defined as town-level markets in which FDI is appeared (Retail FDI(D_{it}) = 1) in 2005 and 2010 census data but not (Retail FDI(D_{it}) = 0) in 2001 data. This implies that treated towns must receive retail FDI during the period of 2002 to 2005.[6] Next, control group is defined as a set of towns that have never experienced foreign retail entry from 2001 to 2010. Pre-treatment market characteristics by which town-level propensity scores are estimated are constructed from 2001 census data. Lastly, post-treatment outcomes and pre-treatment outcomes are built on 2001 and 2010 data respectively.

Table 4.6 provides estimation results of treatment effects (ATTs) on differ-

[5]As a second test, we apply a balancing test method which was proposed by Imai and Ratkovic (2014). Imai and Ratkovic (2014) derived a balance test statistics by considering the restriction imposed by balancing as overidentifying conditions. Doing the overidentification test, we find that we cannot reject the null that covariates are balanced between treatment and control groups.

[6]Note that, if a town receives FDI between 2006 and 2010, the town is not considered as the treated according to the definition above. It is because we do not have any data after 2010 and thus cannot construct *post-treatment* outcomes.

ence in average domestic markups (price-cost margins) between pre- and post-treatment periods. Column (1) reports the unweighted difference-in-difference (DD) estimate. Meanwhile, column (2) presents the DD estimate with weighted data, i.e., DDIPW estimate formulated by Equation (4.11). Regarding the regression results in the unweighted sample, the impact of foreign store treatment on markup differences is estimated to be statistically significant and positive (0.026). Notwithstanding, the DD model without weighting does not deliver reliable results if the treatment is not randomly assigned. Thus, the result of DDIPW estimator in column (2) is of utmost interest for this study. Column (2) points out that retail FDI negatively affect the difference of domestic mean markups between 2001 and 2010. As for the magnitude, the estimate indicates that the markup differential is lower in foreign-treated towns than in non-treated towns by 0.008 % points.

The DDIPW result suggests that foreign store entry causes fall in average markups of domestic stores. However, one may argue that the result is not sufficient enough to be interpreted as the *conclusive* evidence of treatment effects, expressing doubts by two perspectives: First, it only answers whether the towns treated from 2002 to 2005 are affected by FDI. Conversely, it does not answer anything about towns treated at the second half of 2000s. Second, even though we know that 9-year difference of markups reacts to foreign retail entry, we do not know what happens in the middle of the period.

To find further evidence as indisputable as possible, we repeat the experiment in which identification strategy is same as before, whereas definitions of *treatment periods* are differently applied. We already mentioned that the definition of treatment illustrated by Table 4.5 is unavoidably made due to the structure of our data set. However, a good news is that KOCA(Korea Chainstore Association) data, which is the source of town-level retail FDI treatment variable, is provided as an annual data. This implies that KOCA data allows me to construct a new retail FDI dummy variable of which treated periods are *in-between* two adjacent census years.

Taking advantage of the attribute that KOCA data exists on a yearly basis, we redefine the treatment as towns in which foreign retailers open stores *from 2002 to 2004* so that we can estimate treatment effects using shorter time horizon. In the new definition, we exclude the year 2005 from the treatment period, because it allows for exploiting the year of 2005 as the new post-treatment period, leading to the analysis built on the shorter time horizon. Table 4.7 illustrates how we set pre- and post-treatment periods, and how we define the treatment and the control groups in this case.

Table 4.8 shows another case in which we consider treatment group as a set of towns where multinational retailers open stores in Korea *from 2006 to 2009*.

KOCA's annual data allows us to utilise a new definition of treatment period focusing on the second half of 2000s. Accordingly, pre-treatment period (t-1) and post-treatment period (t-1) are 2005 and 2010 respectively in this case.

In Table 4.9, we present estimation results when we applied the treatment definitions specified in Table 4.7 and/or Table 4.8. First, column (1) aims to answer a question of what happens to the result when we account for both treatments occurring from 2002 to 2004 (Table 4.7) and those happening from 2006 to 2009. To answer it, we firstly construct a cross-section data depending on the case of Table 4.7 and do the same thing based on Table 4.8. Then, we append two data sets so that we can build a town-level panel data in which there are two periods of time. Column (1) is the DDIWP estimation result of the panel data analysis. Arguably, column (1) thus shows how we address the first concern that the estimation results based on the baseline treatment definition are limited. Column (1) apparently presents that, despite further accounting for towns treated during the period of 2006-2009, the treatment effect is almost identical to the baseline results in Table 4.6, indicating that the entry of multinationals causes fall in local retail markups all over the 2000s.

Next, column (2) reports estimation result for the case of Table 4.7, and column (3) shows the result of Table 4.8. They targets at answering a question of whether treatment effects matters with markup differences in shorter time horizon, not nine-year difference like Table 4.5. Hence, they are relevant to the second counterargument mentioned above, which is that we do not know what happens in the middle of the period from 2001 to 2010. Column (2), the result when we use a markup difference between 2005 and 2001, and 2002-2004 treatment (Table 4.7), shows that the estimate is highly significant and the negative effect is bigger (-0.012) than when we exploit nine-year difference of markup as the dependent variable (-0.008). On the other hand, when exploiting the markup difference between 2010 and 2005 and the 2006-2009 treatment (Table 4.8), the DDIPW regression provides an insignificant estimate as reported in column (3).

The estimation results reported in Table 4.9 overcome the limitation of the baseline model in Table 4.6. Using modified definitions of treatment periods, we firstly find that foreign stores negatively affect domestic mean markups regardless of whether we consider long-term time perspective or relatively shorter-term time perspective. On top of that, even if treatment effects exist over the period of 2001-2010, our findings suggest that the treatment effect seems to occur intensively at the first half of 2000s.

The result in Column (1) suggests that the entry of foreign retail store lowers the differentials of average local markups between 2001 and 2010 by 0.008% points, compared to the case where it had not been for the foreign entry. Even

though the impact is economically significant at the 1% level, it is not easy to determine how large the pro-competitive effect of foreign retail entry is. Above all, there is almost no empirical evidence on the pro-competitive effect of FDI in the literature, which makes us less able to compare our results with others. To our understanding, Sembenelli and Siotis (2008) is an exceptional empirical study showing evidence on the pro-competitive effect. According to their results, foreign presence in a sector leads to reduction in average markups by 0.019% points in the short run. Seemingly, there is to some degree of discrepancy between ours and the result of Sembenelli and Siotis (2008) in terms of the magnitude. However, it is hard to directly compare the magnitudes because our analysis is conducted focusing on the retail sector whereas Sembenelli and Siotis (2008) estimate the pro-competitive effects of FDI based on a comprehensive of sectors. Nevertheless, our result clearly shows that even if limiting to the retail sector FDI brings about the pro-competitive effect on local firms.[7]

4.5.2 Do Productivity Spillovers Influence Treatment Effects?

In this subsection, we empirically investigate the impact of foreign retail entry on domestic mean productivities through the lens of ATT described earlier. In particular, we use the same regression model, the DDIWP estimator to find the ATT on average productivity of domestic firms. The productivity measure that we use in this subsection is labour productivities defined as sales divided by employments. Researchers who study FDI have paid attention to so-called productivity spillovers. Productivity spillovers imply that once multinational enterprises are located in host-country markets, domestic firms adjacent to the foreign firms can enjoy productivity enhancements from the existence of FDI (e.g., Aitken and Harrison (1999), Girma et al. (2015), etc.). The theoretical motivation about why productivity spillovers arise is that domestic firms can learn from foreign firms. For instance, domestic retailers are able to improve their productivities by observing and imitating foreign retailers' advanced management know-how such as displaying ability and customer services.(Blomström and Kokko, 1998)

When it is confirmed from our data that spillover effects or treatment effect of FDI on domestic stores' productivities exist, the estimates on markup effects in the previous section may be incorrect because productivity improvement by spillovers

[7]I noted in Section 4.1 that there are some empirical papers showing the markup effects of trade or openness. Because discrepancy in both markup measures and datasets chosen, it is also tricky to compare the pro-competitive effects between our results with the markup effects in the trade literature. Nevertheless, most of studies in the trade literature document that openness up to trade or removal of tariff barriers leads to reductions in markup.

would influence markups of domestic stores and thus decrease in domestic mean markups. The mechanism behind this argument that markups are associated with productivities at the firm level is as follows: Increase in productivity will reduce marginal costs, thus leading to rise in price-cost markups. In turn, the change in average markups after FDI treatment may be entangled by not only competitive pressures but also productivity spillovers. Therefore, if possible, it is important to control for the productivity component on the ATT estimation. As noted before selection on treatment via productivity has been already controlled by including market average productivity as one of pre-treatment market characteristics into the propensity-score estimation. Nonetheless, what we am concerned is that FDI would affect markups and productivities contemporarily. If so, the negative effect of FDI on domestic mean markups is likely to be underestimated as foreign retail entry would affect markups and productivities in the opposite direction. That is the reason why we check if productivity spillovers are present in Korea.

Table 4.10 reports the estimation results of whether foreign store entry causes changes in average productivities across domestic stores within a market. Column (1) reports estimation results of regression specifications in the baseline model of Table 4.5. On the other hand, columns (2) - (4) show the results based on treatment definitions noted in Table 4.7 and/or Table 4.8. Significant estimate of the DDIPW model is only appeared in column (1) whereas there are all insignificant estimates in other columns. As discussed in the previous subsection, column (1) might be insufficient to be complete evidence of productivity spillovers, because it excludes towns treated after 2006 and rests on a relatively long-term difference of outcome variable. In the regard, columns (2)-(3) should be complemented to judge the productivity spillover hypothesis of our interest. Unlike the markup effects, the DDIWP estimates in columns (2)-(4) show that there are no evidence on the existence of productivity spillover effects. Relying on the findings above, we draw a conclusion that the markup effects of retail FDI are purely attributed to increase in competitive pressures following the entry of foreign retailers.

4.5.3 Market Definition and Treatment homogeneity

Figure 4.3 is a map of a Korean county, Eun-Pyeong Gu. As illustrated, there are sixteen subordinate towns in the county. A foreign retail store is located on one of the towns, Deajo-dong, which is marked by a grey asterisk. In section 4.4.3, we explain that, in this chapter, the treated towns are defined as every town belonging to the counties FDI stores enter in. Coming back to the example of Enu-Pyeong Gu, the treated town is not only Deajo-dong but also all other fifteen subordinate towns of Enu-Pyeong Gu.

In terms of our empirical model, it can be said that, as long as a town is subordinate to a county FDI enter in, treatment happens homogeneously across all towns within the county, no matter whether the town is directly settled by FDI within its own border or not. This means that there is no friction on treatments within a county. In other words, treatments do not decay with distance within a county.

The treatment homogeneity assumption eventually relates to the question of what is the geographical market definition of foreign stores. Suppose relevant geographical market is a narrower region than a county. Then the entry of foreign stores would not influence pricing behaviours of domestic stores located out of the market. It means treatment cannot be homogeneous within a county, which is forced to modify the treatment assumption above. Thus, the treatment homogeneity assumption requires relevant market definition for foreign stores to be a county.

Literature on large discount stores[8] mostly sets market definitions as counties. (e.g., Jia (2008), Cohen and Mazzeo (2007)) Even restricting to studies on Korean retail sector, most papers also suggest that geographical scope of large discount stores is a county. Jeon and Whang (2010) suggest that geographical market delineation of large discount stores is defined as 20-minutes driving time. According to Lee et al. (2014), 20-minutes driving time is able to be converted into about 5 kilometre's distance. Pointing out that the distance corresponds well to average radius (5.5 kilometres) of Korean counties, Lee et al. (2014) argue that a county is an appropriate market definition for large discount stores in Korea. Recent examples of antitrust geographical market definitions on large discount stores have also presented the diameter of 5 kilometre's distance from the stores.[9] In view of consumers, the treatment homogeneity would hold if within-county consumers shop at large discount stores regardless of distances. In particular, Kim et al. (2002) analyse that diffusion of cars in Korea expands consumers' available shopping area, which makes consumers easily access to relatively remote large discount stores.

Our analyses crucially rest on the assumption about homogeneous treatments within a county. Thus, it is important to test whether the assumption holds by data, otherwise some may think the estimation results so far are questionable. To attenuate the concern, we attempt two experiments via the same methodology that we have used throughout this chapter (i.e., DDIPW estimator). The first experiment is to estimate FDI treatment effect on domestic average markups

[8]Multinational retailers who enter Korean markets establish and operate only a specific store format, large discount stores. In this sense, market definition of foreign stores and that of large discount stores are used interchangeably.

[9]See for example, KFTC (Korea Fair Trade Commission) Decision 2008-285.

with the county-level data. We construct county-level data which contains same variables as before, but all of them are county-specific variables. The dependent variable becomes county average markup across all Korean stores other than foreign ones. The treatment variable is defined as county-level dummy indicating whether or not foreign stores enter the county. Of course, market characteristics used for estimating propensity scores are all county-level variables.

Table 4.11 is results of county level analyses. Column (1) shows estimation results of baseline regression model that we specified via Table 4.5 . Columns (2)-(4) are results when we consider different periods of treatment. In addition, column (1) is county-level counterpart of column (2) in Table 4.6, and columns (2)-(4) mirror columns (1)-(3) in Table 4.9. If treatment is heterogeneous across towns, how would county-level results be estimated? What we expect is that county-level results should be significantly different from the previous town-level results in which treatments are assumed to be homogeneous. However, Table 4.11 definitely shows that both county-level and town-level effects are almost identical. Particularly, magnitudes as well as significance of every estimate make no difference compared to town-level results.

The second empirical experiment is to estimate ATTs with modified data in which towns receiving retail FDI within its own borders (e.g., Deajo-dong for Eun-Pyeong County) are excluded. Suppose a geographical market definition for large discount stores is not a county but a town. The reasonable conjecture would be that ATT estimate of the second experiment should be statistically insignificant or at least quite different from the original results. If this is confirmed by data, treatment homogeneity assumption should be rejected. Table 4.12 reports estimation results with the modified data that we drop towns in which foreign stores exist within their own borders. As expected, ATT results are very robust even in this data. Estimated coefficients and their significance levels are almost same as the original results and the first experiment. This would critically indicate that defining foreign stores' market as a county is suitable and that treatments happen homogeneously within a county.

4.5.4 Korean Large Discount Stores

The prevalence of large discount stores as one of distinct features after FDI liberalisation in Korea, because multinationals establish and operate only a new retail store format, large discount stores. The feature also significantly results from continuous exertions of a few Korean conglomerates ('Chaebol') who adapt the new store format and establish large discount stores over the Korean markets. The Korean firms' endeavours aim to react to market penetrations by multinationals

and to protect Korean markets from FDI. In the meanwhile, large discount stores have dominated the entire retail industry over the two decades in Korea, and both the foreign and the Korean large retail firms are main players in the discount retailing industry. Therefore, in our study, there may be concern that the results on markup effects so far are jointly determined by foreign and Korean large discount stores, not solely by foreign large discount stores.

Table 4.13 is estimation results of treatment effects (ATTs) of Korean large discount store (KLDS) entry on market average markups. Recall that outcome variable used so far is domestic mean markups which seeks to figure out the indirect markup effects of foreign retail entry. However, the current market average markups used in this subsection are defined as town-level average markup across all stores other than *Korean* large discount stores. Also, treatment is defined as towns in which KLDS exists. That is, the treatment variable here is a dummy having values as one if KLDS is present. Table 4.13 thus shows whether KLDS's entry influences market average markups *after* direct impacts of KLDS on itself are excluded. With regards to methodology, we exploit the same model, the DDIPW regression, to find the ATT on market average markup.

The results in Table 4.13 suggest that the issue is not a concern in the context of Korean retail sector. Regardless of treatment definitions and treatment periods, every ATT estimate in columns (1) - (4) indicates that there is no effect of KLDS on average market markup. Some might be puzzled by the results. While literature on Wal-Mart effects has studied the impact of large discount stores from various aspects of economies, it has never paid attention to the possibility of heterogeneous effects between foreign and domestic ownership. Although the question about why this heterogeneous reactions happen seems to be critical, the topic is beyond this chapter's scope.

In conclusion, Table 4.13 apparently reveals that decrease in markups of domestic stores is primarily attributed to the entry foreign stores, not to that of KLDS, ensuring that our results about treatment effects of FDI on markups are truly reliable. That is to say, average markups fall down as a result of the entry of foreign stores independently of KLDS.

4.6 Conclusion

In this chapter, we examine how multinational enterprises affect domestic firms' markups in the Korean retail sector. Using one of propensity-score based methods, the inverse propensity-score weighting (IPW) regression model, we construct a pseudo-randomised sample. Then, we find that the entry of foreign stores has a negative causal impact on average markups of Korean firms through the difference-

in-difference regression. We also find that the estimated markup effect is not entangled with both productivity spillovers and influences from Korean large discount stores. Therefore, our findings indicate that the decline in markups are only driven by competition effects through the entry of foreign stores.

Table 4.1: Numbers of Stores by Retail Multinationals in Korea

		Number of Stores					
Name	Year of Entry	1997	2000	2003	2006	2009	2012
Carrefour	1996	3	19	28	30 (withdraw)		
Marko	1996	2 (withdraw)					
Wal-Mart	1998		6	13	14 (withdraw)		
Costco	1998	3	4	5	5	7	9
Tesco	1997	1	7	28	51	114	125
Total		9	34	74	111	121	134

Source: Korea Chainstore Association, the 2004 Year Book of the Retail Industry

Table 4.2: Association between Price-Cost Margin and Retail FDI

	Pool OLS	Town FE	Wave FE	Town FE Year FE
	(1)	(2)	(3)	(4)
Retail FDI	-0.056***	-0.087***	-0.014***	-0.005*
	(0.002)	(0.002)	(0.002)	(0.003)
Town FE	No	Yes	No	Yes
Census-Year FE	No	No	Yes	Yes
N	12,498	12,498	12,498	12,498
R^2	0.052	0.082	0.253	0.336
Clusters	3,468	3,468	3,468	3,468

Notes: This table reports results of regressing town-level price-cost margins on retail FDI treatment dummy. Retail FDI treatment indicates whether a town is under the influence of foreign retail stores. More specifically, we define the value of retail FDI dummy as one when the town is subordinated to the county in which foreign store have entered. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01. Standard errors in parentheses are clustered at the town level.

Table 4.3: Probit Estimation Results

	Probit
Markup_{2001}	0.318
	(0.754)
$\Delta\text{Markup}_{2001}$	0.069
	(0.085)
$\text{Labour Productivity}_{2001}$	0.078
	(0.076)
$\Delta\text{Labour Productivity}_{2001}$	-0.044
	(0.053)
Employment_{2001}	0.050
	(0.116)
$\Delta\text{Employment}_{2001}$	-0.213*
	(0.121)
Wage_{2001}	0.069
	(0.148)
ΔWage_{2001}	-0.037
	(0.133)
$\text{Net Entry Rate}_{2001}$	0.070
	(0.071)
$\text{Store Floor Space}_{2001}$	0.101
	(0.100)
Rent Cost_{2001}	0.195***
	(0.063)
Tax Cost_{2001}	-0.097**
	(0.039)
$\text{Korean Large Discount Store}_{2001}$	0.622***
	(0.088)
N	1,634
Pseudo R^2	0.168

Notes: This table shows results of the probit regression of Retail FDI treatment dummy on observable covariates. The estimation results are used to compute propensity scores on which the DDIPW (diff-in-diff inverse propensity-score weighting estimator). Levels variables enter in the log form except Price-Cost Margin, Net Entry Rate and Korean Large Discount Store. * p-value < 0.10, ** p-value < 0.05, *** p-value < 0.01.

Table 4.4: Balancing Test

	Standardised Difference		*Ratio of Variance*	
	Unweighted	Weighted	Unweighted	Weighted
Markup_{2001}	-0.222	0.044	0.537	0.753
$\Delta\text{Markup}_{2001}$	-0.154	0.042	0.942	1.377
$\text{Labour Productivity}_{2001}$	0.429	-0.020	0.896	0.867
$\Delta\text{Labour Productivity}_{2001}$	0.068	-0.030	1.172	0.992
Employment_{2001}	0.872	-0.071	0.389	0.597
$\Delta\text{Employment}_{2001}$	-0.033	-0.054	1.547	1.505
Wage_{2001}	0.439	-0.044	0.683	0.708
ΔWage_{2001}	0.359	-0.047	0.604	0.686
$\text{Net Entry Rate}_{2001}$	0.162	-0.048	0.343	0.846
$\text{Store Floor Space}_{2001}$	0.892	-0.062	0.417	0.641
Rent Cost_{2001}	0.970	-0.092	0.308	0.621
Tax Cost_{2001}	0.591	-0.090	0.547	0.595
$\text{Korean Large Discount Store}_{2001}$	0.783	-0.013	1.144	1.008

Notes: This table reports of the balancing test checking whether the treatment variable (retail FDI entry indicator) becomes randomly assigned in the weighted data.

Table 4.5: Baseline Treatment Definition

	Pre-Treatment $(t-1)$	Treatment (t)	Post-Treatment $(t+1)$
Year	2001	2005	2010
Treatment	Retail FDI$(D_{it}) = 0$	Retail FDI$(D_{it}) = 1$	Retail FDI$(D_{it}) = 1$
Control	retail FDI$(D_{it}) = 0$	retail FDI$(D_{it}) = 0$	retail FDI$(D_{it}) = 0$

Notes: This table aims to illustrate how we define the treatment group (and the control group) using Korea's retail census data for the three periods of time. Specifically, the treatment is defined as towns in which retail FDI is appeared in both 2005 and 2010 census data, but not in 2001 census data. Therefore, treated towns must receive retail FDI during the period of 2002 to 2005.

Table 4.6: Baseline DDIPW Estimation Results

	Unweighted: Diff-in-Diff (1)	*Weighted*: DDIWP (2)
Retail FDI Treatment	0.026***	-0.008**
	(0.006)	(0.004)
Dependent Variable	$\mu_{2010} - \mu_{2001}$	
Treatment	263	263
N	1,634	1,634

Notes: This table shows a simple diff-in-diff regression result relying on the unweighted data in column (1). On the other hand, column (2) presents the estimation result of diff-in-diff inverse propensity score weighting (DDIPW) regression. The results in this table are based on the treatment definition specified in Table 4.5.

Table 4.7: Treatment Definition in Shorter Time Horizon: 2002-2004

	Pre-Treatment $(t-1)$	Treatment (t)	Post-Treatment $(t+1)$
Year	2001	2002-2004	2005
Treatment	retail FDI$(D_{it}) = 0$	retail FDI$(D_{it}) = 1$	retail FDI$(D_{it}) = 1$
Control	retail FDI$(D_{it}) = 0$	retail FDI$(D_{it}) = 0$	retail FDI$(D_{it}) = 0$

Notes: This table illustrates the treatment re-defined by towns receiving retail FDI during the priod of *2002-2004*. The new treatment definition based on KOCA(Korea Chainstore Association) data by which we extract information of *annual* entry status of retail FDI across towns.

Table 4.8: Treatment Definition in Shorter Time Horizon: 2006-2009

	Pre-Treatment $(t-1)$	Treatment (t)	Post-Treatment $(t+1)$
Year	2005	2006-2009	2010
Treatment	retail FDI$(D_{it}) = 0$	retail FDI$(D_{it}) = 1$	retail FDI$(D_{it}) = 1$
Control	retail FDI$(D_{it}) = 0$	retail FDI$(D_{it}) = 0$	retail FDI$(D_{it}) = 0$

Notes: This table illustrates the treatment re-defined by towns receiving retail FDI during the priod of *2006-2009*. The new treatment definition based on KOCA(Korea Chainstore Association) data by which we extract information of *annual* entry status of retail FDI across towns.

Table 4.9: DDIPW Results with Different Periods of Treatments

	(1)	(2)	(3)
Retail FDI Treatment	-0.007***	-0.012***	-0.001
	(0.006)	(0.004)	(0.003)
Dependent Variable	$\mu_{2010} - \mu_{2005}$, $\mu_{2005} - \mu_{2001}$	$\mu_{2005} - \mu_{2001}$	$\mu_{2010} - \mu_{2005}$
Treatment Period	2002-2004, 2006-2009	2002-2004	2006-2009
Data Structure	Panel	Cross-Section	
Treatment	612	177	435
N	3,334	1,554	1,780

Notes: This table displays estimation results of the diff-in-diff inverse propensity score weighting (DDIPW) regression under the treatment definitions noted in Table 4.7 and/or Table 4.8. Column (1) shows the treatment effect of retail FDI on average local markups when we analyse the treatment effect based on a panel data so that we can account for treatments occurring from 2002 to 2004 as well as those happening during the period of 2006-2009. Column (2) is the result when the treatment is defined as towns receiving FDI from 2002 to 2004 (Table 4.7). Meanwhile, column (3) reports the result when defining the treatment as towns receiving retail FDI during the period of 2006 - 2009 (Table 4.8).

Table 4.10: Treatment Effects of Retail FDI on Productivities

	(1)	(2)	(3)	(4)
Retail FDI Treatment	0.108*	0.034	-0.002	0.038
	(0.062)	(0.028)	(0.058)	(0.036)
Dependent Variable	$\text{Prod}_{2010} - \text{Prod}_{2001}$	$\text{Prod}_{2010} - \text{Prod}_{2005}$, $\text{Prod}_{2005} - \text{Prod}_{2001}$	$\text{Prod}_{2005} - \text{Prod}_{2001}$	$\text{Prod}_{2010} - \text{Prod}_{2005}$
Treatment Period	2002-2005	2002-2004, 2006-2009	2002-2004	2006-2009
Data Structure	Cross-Section	Panel	Cross-Section	Cross-Section
Treatment	263	612	177	435
N	1,634	3,334	1,554	1,780

Notes: This table presents the estimation results of the treatment effect of foreign retail entry on domestic mean productivity (i.e., productivity spillover effects), using the diff-in-diff inverse propensity score weighting (DDIWP) regression. Column (1) is the result based on the baseline treatment definition noted in Table 4.5. Column (2) reports the result when we run the regression based on a panel data to investigate whether the spillover effects occur all over the 2000s. Column (3) is the result when defining the treatment as towns receiving retail FDI during the period of 2002-2004 (Table 4.7), whereas column (4) shows the result when the treatment is defined as towns receiving from 2006 to 2009 (Table 4.8).

Table 4.11: Treatment Effects of Retail FDI on Markups: County-level Analyses

	(1)	(2)	(3)	(4)
Retail FDI Treatment	-0.012*	-0.006*	-0.012***	0.0004
	(0.007)	(0.004)	(0.006)	(0.0045)
Dependent Variable	$\mu_{2010} - \mu_{2001}$	$\mu_{2010} - \mu_{2005}$, $\mu_{2005} - \mu_{2001}$	$\mu_{2005} - \mu_{2001}$	$\mu_{2010} - \mu_{2005}$
Treatment Period	2002-2005	2002-2004, 2006-2009	2002-2004	2006-2009
Data Structure	Cross-Section	Panel	Cross-Section	Cross-Section
Treatment	21	45	15	30
N	161	325	155	170

Notes: This table presents the estimation results of the treatment effect of foreign retail entry on domestic mean markups using county-level panel data. We make use of the same empirical methodology, the diff-in-diff inverse propensity score weighting (DDIWP) regression. Column (1) is the result based on the baseline treatment definition noted in Table 4.5. Column (2) reports the result when we run the regression based on a panel data to investigate whether the treatment effects occur all over the 2000s. Column (3) is the result when defining the treatment as towns receiving retail FDI during the period of 2002-2004 (Table 4.7), whereas column (4) shows the result when the treatment is defined as towns receiving from 2006 to 2009 (Table 4.8).

Table 4.12: Town-level Analyses Excluding Direct Treated Towns

	(1)	(2)	(3)	(4)
Retail FDI Treatment	-0.009**	-0.007***	-0.012**	0.0005
	(0.003)	(0.002)	(0.005)	(0.0034)
Dependent Variable	$\mu_{2010} - \mu_{2001}$	$\mu_{2010} - \mu_{2005}$, $\mu_{2005} - \mu_{2001}$	$\mu_{2005} - \mu_{2001}$	$\mu_{2010} - \mu_{2005}$
Treatment Period	2002-2005	2002-2004, 2006-2009	2002-2004	2006-2009
Data Structure	Cross-Section	Panel	Cross-Section	Cross-Section
Treatment	246	570	160	410
N	1,617	3,292	1,537	1,755

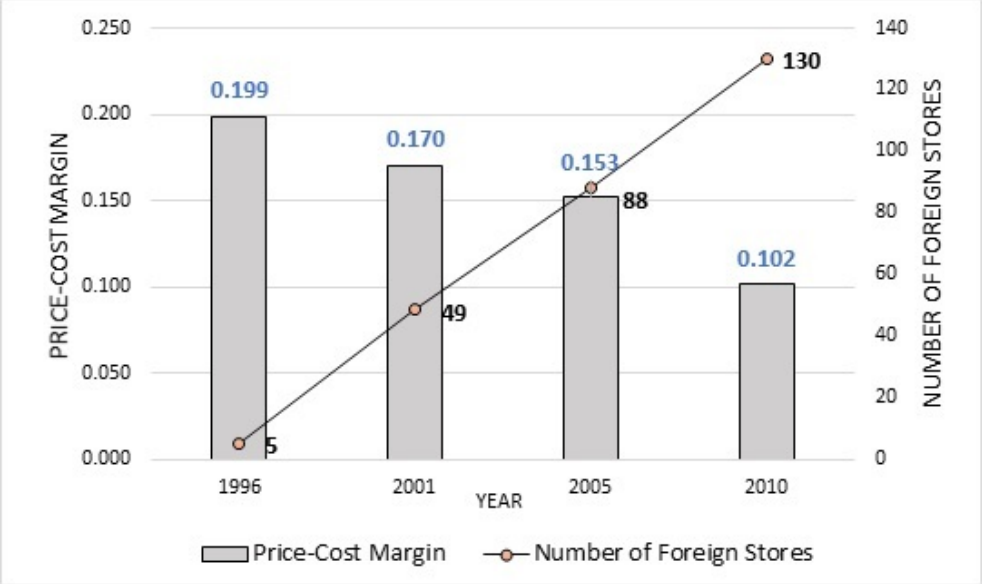
Notes: This table presents the estimation results of the treatment effect of foreign retail entry on domestic mean markups exploiting the diff-in-diff inverse propensity score weighting (DDIWP) regression. Data used in this analyses is town-level panel data in which the direct treated towns is excluded. The direct treated towns imply towns in which retail FDI enter within their own borders. Column (1) is the result based on the baseline treatment definition noted in Table 4.5. Column (2) reports the result when we run the regression based on a panel data to investigate whether the treatment effects occur all over the 2000s. Column (3) is the result when defining the treatment as towns receiving retail FDI during the period of 2002-2004 (Table 4.7), whereas column (4) shows the result when the treatment is defined as towns receiving from 2006 to 2009 (Table 4.8).

Table 4.13: Korean Large Discount Stores(KLDS) and Average Market Markups

	(1)	(2)	(3)	(4)
Korean LDS Treatment	-0.007	-0.003	-0.005	0.002
	(0.007)	(0.004)	(0.006)	(0.009)
Dependent Variable	$\mu_{2010} - \mu_{2001}$	$\mu_{2010} - \mu_{2005}$, $\mu_{2005} - \mu_{2001}$	$\mu_{2005} - \mu_{2001}$	$\mu_{2010} - \mu_{2005}$
Treatment Period	2002-2005	2002-2004, 2006-2009	2002-2004	2006-2009
Data Structure	Cross-Section	Panel	Cross-Section	Cross-Section
Treatment	335	583	285	198
N	1,230	2,362	1,186	1,175

Notes: This table presents the estimation results of the treatment effect of the entry of Korean large discount stores (KLDS) on average market markups exploiting the diff-in-diff inverse propensity score weighting (DDIWP) regression. The dependent variable, average market markups, is measured by town-level average markup across all stores other than *Korean large discount stores*. The treatment is defined as towns in which KLDS enters. Column (1) is the result based on the baseline treatment definition noted in Table 4.5 where KLDS is applied instead of retail FDI. Column (2) reports the result when we run the regression based on a panel data to investigate whether KLDS's treatment effects occur all over the 2000s. Column (3) is the result when defining the treatment as towns with KLDS during the period of 2002-2004 (Table 4.7), whereas column (4) shows the result when the treatment is defined as towns with KLDS from 2006 to 2009 (Table 4.8).

Figure 4.1: Price-Cost Margin and Retail FDI over Years



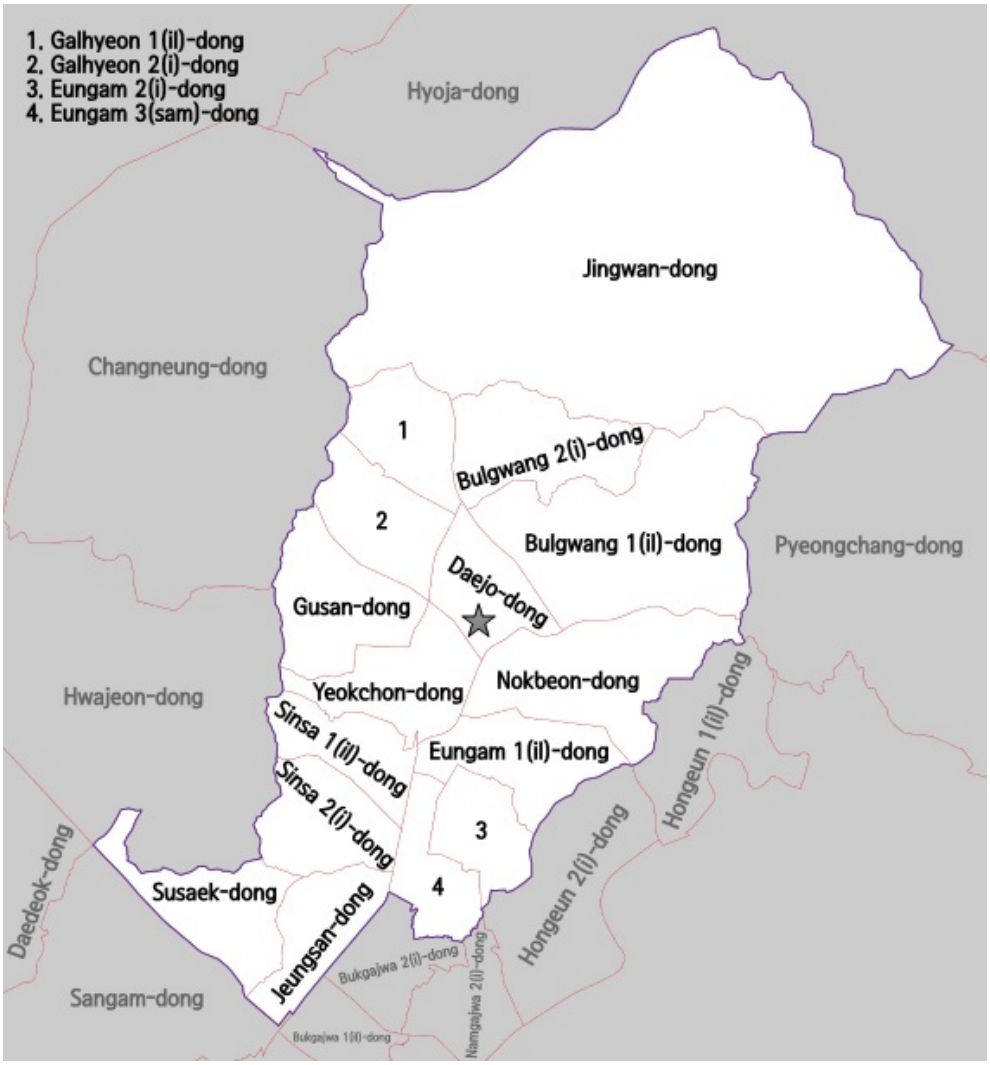
Notes: This figure presents time trends in both (weighted) average price-cost margins of all Korean retail stores and the numbers of foreign stores.

Figure 4.2: Average PCM: County with FDI vs. County without FDI



Notes: This figure displays how the evolutions of (weighted) average price-cost margins differ between counties with retail FDI and counties without receiving retail FDI.

Figure 4.3: An Example of Market Definitions and Treatment



Source:
https://upload.wikimedia.org/wikipedia/commons/2/27/01-12-eunpyeong-en.svg

Note: This figure is a map of a Korean county, Eun-Pyeong Gu. There are a total of sixteen subordinate towns in the county, and a foreign retail (Tesco) store is located on one of towns, Deajo-dong, which is marked by a grey asterisk.

Chapter 5

Conclusion

This thesis has scrutinised critically the significance of retail sector in the global economy, focusing on the context of the Republic of Korea. As globalisation progressed, Korean retail sector has undergone structural changes. In particular, entry of retail multinationals led to fiercer market competitions, and thus retail market structure in Korea became more concentrated towards large-scaled retailers. The structural transformations induced by retail FDI liberalisation, again, have had influences on the outcome of trade reforms in Korea. In Chapter 2, we investigated how the concentrated retail market structure affects consumer price effects of tariff reductions. We found that increase in market concentrations lowered tariff pass-through to consumer prices over the last two decades in Korea. Chapter 3 explored the link between retail market structure and multinationals' entry across Korea's local markets. We showed that, under imperfectly competitive markets, a multinational retailer (Tesco) makes a strategic entry decision in response to its Korean competitor(Emart)'s entry choices. Lastly, Chapter 4 examined whether retail FDI causes local retail firms to charge lower markups. We found causal evidence that retail globalisation leads to such a pro-competitive effect in Korea.

This thesis attempts empirical analyses on the basis of a new data source which has rarely utilised. Throughout the chapters, Korea's retail census is used as one of main data sources. Since many of our analyses are made at the local market level, it is important to account for market-level heterogeneities to correctly answer our research questions. Korean retail census data provided very good data source to construct observable control variables. We also conducted matching consumption-item-level CPI with product-level tariff data with a new matching strategy. In fact, difficulty in matching two data sets has been regarded as one of obstacles to estimate the price effects of international trade. Our matching results expectedly make empirical economists access to research on international trade and consumer price index.

In addition, this thesis is built on various econometric methodologies. Our identification strategies range from panel data estimation model, two stage least square instrument variable estimations, bivariate probit regression model to spatial autoregressive model. Further to reduced-form estimations, we attempted structural model estimations relying on the approach used in empirical IO field, as specified in Chapter 3. To find a causal impact of retail FDI rigorously, we also used microeconometric evaluations techniques. Moreover, our empirical analyses are complemented by theoretical guidelines. In Chapter 2, we proposed a theoretical framework based on an oligopolistic competition model with a nested CES demand. Chapter 3 also presented an empirical model of market entry games between foreign and domestic firms.

While we study the Korean case, it is unlikely that the importance of retail intermediaries is restricted to Korea's economy. Wal-Mart solely accounts for 15% of total imports of consumer goods from China to the U.S. (Basker and Pham Hoang, 2011). The presence of foreign retail stores enhances Chinese exports (Head et al., 2014). Multinational retailers contribute to increase in the total factor productivity (TFP) in the supplying manufacturing industries in Romania (Javorcik and Li, 2013). Empirical evidence has been accumulated by a recent body of research. On top of the literature, we add another empirical evidence on this topic. However, this thesis is differentiated from existing research in the sense that we investigated a unique case of Korea, a country which is unlikely to be categorised as developing country but, at the same time, that has profoundly affected by the tide of retail globalisation since the mid 1990s.

There is still a lot of work to be done in examining the relationship between retail sector and trade. While our attention has been given to consumer prices in this thesis, the role of the retailers could be extended to further issues such as welfare gains of trade and the effect on inequality via the price channel. In addition, although there is a strand of papers on the role of exporting intermediaries initiated by Ahn et al. (2011), the role of retailers as an importing intermediary has not been uncovered yet. This research area would be explored by future research.

As a matter of fact, the findings in Chapter 2 and those in Chapter 4 are interconnected to each other. In particular, we know that retail markups exacerbate the price effect of trade reforms (Chapter 2) and that retail FDI liberalisation decrease average markups of local firms (Chapter 4). Thus, we would infer that retail FDI liberalisation helps to increase tariff pass-through rates into CPI by reducing retail markups. In other words, it would seem that tariff reductions and retail FDI liberalisation have a complementary relationship in terms of consumer price effects. The specific topic is worthy of exploring for another future research.

FDI literature has mainly been involved in manufacturing sectors. On the

other hand, service sectors has been away from the center of FDI literature. Furthermore, the nature and the motivation of FDI in oligopolistic industries has little formal explanation both theoretically and empirically. These gaps in FDI literature would be a valuable extension of this thesis. Lastly, while we investigated the impact of multinational retailers on local firms' markups, it could be extended to other economic indicators such as employments, firm-churning rates, productivities and so on. Through the lens of IPW and DD, it would be possible to find causal impacts of foreign retail stores on those variables. Furthermore, it may also be possible to integrate every facet of economic outcomes in order to make synthetic judgement on what happens to the Korean retail sector after foreign store's entry. The topic is also worth analysing further in the future.

Bibliography

Ackerberg, D., Benkard, C. L., Berry, S. and Pakes, A. (2007), 'Econometric tools for analyzing market outcomes', *Handbook of econometrics* **6**, 4171–4276.

Aguirregabiria, V. (2014), 'Empirical industrial organization: Models, methods, and applications', *Mimeo* .

Aguirregabiria, V. and Vicentini, G. (2016), 'Dynamic spatial competition between multi-store retailers', *The Journal of Industrial Economics* **64**(4), 710–754.

Ahn, J., Khandelwal, A. K. and Wei, S.-J. (2011), 'The role of intermediaries in facilitating trade', *Journal of International Economics* **84**(1), 73–85.

Aitken, B. J. and Harrison, A. E. (1999), 'Do domestic firms benefit from direct foreign investment? evidence from venezuela', *American economic review* **89**(3), 605–618.

Amiti, M., Itskhoki, O. and Konings, J. (2014), 'Importers, exporters, and exchange rate disconnect', *American Economic Review* **104**(7), 1942–78.

Angrist, J. D. and Pischke, J.-S. (2008), *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Arnold, J. M. and Javorcik, B. S. (2009), 'Gifted kids or pushy parents? foreign direct investment and plant productivity in indonesia', *Journal of International Economics* **79**(1), 42–53.

Atkeson, A. and Burstein, A. (2008), 'Pricing-to-market, trade costs, and international relative prices', *American Economic Review* **98**(5), 1998–2031.

Atkin, D. and Donaldson, D. (2015), Who's getting globalized? the size and implications of intra-national trade costs, Technical report, National Bureau of Economic Research.

Basker, E. and Noel, M. (2009), 'The evolving food chain: competitive effects of wal-mart's entry into the supermarket industry', *Journal of Economics & Management Strategy* **18**(4), 977–1009.

Basker, E. and Pham Hoang, V. (2011), Wal-mart as catalyst to us-china trade, Technical report, Development and Policies Research Center (DEPOCEN), Vietnam.

Behrens, K. and Robert-Nicoud, F. (2014), 'Survival of the fittest in cities: Urbanisation and inequality', *The Economic Journal* **124**(581), 1371–1400.

Bernard, A. B., Jensen, J. B., Redding, S. J. and Schott, P. K. (2010), 'Wholesalers and retailers in us trade', *The American Economic Review* pp. 408–413.

Berry, S. T. (1992), 'Estimation of a model of entry in the airline industry', *Econometrica: Journal of the Econometric Society* pp. 889–917.

Blomström, M. and Kokko, A. (1998), 'Multinational corporations and spillovers', *Journal of Economic surveys* **12**(3), 247–277.

Bresnahan, T. F. (1989), 'Empirical studies of industries with market power', *Handbook of industrial organization* **2**, 1011–1057.

Bresnahan, T. F. and Reiss, P. C. (1990), 'Entry in monopoly market', *The Review of Economic Studies* **57**(4), 531–553.

Broda, C. and Romalis, J. (2008), Inequality and prices: Does china benefit the poor in america, Technical report, University of Chicago (mimeo).

Burstein, A. and Gopinath, G. (2014), International prices and exchange rates, *in* 'Handbook of international economics', Vol. 4, Elsevier, pp. 391–451.

Burstein, A. T., Neves, J. C. and Rebelo, S. (2003), 'Distribution costs and real exchange rate dynamics during exchange-rate-based stabilizations', *Journal of monetary Economics* **50**(6), 1189–1214.

Busso, M., DiNardo, J. and McCrary, J. (2014), 'New evidence on the finite sample properties of propensity score reweighting and matching estimators', *Review of Economics and Statistics* **96**(5), 885–897.

Cheung, J. and Hongsik, L. (2010), 'Changes in korean tariff structure under the new free trade agreements', *Korean Institute of Public Finance* .

Cho, J., Chun, H. and Lee, Y. (2015), 'How does the entry of large discount stores increase retail employment? evidence from korea', *Journal of Comparative Economics* **43**(3), 559–574.

Choi, J., Chung, D., Park, H. and Song, J. (2012), Trade policy of promotion of food industry in korea, Technical report, Korean Rural Economic Institite.

Ciliberto, F. and Tamer, E. (2009), 'Market structure and multiple equilibria in airline markets', *Econometrica* **77**(6), 1791–1828.

Coe, N. M. and Lee, Y.-S. (2013), "we've learnt how to be local': the deepening territorial embeddedness of samsung–tesco in south korea', *Journal of Economic Geography* **13**(2), 327–356.

Cohen, A. M. and Mazzeo, M. J. (2007), 'Market structure and competition among retail depository institutions', *The Review of Economics and Statistics* **89**(1), 60–74.

Corsetti, G. and Dedola, L. (2005), 'A macroeconomic model of international price discrimination', *Journal of International Economics* **67**(1), 129–155.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K. and Pavcnik, N. (2016), 'Prices, markups, and trade reform', *Econometrica* **84**(2), 445–510.

De Loecker, J. and Warzynski, F. (2012), 'Markups and firm-level export status', *American economic review* **102**(6), 2437–71.

Dornbusch, R. (1987), 'Exchange rates and prices', *American Economic Review* **77**(1), 93–106.

Ekholm, K., Forslid, R. and Markusen, J. R. (2007), 'Export-platform foreign direct investment', *Journal of the European Economic Association* **5**(4), 776–795.

Ellickson, P. B., Houghton, S. and Timmins, C. (2013), 'Estimating network economies in retail chains: a revealed preference approach', *The RAND Journal of Economics* **44**(2), 169–193.

Ellickson, P. B. and Misra, S. (2011), 'Structural workshop paper—estimating discrete games', *Marketing Science* **30**(6), 997–1010.

Faber, B. (2014), Trade liberalization, the price of quality, and inequality: Evidence from mexican store prices, Technical report, Mimeo, University of California at Berkeley.

Feenstra, R. C. (1989), 'Symmetric pass-through of tariffs and exchange rates under imperfect competition: An empirical test', *Journal of international Economics* **27**(1), 25–45.

Feenstra, R. C. and Weinstein, D. (2016), 'Globalization, competition, and us welfare', *Journal of Political Economy* .

Gaudin, G. (2016), 'Pass-through, vertical contracts, and bargains', *Economics Letters* **139**, 1–4.

Girma, S. (2005), 'Safeguarding jobs? acquisition fdi and employment dynamics in uk manufacturing', *Review of World Economics* **141**(1), 165–178.

Girma, S., Gong, Y., Görg, H. and Lancheros, S. (2015), 'Estimating direct and indirect effects of foreign direct investment on firm productivity in the presence of interactions between firms', *Journal of International Economics* **95**(1), 157–169.

Girma, S. and Görg, H. (2007), 'Evaluating the foreign ownership wage premium using a difference-in-differences matching approach', *Journal of International Economics* **72**(1), 97–112.

Goldberg, L. G. and Grosse, R. (1994), 'Location choice of foreign banks in the united states', *Journal of Economics and Business* **46**(5), 367–379.

Goldberg, L. S. and Campa, J. M. (2010), 'The sensitivity of the cpi to exchange rates: Distribution margins, imported inputs, and trade exposure', *The Review of Economics and Statistics* **92**(2), 392–407.

Görg, H., Warzynski, F. et al. (2003), Price cost margins and exporting behaviour: Evidence from firm level data, Technical report, Research paper, Leverhulme Centre for Research on Globalisation and Economic Policy.

Greene, W. H. (2007), 'Econometric analysis'.

Han, J., Liu, R., Marchand, B. U. and Zhang, J. (2016), 'Market structure, imperfect tariff pass-through, and household welfare in urban china', *Journal of International Economics* **100**, 220–232.

Han, S. and Vytlacil, E. (2016), Identification in a generalization of bivariate probit models with dummy endogenous regressors, Technical report, working Paper, University of Texas at Austin and New York University. 2.3, 6, B.

Harrison, A. E. (1994), 'Productivity, imperfect competition and trade reform: Theory and evidence', *Journal of international Economics* **36**(1-2), 53–73.

Haskel, J. E., Pereira, S. C. and Slaughter, M. J. (2007), 'Does inward foreign direct investment boost the productivity of domestic firms?', *The review of economics and statistics* **89**(3), 482–496.

Head, K., Jing, R. and Swenson, D. L. (2014), 'From beijing to bentonville: Do multinational retailers link markets?', *Journal of Development Economics* **110**, 79–92.

Head, K., Mayer, T. and Ries, J. (2002), 'Revisiting oligopolistic reaction: are decisions on foreign direct investment strategic complements?', *Journal of Economics & Management Strategy* **11**(3), 453–472.

Heckman, J. J. (1977), 'Dummy endogenous variables in a simultaneous equation system'.

Heckman, J. J., Ichimura, H. and Todd, P. E. (1997), 'Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme', *The review of economic studies* **64**(4), 605–654.

Helpman, E. (1984), 'A simple theory of international trade with multinational corporations', *Journal of political economy* **92**(3), 451–471.

Hoenen, A. K. and Hansen, M. W. (2009), Oligopolistic competition and foreign direct investment, Technical report, Working Paper.

Holmes, T. J. (2011), 'The diffusion of wal-mart and economies of density', *Econometrica* **79**(1), 253–302.

Hong, G. H. and Li, N. (2017), 'Market structure and cost pass-through in retail', *Review of Economics and Statistics* **99**(1), 151–166.

Horvitz, D. G. and Thompson, D. J. (1952), 'A generalization of sampling without replacement from a finite universe', *Journal of the American statistical Association* **47**(260), 663–685.

Hotelling, H. (1929), 'Stability in competition', *The economic journal* **39**(153), 41–57.

Hotz, V. J. and Miller, R. A. (1993), 'Conditional choice probabilities and the estimation of dynamic models', *The Review of Economic Studies* **60**(3), 497–529.

Iacovone, L., Javorcik, B., Keller, W. and Tybout, J. (2015), 'Supplier responses to walmart's invasion in mexico', *Journal of International Economics* **95**(1), 1–15.

Imai, K. and Ratkovic, M. (2014), 'Covariate balancing propensity score', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **76**(1), 243–263.

Imbens, G. W. (2004), 'Nonparametric estimation of average treatment effects under exogeneity: A review', *Review of Economics and statistics* **86**(1), 4–29.

Javorcik, B. S. and Li, Y. (2013), 'Do the biggest aisles serve a brighter future? global retail chains and their implications for romania', *Journal of International Economics* **90**(2), 348–363.

Jeon, S. and Whang, Y.-J. (2010), 'An econometric analysis of competitive effects of eland-carrefour merger in 2006', *Journal of Regulation Studies* **19**(2), 75–103.

Jia, P. (2008), 'What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry', *Econometrica* **76**(6), 1263–1316.

Jones, J. and Wren, C. (2016), 'Does service fdi locate differently to manufacturing fdi? a regional analysis for great britain', *Regional Studies* **50**(12), 1980–1994.

Keller, W. and Yeaple, S. R. (2009), 'Multinational enterprises, international trade, and productivity growth: firm-level evidence from the united states', *The Review of Economics and Statistics* **91**(4), 821–831.

KFTC (2009), Sales commision of large retail stores and policy implications, Technical report, Korea Fair Trade Commission.

Kim, J.-I. and Kim, J.-D. (2003), Liberalization of trade in services and productivity growth in korea, *in* 'Trade in Services in the Asia Pacific Region, NBER East Asia Seminar on Economics (EASE), Volume 11', University of Chicago Press, pp. 179–208.

Kim, S.-K., Kim, C.-K. and Kim, K.-H. (2002), 'Structural change analysis and policy suggestions for strengthening competitiveness of korean retail industry', *Korean Institute for Industrial Economics and Trade (KIET)* .

Knickerbocker, F. T. (1973), 'Oligopolistic reaction and multinational enterprise', *Thunderbird International Business Review* **15**(2), 7–9.

Lee, Y., Chun, H. and Hur, S. (2014), 'Entry and competition of large discount stores in korea', *The Korean Journal of Economic Studies* **62**(3), 93–118.

Levinsohn, J. (1993), 'Testing the imports-as-market-discipline hypothesis', *Journal of International Economics* **35**(1-2), 1–22.

Maddala, G. S. (1986), *Limited-dependent and qualitative variables in econometrics*, number 3, Cambridge university press.

Marchand, B. U. (2012), 'Tariff pass-through and the distributional effects of trade liberalization', *Journal of Development Economics* **99**(2), 265–281.

Markusen, J. R. (1984), 'Multinationals, multi-plant economies, and the gains from trade', *Journal of international economics* **16**(3-4), 205–226.

Melitz, M. J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica* **71**(6), 1695–1725.

Melitz, M. J. and Ottaviano, G. I. (2008), 'Market size, trade, and productivity', *The review of economic studies* **75**(1), 295–316.

Morrissey, O. and te Velde, D. W. (2003), 'Trade, fdi and equity in the republic of korea in the 1990s'.

Muraközy, B. and Russ, K. N. (2015), Competition with multinational firms: Theory and evidence, Technical report, IEHAS Discussion Papers.

Nakamura, E. (2008), 'Pass-through in retail and wholesale', *American Economic Review* **98**(2), 430–37.

Nakamura, E. and Zerom, D. (2010), 'Accounting for incomplete pass-through', *The Review of Economic Studies* **77**(3), 1192–1230.

Newey, W. K. (1987), 'Efficient estimation of limited dependent variable models with endogenous explanatory variables', *Journal of Econometrics* **36**(3), 231–250.

Neyman, J. and Scott, E. L. (1948), 'Consistent estimates based on partially consistent observations', *Econometrica: Journal of the Econometric Society* pp. 1–32.

Nicita, A. (2009), 'The price effect of tariff liberalization: Measuring the impact on household welfare', *Journal of Development Economics* **89**(1), 19–27.

Oster, E. (2017), 'Unobservable selection and coefficient stability: Theory and evidence', *Journal of Business & Economic Statistics* pp. 1–18.

Porto, G. G. (2006), 'Using survey data to assess the distributional effects of trade policy', *Journal of International Economics* **70**(1), 140–160.

Ramondo, N. and Rodríguez-Clare, A. (2013), 'Trade, multinational production, and the gains from openness', *Journal of Political Economy* **121**(2), 273–322.

Robins, J. M., Rotnitzky, A. and Zhao, L. P. (1994), 'Estimation of regression coefficients when some regressors are not always observed', *Journal of the American statistical Association* **89**(427), 846–866.

Rosenbaum, P. R. and Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Rust, J. (1987), 'Optimal replacement of gmc bus engines: An empirical model of harold zurcher', *Econometrica: Journal of the Econometric Society* pp. 999–1033.

Sembenelli, A. and Siotis, G. (2008), 'Foreign direct investment and mark-up dynamics: Evidence from spanish firms', *Journal of International Economics* **76**(1), 107–115.

Siotis, G. (2003), 'Competitive pressure and economic integration: An illustration for spain, 1983–1996', *International Journal of Industrial Organization* **21**(10), 1435–1459.

Spengler, J. J. (1950), 'Vertical integration and antitrust policy', *Journal of political economy* **58**(4), 347–352.

Stock, J. H. and Yogo, M. (2005), 'Testing for weak instruments in linear iv regression'.

Suh, Y.-G. and Hahn, G.-D. (2004), 'Analysis on saturation factors of korean discount stores', *Journal of Distribution Research* **9**(4), 65–83.

Tamer, E. (2003), 'Incomplete simultaneous discrete response model with multiple equilibria', *The Review of Economic Studies* **70**(1), 147–165.

Tintelnot, F. (2017), 'Global production with export platforms', *The Quarterly Journal of Economics* **132**(1), 157–209.

Topalova, P. (2010), 'Factor immobility and regional impacts of trade liberalization: Evidence on poverty from india', *American Economic Journal: Applied Economics* **2**(4), 1–41.

UNCTAD (2015), *World investment report 2015: Reforming international investment governance*, UN.