Essays on the Discrete Choice Model: Application and Extension

Inauguraldissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

Aiyong Zhu

vorgelegt im Sommersemester 2014

Abteilungssprecher: Prof. Dr. Eckhard Janeba

Referent: Prof. Philipp Schmidt-Dengler, PhD

Koreferent: Prof. Kathleen Nosal, PhD

Verteidigung: 20. Mai 2014

Acknowledgments

During my four years' PhD study at University of Mannheim, I benefited a lot from so many people in the economic department. I would like to take this opportunity to express my sincere appreciation and gratitude to all of them.

First and foremost, I am deeply indebted to my supervisor: Philipp Schmidt-Dengler for his continuous support and encouragement of my Ph.D study and research, and for his patience and guidance to help me in all the time of writing this thesis. Moreover, I owe special thanks to my co-advisors: Kathleen Nosal and Yuya Takahashi for their invaluable feedback on my work and countless discussions, in particular on the essence of dynamic models.

I am also grateful to insightful comments on my thesis from Isis Durrmeyer, Qiang Guo, Christopher Koch, Volker Nocke, Alex Shcherbakov, Konrad Stahl, Naoki Wakamori, and Stefan Weiergraeber. My sincere thanks also goes to the fellow students: Agustin H. Arias, Johannes Beutel, Jieping Chen, Robert Heimbach, Hannes Kammerer, Christian Koch, Michael Kummer, Vera Molitor, Cornelius Müller, Alexander Paul, Elena Rancoita, Johannes Schoch, Michele Weynandt, for their numerous help to smooth my life in Germany. In addition, I gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft and I also highly appreciate Deutsche Bundesbank for their hospitality as well as the access to the database.

Last but not the least, I would like to thank my parents for their constant support throughout my life.

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

My thesis consists of three chapters that address very diverse research questions involving foreign direct investment and audit market competition. The common underlying methodology-discrete choice model is the key factor that connects all chapters covering the different research fields. In each chapter, I applies a variation or extension of the class to discrete choice models to pursue the corresponding research topics.

The origins of discrete choice model are rooted in the early studies of psychophysics (the physical study of the relations between physical stimuli and sensory response) at 1860. A pathbreaking approach for modeling individual decisions using discrete choice models emerged in the 1970's. It was pioneered by the work of McFadden on Random Utility Maximization theory (McFadden, 1974, 1981). Suppose that an individual has to choose among N mutually exclusive alternatives of a product or of a service. Given the individual's preference, he derives a utility consisting of both deterministic and random components from each alternative. The deterministic part depends on the on the characteristics of the individual and the attributes of the alternative; while the random part of the utility is generally assumed to follow certain distribution to capture the decision maker's idiosyncratic preference. The individual then chooses an alternative to maximize his utility. With the idiosyncratic preference, any choice observed in the data can be rationalized by Random Utility Maximization theory. Maximum likelihood and simulated maximum likelihood based on individual choice data have been developed to estimate the parameters of these choice models from data on revealed or stated preferences, using a wide range of structural specifications.

Berry (1994) and Berry, Levinsohn and Pakes (henceforth BLP, 1995, 2004) have vastly extend the discrete choice model in many dimensions to capture a rich set of individual heterogeneity and estimate the heterogeneous preference only using aggregated market level data. As BLP approach has become a workhorse in empirical industrial organization, I applied it in another field of international trade to study how firms choose location to engage in FDI (foreign direct investment) in Chapter 2. In Chapter 4, we study how public listed companies choose audit firms by using an extension of BLP approach. Different from standard demand estimation for differentiated products, one important feature of the audit market is that product characteristics (audit service) are not constant across all listed firms that are obligated to buy them. Nested with BLP approach, Gowrisankaran and Rysman (2012) first develop a dynamic discrete choice model to study how consumers choose durable goods. Building on this state-of-the-art approach, I extend the dynamic discrete choice model to quantify adaptation costs in firms' sequential FDI location choice pattern in Chapter 3.

In Chapter 2, I apply the BLP approach to investigate the pattern and determinants of horizontal foreign direct investment. To fit in the framework of discrete choice models, locations are characterized by a bundle of both observed and unobserved attributes. Firms are endowed with heterogeneous preference for location attributes, and choose at most one location to engage in FDI each period to maximize their (latent) profit. Based on the panel dataset on German firms' outward FDI behavior, I find a persistence pattern in firm's FDI location choice that firms are more likely to keep investing in the same location where they have invested before. This persistence pattern is essentially the state dependence that has been widely observed and studied in many other economic fields, e.g., labor economics, marketing, etc. As Heckman (1981) has pointed out the decision maker's heterogeneity should be taken into account in order to identify the true state dependence. In general, the observed persistence pattern in firm's FDI location choice may be attributed to three channels: first, true state dependence in firms' decision making; second, location attractiveness which can be measured by market size, economic growth, labor cost, etc.; third firms' heterogeneous preference for some specific locations. To disentangle from other two channels, I model a great variety of firm heterogeneity and control for location attractiveness at the same time. The estimation results from all specifications confirm that firms' previous investment experience does have a significantly positive effect on the FDI location choice. In addition to identifying the state dependence, I find that the market size and similarity between economic endowment also have a positive effect in attracting horizontal FDI.

Given the findings in Chapter 2, I seek to explain why we observe such a special pattern in firms' FDI location choice in Chapter 3. I interpret this state dependence as an evidence of the presence of adaptation costs for firm to engage in FDI. When engaging in FDI for the first time, multinational firms often incur large costs to adapt to different business practices, ethical norms and regulations in a new country. Therefore, prior investment experience in a location can influence firms' FDI decisions in subsequent periods, since firms do not need to pay the adaptation costs again if they invest in the same location. I define these one-shot fixed entry costs as adaptation costs and then impose a dynamic structural model to quantify the magnitude of these adaptation costs. To my best knowledge, this is the first paper to address the issue about the magnitude of entry costs in FDI. Similarly in the related literature of exporting, Roberts and Tybout (1997) and Das Roberts and Tybout (2007) first identify the presence of entry costs in dynamic exporting and then structurally estimate the sunk entry costs using the dataset on exporting behavior of Colombian manufacturing firms.

To incorporate the state dependence in multinational firms' FDI location decisions, I propose a dynamic discrete choice model, in which firms sequentially choose a location for FDI to maximize the expected discounted profit with adaptation costs and firm heterogeneity taken into account. The gross one-period profit flow earned by firms is approximated as a function of both location and firm characteristics. Adaptation costs are only paid these firms that enter the location to engage in FDI for the first time. The empirical results show that adaptation costs are substantial and varying across locations, ranging from $0.5\% - 29\%$ of the average expected discounted profit[1]. With these estimation results, I then simulate several policy experiments involving FDI promotion schemes that are associated with reduction in adaptation costs. These counterfactuals suggest that firms with distinct experience states respond systematically differently to different FDI promotion policies. In general, the expected discounted profit across all firms would increase by around 15%, on average over time if adaptation costs were eliminated in every location completely. More importantly, this increment in profit is largely due to better matching between locations and firms. Over the sample period, around 21% would change another location to engage in FDI if adaptation costs in all locations were subsidized to be zero. The information that adaptation costs have a significant influence on firms' FDI location choices is particularly important for both host and home countries. For home countries, the results provide some insight into the benefit of the policy on subsidizing firms when entering foreign markets. For host countries, the right promotion policy will effectively attract FDI which may boost local economic growth.

Chapter 4 is a coauthored paper about demand estimation in the audit market with Qiang Guo and Christopher Koch. In this paper, we extend the discrete choice model to study the potential impact of several policies proposed by the European Commission on the audit market. In contrast to the standard discrete choice model, one special feature in modeling how audit firm choice lies in that some product characteristics, e.g., audit fees charged by the same audit firm vary across clients, i.e., public listed companies in the stock market. Moreover, we only observe audit fees for actual matches between audit firm and clients. This is another unusual feature in audit firm choice model that we need to predict what audit fees a client would have expected to pay, if it had chosen another audit firm than the one we observed in the data. When predicting audit fees, we also need to deal with the typical endogenous price problem in demand estimation. In this case, the audit fees are usually correlated with audit quality or other unmeasured characteristics of audit firms left in the error term. As suggested by Gerakos and Syverson (2013), the exogenous supply shock from merger and acquisitions between clients can be used as an

[1] "average" is referring to the mean value across all observed German non-manufacturing firms over the entire sample period.

instrument for predicting audit fees. We first test the validity of this instrument in our model and then apply the control function approach to address endogenous predicted audit fees for the nonlinear discrete choice model. In the second stage, we conduct the standard demand estimation by using the predicted audit fees and conditioning on the control function. After recovering clients' preference towards audit firms, we are able to study how the proposed joint audit policy[2] would affect the audit market structure in the UK.

Audit market in the UK is the largest and most concentrated in Europe; whereas French audit market has the lowest concentration ratio and France is the only country that implements joint audit policy in Europe. We are particulary interested in whether the joint audit policy would affect market concentration in the UK. Intuitively, since the high dominance of big four audit firms[3] in the UK reveals that listed companies prefer big 4 audit firms because of their high reputation or outstanding service, we should not expect that these companies would choose another small audit firms if they were obligated to hire two audit firms. Then how do we explain the lowest market share of big four audit firms in France? With a deeper dig into the French data, we find that the most prevalent combinations of the two audit firms consists of one from big four audit firm and the other one from a small audit firm. Accordingly, we propose there may exist different synergy (pair effect) in different combinations of audit firms and clients also have heterogeneous preference towards these combinations. Thus, we first use French audit market to identify the synergy (pair) effect between different pair types in joint audit and in the meantime estimate clients' preference for the individual audit firms. Then we conduct the counterfactual analysis of implementing joint audit policy in the UK. The simulation results show that joint audit policy would reshape the market structure substantially and reduce the market concentration of big four audit firms to a mild extent. The counterfactuals also indicate that joint audit would increase the welfare significantly due to the positive pair effect, but this pair benefit is not high enough compensate the

[2]Detailed definitions are present in Chapter 4. In brief, joint audit requires that the client has to hire two audit firm issuing one independent auditing report annually.

[3]Big four audit firms are referring to PricewaterhouseCoopers, Deloitte, KPMG and Ernst & Young.

welfare loss from being forced to choose another auditor. As a consequence, the net welfare change for clients in the UK would be negative and decrease by 7.2 million GBP on average over the sample period.

Chapter 2

Location Patterns and Determinants of FDI: A View from Firm-level Data

2.1 Introduction

There has been a long standing debate about what factors most influence the location decision of foreign direct investment (henceforth FDI) in the global market. Most countries has set up a top policy agenda to attract FDI inflow to boost local economic growth, even though how multinational firms choose locations[1] to invest is not well understood yet. Theoretical studies often break down FDI into horizontal FDI and vertical FDI due to different incentives. Horizontal FDI is defined as an activity that multinational firms produce the same goods and services in multiple countries, aiming for serving local markets. Vertical FDI is defined as that firms locate different stages of production in different countries, taking advantage of international factor-price differences. An extensive empirical study in most of the literature (Carr, Markusen, and Maskus (2001), etc) has reached an conclusion that large market size and high similarity in economic endowment between countries are most attractive to horizontal FDI; while low labor cost is an important determinant in the vertical FDI location choice.

While most of the empirical paper use firm level data to study the determinants of different type of FDI, they often underemphasize or can not observe due to data availabil-

[1] Location and country are used interchangeably throughout this paper.

ity the repeated investment behavior by the same firm over time. A rich panel dataset -Microdatabase Direct Investment (henceforth MIDI) on German firms' outward FDI enables me to bridge the current gap to study the location patterns and determinants of FDI in a more comprehensive way. By tracking each individual multinational firm's investment behavior over time, I find an interesting pattern that firms' FDI location choice persists over time. That is to say, firms are more likely to keep investing in the same country where they invested before. This finding resembles the pattern of multinational firms' global expansion dicussed in the earlier management literature (Johanson and Vahlne (1977), etc). They claim that firms in the initial stage of foreign expansion exhibit a strong preference for markets which are culturally and economically similar to the home country. In this paper, I do not plan to explain why we can observe such a persistent pattern in FDI location choice from the data. Instead, I focus on the empirical question: whether firms' investment history really matters for the FDI location choice. It turns out that this persistent pattern is very robust even after controlling for various location characteristics as well as both observed and unobserved firm heterogeneity. The control for firm heterogeneity is very necessary because it helps to rule out the "spurious state dependence" pointed out by Heckman (1981).

My paper complements a growing literature about how heterogeneous firms participate in international markets. Until recently, the issue of firm heterogeneity in the FDI location decision process starts to draw attention of researchers. The pioneering work of Helpman, Meltiz and Yeaple (2004) presents a theoretical model about how heterogeneous firms in productivity endogenously choose the way of foreign market access: exporting, FDI or only serving domestic market. Tomiura (2007) support their conclusion with empirical evidence from Japanese manufacturing firms that exporting firms tend to be less productive than firms active in FDI, but more productive than domestic firms. Aw and Lee (2008), using a firm-level data in Taiwan also confirm that more productive firms engage in FDI. In addition, they show that the most productive firms tend to invest in multiple countries and firms choosing the locations with high investment cost are relatively more productive than those choosing locations with low investment cost. Yeaple (2009) demonstrate that this sorting in productivity can be extended to the scale and scope of multinational firms. Based on a firm level data of US multinational activity in manufacturing industries, he finds that more productive firms are able to access a larger set of foreign markets and set up affiliates there with larger size than those less productive firms. Chen and Moore (2010) explicitly model and examine the relationship between the market attractiveness and firm heterogeneity measured by total factor productivity (TFP). Using a rich dataset on French manufacturing firms' international activity, they find that more productive firms are self-selected to invest in relatively tough foreign markets, e.g., smaller market potential and higher investment costs.

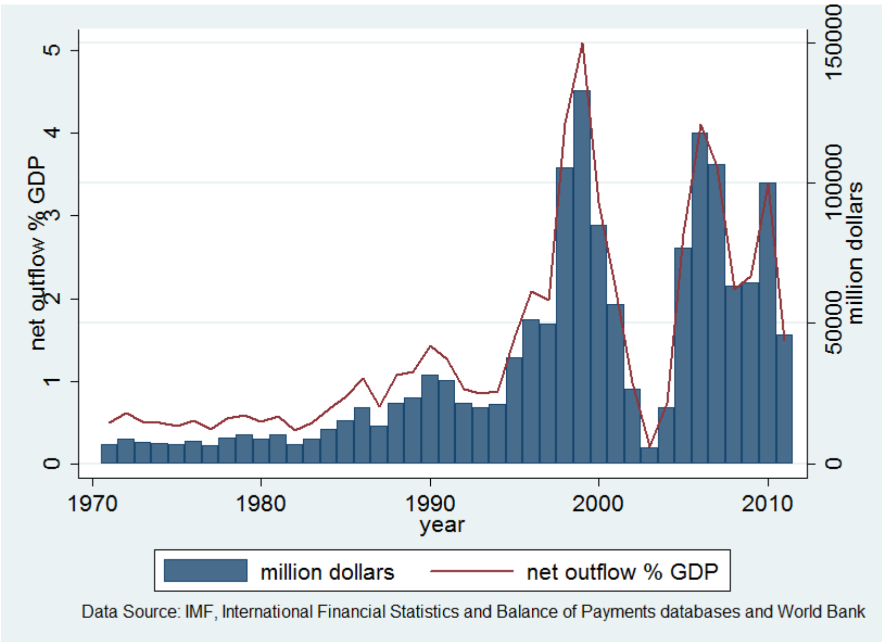
The main contribution of this paper to the FDI literature is twofold. First and foremost, by using detailed plant-level panel data, I clearly identify the state dependence in multinational firms' FDI location choice. In the previous papers by Barkema, Bell, and Pennings (1996), Shaver, Mitchell, and Yeung (1997), and Delios and Henisz (2003), they also provide empirical evidence that firms' country-specific investing experience affects the FDI location decision because it can help firms reduce entry barriers like cultural distance and political hazard for the subsequent FDI in the same or related countries. However, they ignore unobserved firm heterogeneity that could also generate the same persistence pattern as observed in the data. In this paper, I explicitly account for observed and unobserved firm heterogeneity and show that the observed state dependence in the data is indeed related to firms' investment history. Second, contrary to most empirical papers, I focus on multinational firms in non-manufacturing industry, whose firm observations take the majority, around 60% of all German multinational firms. Due to the nature of non-manufacturing industry, most of the firms there produce non-tradable goods or provide local service. Thus the FDI engaged by non-manufacturing firms are more likely to be horizontal FDI because they can only serve the local markets. This focus is particularly important for policy consideration since the high prevalence of horizontal FDI are most engaged by non-manufacturing firms.

The rest of the paper is organized as follows. I first provide a detailed description the data in section 2 and then present empirical models to identify the state dependence in section 3. Estimation results and analysis are discussed in section 4. In the end, I conclude in section 5.

2.2 Data Description

The standard definition of FDI is the net inflows of the direct investment to expand production or business operating in a foreign country. It is the sum of equity capital, reinvestment of earnings, other long-term capital, and short-term capital as shown in the balance of payments. Figure (2.1) shows the long run FDI trend in Germany. It can be seen that the aggregated volume of outward FDI grows over time and reaches the peak at 1999, around 5% of GDP.[2]

Figure 2.1: Aggregated volume of outward FDI in Germany during 1971-2011



The large amount of aggregated volume of FDI on one hand shows that German multinational firms are the major players actively participating international markets, but on the other hand, it does not provide any information about the FDI location distribution. Since I am focusing on the location patterns of FDI, I will describe in length the MIDI panel database provided by the Bundesbank (German Central Bank).

[2]The sharp drop at 2003 may be due to the change of reporting threshold.

The Bundesbank has been collecting annual statistics on foreign direct investment stocks in accordance with the provision of the Foreign Trade and Payments Regulation since 1976. German enterprises need to report their international capital links if the direct investment enterprise abroad meets reporting requirements involving both its total assets and shareholding of the associated German parent firm.

Table 2.1 shows that reporting thresholds have been altered many times during the last several decades due to changes in Accounting and Reporting Law. The MIDI forms a unbalanced panel dataset[3] on the subsidiary level from 1996 onwards, and since 2002, the information about the German parent firms' characteristics including total assets, turnover, number of employees, etc of reporting enterprises become available.

Table 2.1: Reporting thresholds

Reporting year	Shares of parent firms	Total assets(subsidiary)
1993	$\geq 20\%$	> DM 1 million
1999	10%-50%	> DM 10 million
	$\geq 50\%$	> DM 1 million
2002	$\geq 10\%$	> Euro 3 million

2.2.1 Location distribution of German outward FDI

Since different types of FDI arise from different incentives, this paper will concentrate on the more prevalent horizontal FDI. Compared with vertical FDI, the profit from horizontal FDI can be less controversially assumed to be a function of the local location attributes as well as firm characteristics. It is common knowledge that firms in manufacturing sectors can engage in both kinds of FDI, while investment by firms in non-manufacturing industries is often regarded as horizontal FDI because their non-tradable products can only be consumed in local market. Since the FDI engaged by non-manufacturing firms has not been paid much attention in most of the empirical papers, I in particular focus on firms in non-manufacturing industry. In contrast to the standard definition, FDI in this paper is defined as a new affiliate over the reporting threshold set up by German parent

[3] The unbalanced panel is mainly driven by the new firms entering into the sample every year. Meanwhile, the change in the reporting threshold is another explanation.

firms. In particular, the shareholding by the parent firm must be larger than 50%, i.e., an absolute majority shareholder for the investment to be considered as an FDI.[4] In the remainder of this paper, “German firms" refers to German non-manufacturing firms for short and both are used interchangeably.

According to the MIDI database covering from 1996 to 2009, the top 10 preferred locations chosen by German non-manufacturing firms for FDI are indicated in Figure (2.2). More precisely, Figure (2.2) shows the percentage of the number of stocks of German foreign affiliates among the top 10 countries. Not surprisingly, nine out of 10 are European countries and the most preferred country is the USA with the largest market size measured by GDP in the world. This observation is consistent with conventional theory that horizontal FDI is attracted by large market size and high similarity between economic endowments (Carr, et al., 2001). In this case, German non-manufacturing firms prefer to setting up subsidiaries in countries with large GDP and high GDP per capita.

Figure 2.2: Top 10 FDI locations during 1996-2009

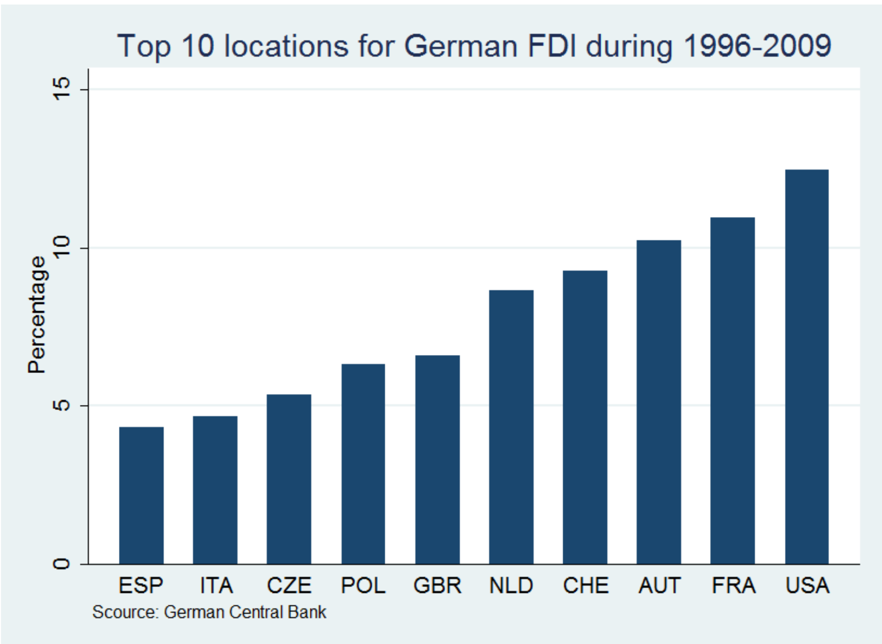


Table (2.2) shows more interesting patterns of the FDI location choice. The first col-

[4]Due to data availability, I can not distinguish different FDI entry mode, e.g. greenfield and merge and acquisition (M&A). For greenfield subsidiaries, parent firms are almost absolute majority shareholders, thus 50% restriction is mainly set for subsidiaries acquired by M&A.

umn of Table 2.2 shows the overall top 10 locations among all German non-manufacturing firms, which coincides with Figure (2.2). The second column presents the top 10 locations for firms that only invested once during 1996 to 2009. The three columns on the right are related to firms that invested more than once and t_0 denotes the first year in the sample in which firms first engaged in FDI. A large number of firms, around 60%, only invested once during the whole period as in the second column and among these firms, the USA is their most preferred investment location choice. On the contrary the remaining 40% firms who have invested multiple times choose France and Austria as their top two locations. However, if I focus on the timing ($t = t_0$, $t > t_0$) of FDI decisions, I find that their first-time ($t = t_0$) FDI top two location choices are the same as the overall pattern for firms investing multiple times, but in the subsequent periods ($t > t_0$) France and Poland become the top two locations. The top 10 locations except Belgium for the subsequent investment exactly coincide with the most preferred locations for their first-time decision even though there is slight difference between the respective ordering.

Table 2.2: Top 10 locations for different groups of German firms 1996-2009

Overall	Single time	Multiple Times		
		overall	$t = t_0$	$t > t_0$
USA	USA	FRA	FRA	FRA
FRA	AUT	AUT	AUT	POL
AUT	CHE	CHE	NLD	AUT
CHE	FRA	NLD	USA	USA
NLD	POL	GBR	GBR	GBR
GBR	NLD	USA	BEL	NLD
POL	GBR	ITA	CHE	CHE
CZE	CZE	POL	ITA	CZE
ITA	HUN	CZE	POL	ITA
ESP	ESP	ESP	CZE	ESP
6,408	4,041	2,367	2,367	2,367

t_0: first period (year) when the firm engaged in FDI

Two different conditional probabilities in Table 2.3 shed some light on whether investment history matters for German firms' FDI location choices. Variable $a_{it} = j$ denotes that firm i chooses location j for FDI in period t, and variable t_0 denotes the first period (year) when firms engaged in FDI in the sample. In brief, these two conditional probabilities show how large the difference in the probability of choosing the same location is given different initial conditions. The conditional probability ($Pr(a_{it} = j|a_{it_0} = j, t > t_0, a_{it} \neq 0)$) refers to the probability of choosing the same location as their first-time ($t = t_0$) decision conditional on that these firms make a new investment ($a_{it} \neq 0$) in the subsequent periods ($t > t_0$). As in Table 2.3, this conditional probability is significantly higher than another conditional probability, ($Pr(a_{it} = j|a_{it_0} \neq j, t > t_0, a_{it} \neq 0)$), defined as the probability of choosing the same location, but conditional on firms having not chosen it in the first period (($t = t_0$)). The contrast between these two conditional probabilities clearly presents an interesting persistence pattern in sequential FDI location choices. Location attractiveness, firm heterogeneity and state dependence are three main sources that could generate the observed persistence pattern in the data. If the incentive of firms to invest in the same location repeatedly over time is that the economic situation is attractive there, firms will engage in FDI in these countries independent of their previous experience. In this case, the two conditional probabilities discussed above should be very close to each other rather than differ as much as in Table 2.3. Thus the difference indicates that location attractiveness is not the only source for observed persistence and the remaining two channels play a vital role. However it is very hard to disentangle state dependence from firm heterogeneity based on mere descriptive statistics.

Table 2.3: Probability of choosing a location given different initial conditions

Location	$Pr(a_{it} = j\|a_{it_0} = j, t > t_0, a_{it} \neq 0)$	$Pr(a_{it} = j\|a_{it_0} \neq j, t > t_0, a_{it} \neq 0)$
FRA	0.21	0.08
AUT	0.13	0.06
CHE	0.17	0.06
NLD	0.20	0.06
GBR	0.16	0.08
USA	0.42	0.05
ITA	0.12	0.05
ESP	0.13	0.05
POL	0.39	0.05
CZE	0.27	0.05
China	0.65	0.01

t_0: the first period (year) when firm engaged in FDI.

2.2.2 Summary of the Data

To give a clear picture of which sources drive this location persistence in German firms' FDI decisions, I will provide the empirical evidence in the next subsection. It is necessary to summarize firm-year observations I am going to use for empirical estimation in the first place. Firstly, I focus on most of the OECD countries plus China and Hong Kong due to data availability on country characteristics. In spite of this relatively small choice set with only 24 locations, it covers around 83% of the locations chosen by German firms during the sample period. Secondly, I drop firms that have ever invested in countries outside of the newly defined "small choice set".[5] In addition, I also group 24 locations into different regions, mainly based on the geographical proximity and similarity in culture and economic development.[6] The reason for grouping these locations is discussed in next section. United States (USA), Canada (CAN) grouped as North America; France (FRA), Austria (AUT), Switzerland (CHE), Belgium (BEL), Ireland (IRL), Luxembourg (LUX) Netherland (NLD), Great Britain (GBR) grouped as Western Europe; Italy (ITA), Spain (ESP), Greece (GRC), Portugal (PRT) grouped as Southern Europe; Poland (POL), Czech Republic (CZE), Hungary (HUN), Slovak Republic (SVK) grouped as Eastern Europe; Denmark (DNK), Finland (FIN), Norway (NOR), Sweden (SWE) grouped as Northern Europe; plus China (CHN) and Hong Kong (HKG) grouped as Eastern Asia. Finally, I restrict the sample period to the years from 2002 to 2009 because of the availability of firms' attributes. During the sample period from 2002 to 2009, the most active German non-manufacturing firms engaging in FDI are in the following sectors: wholesale[7] 23.73%, household-related services 14.24%, real estate 7.92%, retail[8] 4.89% and business activities 4.68%.

A large number of German non-manufacturing firms, around 88.5%, choose at most one location for FDI every year. Of these firms 43% never engage in FDI again during

[5] I will explain why I drop these observations in detail when I present the structural model in the next section.

[6] I use the United Nations geoscheme as the main grouping criteria. The slight modification is that I put the United Kingdom and Ireland into Western Europe group; while leave only nordic countries in the Northern Europe group. This grouping of locations will be used as a basis for defining regional experience at later stages for estimation.

[7] The wholesale sector excludes motor vehicles and motorcycles.

[8] Motor vehicles, motorcycles repair of personal and household goods are excluded in this sector

during 2002 to 2009, because as in the previous table, most firms only invest once and these firms invested before 2002. Only 11.5% of firms choose more than one location for FDI in a given year and these observations will be dropped in order to be consistent with a discrete choice model.[9] Consequently, there are around 1700 firms left in the final sample for estimation. Table 2.4 gives summary statistics of the choices made in the final sample. The first column of Table 2.4 presents the number of locations chosen by German firms every period, while the last column contains information about the maximum number of locations chosen for FDI every year over the whole sample period. In the third column of Table 2.4, it shows that there are around 1100 firms making the first-time FDI during 2002-2009 and more than 80% of them choose only one location at their initial period ($t = t_0$). Consistent with previous observations that most of German firms only invest once, around 90% of firms do not invest every year in the subsequent periods ($t > t_0$). Therefore, I assume that firms can choose at most one location for FDI each period with little loss of generality. In the end, the number of firm and year observation reaches 12000, which provides a large sample for estimation.

Table 2.4: Summary of FDI action 2002-2009

locations(#)	Percentage	Timing	
	(%)	$t = t_0$	$t > t_0$
0	82.26	-	90.93
1	14.62	82.29	7.49
2	1.98	10.82	1.05
≥ 3	1.14	6.89	0.53
OBS(firm-year)	12,019	1,146	10, 873

[9]There are two alternative ways to deal with the multiple choices, i.e., engaging in FDI at more than locations at some period. The first is to redefine the choice set, incorporating all possible combinations of any two locations in the original choice set. But this implies there are potential combination effects between locations to be identified, which is far beyond the framework of this paper. The second approach is to randomly assign one of the location to be choice made in that period, and I then re-estimate the model for robustness check.

2.3 Empirical Model

When firm i sets up a new subsidiary in location j at period t, the latent profit (or attractiveness) π_{ijt} can be approximated by the following equation:

$$\pi_{ijt} = F_{ij} + C_j + \beta_1 s_{ijt} + \sum_k \alpha_k x_{jkt} + \varepsilon_{ijt} \tag{2.1}$$

where $j = \{1, \ldots, 24\}$ denotes locations; x_{jkt}, location characteristics: market size (measured by real gdp), GDP per capita (gdppc), economic growth rate (growth), labor cost, tax rate (tax), unemployment rate (unempt) and investment risk which is captured by corruption perception index (henceforth CI)[10]; Higher CI means less corruption, i.e. less investment risk in this country; C_j denotes location fixed effect, and firm heterogeneity (firm-location pair fixed effect) is captured by (F_{ij}). Variable investment history $s_{ijt} = \{0, 1\}$, denotes firm i's experience in this specific location. If firm i has invested in location j before period t, then $s_{ijt} = 1$; otherwise 0. Idiosyncratic profit shock ε_{ijt} is assumed to follow iid type 1 extreme value distribution across firms locations and time.

For the typical discrete choice model, I need to specify the outside option and normalize its latent profit (or attractiveness) to be a constant term for identification. In the current setting, the outside option would be naturally defined as not choosing one of 24 locations to invest, which includes no FDI or choose some location outside of the choice set to invest. However, the latent profit from investing in some location is systematically different from no investing, this will create a valid concern if I normalize the outside option to be constant value. To mitigate this concern, I will drop the observations of firms that have ever invested outside of the choice set as mentioned before. Consequently, if firm i chooses outside option ($j = 0$) in period t, which means this firm decides not to invest this period, then the corresponding attractiveness is normalized as 0 plus iid type 1 extreme value distributed ε_{i0t},

$$\pi_{i0t} = \varepsilon_{i0t}$$

[10]All these annual location attributes: gdp, gdppc, growth rate, unemployment rate and average tax rate on profit are from World Bank: World Development Indicators. Labor cost is measured by hourly wages in the manufacturing sector in US $, from the source of Bureau of Labor Statistics. CI is from Corruption Perceptions Index, Transparency International.

Therefore, if I observe firm i setting up one new subsidiary at location j in period t, i.e. $a_{it} = j$, it must be that the latent profit in location j is higher than in any other location for firm i in this period,

$$P(a_{it} = j) = P(\pi_{ijt} > \pi_{ikt}, \forall k \neq j)$$

To estimate the above discrete model, I will apply conditional logit approach to obtain these parameters. The key parameter of interest is β_1, because I want to identify whether firms' previous FDI location choice affect their decision on which location to choose for their current FDI. This main identification goal requires further refinement of the observations in the sample. In the current sample, firms face 24 locations plus not investing in the choice set. The estimation results based on this sample make it difficult to interpret the parameter β_1. Because the coefficient β_1 now confounds with effect whether firms' previous FDI experience makes it more likely to invest or not, not only effect that which location is affected to be chosen. For instance, if β_1 is significantly positive, I can only infer that firms' previous FDI experience make them more likely to invest or make them more likely to choose the location where they have invested before to invest. If β_1 is significantly negative, it could be that the effect that firms' previous experience make them less likely to invest dominates the that firms prefer the same location for FDI conditional on investing. Both signs can make sense as follows: firms with more international investment experience seems more active in FDI activity (positive sign); while due to the diminishing return to the new affiliates, firms already having a large number of affiliates abroad may be less likely to invest again (negative sign). Therefore, I will drop year observations when firms did not make a FDI in order to get a pure effect about whether firms' FDI location history affect their current FDI location choice by conditioning on observing a FDI this period.

2.3.1 Models of firm heterogeneity

$$F_{ij} = \omega_{ij} + \lambda C_j z_i \tag{2.2}$$

F_{ij} is used to capture firm heterogeneity in equation 2.1. I decompose firm heterogeneity into observed (z_i) and unobserved parts (ω_{ij}). The observed heterogeneity is measured by firms' persistent attributes: industry and size. Since the sample has been restricted to firms in non-manufacturing industry, I only need to distinguish further whether firms are in financial service industry or not. For firms in financial service sectors, they might have a persistent preference towards certain countries, like Luxembourg, Switzerland because of various regulations. In terms of heterogeneous firm size,[11] small firms may strongly prefer to invest in neighboring countries; while large firms may invest more broadly (Yeaple, 2009). These kinds of observed heterogeneity are captured by the interaction term between country fixed effect and observed firm characteristics.

On the other hand, the unobserved firm heterogeneity reflects some determinants which might affect firms' decision in choosing the location for FDI, but unobserved in the data. For instance, the CEO of a firm comes from certain county, or the products of this firm are made tailed for that specific market (country); therefore, this firm keep investing in that country, which can only be rationalized by unobserved heterogeneity (ω_{ij}). There are two common alternatives to model the unobserved firm heterogeneity in the literature. The first alternative is to assume ω_{ij} is the "fixed effect" between firm i and country j for all i and j without making any additional assumption about the distribution of ω_{ij}. That is to say, we need to create the interaction between a set of individual firm dummy variables and a set of country dummy variables in the estimation procedure.[12] However, this method will generate typical incidental parameters problem which there around 2000 individual firms and 24 locations in my model. Chamberlain (1980) proposes a brilliant approach-conditional maximum likelihood estimation to get around this incidental parameters problem. The intuition of Chamberlain's method is to take out the incidental parameters of the likelihood expression by conditioning on

[11]Firm size is measured by total assets, which seems not to be qualified as persistent attributes. But I group firms into three classes: small, medium and large size according to the mean value of their total assets during 2002-2009. I also use the first and last years' total assets to generate the same class and find that around 89% of firms stay in the same category in the two different periods. Therefore, it is reasonable to assume firms' size is constant during the sample period.

[12]The interactions between individual firm dummy and country dummy will suppress the observed term ($C_j z_i$) in equation (2.2).

their sufficient statistics. The likelihood function is then only a function of structural parameters and can be maximized to yield consistent estimators.

However, if the explanatory variables include strictly exogenous variables, lags of the endogenous variable and the unobserved individual heterogeneity, the estimation results by conditional maximum likelihood will be biased. Honoré and Kyriazidou (2000) consider the identification for that case and propose an consistent estimation method under certain regularity conditions. Chintagunta, Kyriazidou, and Perktold (2001) using a panel data on household purchasing behavior estimate a discrete choice model with observed individual and brand characteristics, the lagged choice and unobserved individual brand fixed effect. They compare the conditional maximum likelihood procedure with the method proposed by Honoré and Kyriazidou (2000) and find that the estimated coefficient for lagged choice under conditional maximum likelihood procedure is significantly underestimated towards zero; while the approach of Honoré and Kyriazidou (2000) enable them to correct that bias. In this paper, I will consider the "fixed effect" model and provide a lower bound for coefficient (β_1) by using the conditional maximum likelihood approach.[13] If I can find positive low bound for β_1, it is enough to show the existence of true state dependence after controlling for firm heterogeneity.

Another alternative is to model ω_{ij} as "random effect", which is also widely used in the literature. This approach requires to specify the conditional distribution of the unobserved heterogeneity. The common assumption is that ω_{ij} follows multivariate normal distribution conditional on the observed covariates. Coupled with the extreme value type 1 distribution of iid profit shock, this standard normal mixture of logit model will yield consistent estimator for β_1. Despite the concerns of misspecification of the distribution of unobserved heterogeneity (ω_{ij}), I will also apply "random effect" approach for robustness check of state dependence.

[13] The method proposed by Honoré and Kyriazidou (2000) rule out the class of models with the time dummy and trend in the covariates, resulting in being unapplicable in my model, because the market size measured by GDP grows over time for most of the countries in the data.

2.4 Estimation Results

Since the latent profit π^*_{ijt} is not observable in the data, I will apply McFadden's conditional logit model to estimate the parameters in equation (2.1). As I mentioned in the previous section the regression results in Table 2.5 is based on the observations in which firms make an investment, because I am interested in the factors that determine firm's FDI location choice conditional on firms making an investment and in particular whether experience rather than heterogeneity matters to that location decision.

It is worth mentioning how I control for the firm-location fixed effect (firm heterogeneity) F_{ij} in different columns. In column 1 of Table 2.5, I do not control any firm heterogeneity, but only country characteristics (C FE denotes country fixed effect). The estimated coefficient for β_1 is not surprisingly positive and significant. As I discussed earlier, this coefficient may confound with firm heterogeneity and does not imply any state dependence. From column S2 to column S4 in Table 2.5, I use the dummies of firm type in terms of sector and size to interact with country fixed effect to control for that observed firm heterogeneity. To obtain consistent estimation of β_1, I need to assume that the all firm-country fixed effect are constant for firms within each type, i.e., there is no other unobserved firm-country fixed effect. The estimation results under this approach shows that the coefficient of firms' experience still has significantly positive effect on their current location choice even after controlling for observed firm heterogeneity. The magnitude of β_1 in column S2 slightly decreases compared with the value in column S1 without controlling for any firm heterogeneity, which implies observed firm heterogeneity does partially explain the persistence in FDI location choice. In column S3, I add the interaction between firms' country-specific investing experience with corruption perceived index (CI). The coefficient of this interaction term is significantly negative, implying that experienced firms are less sensitive to investment risk captured by CI. This negative sign provides a potential explanation why we observe this persistence in FDI location choice, which will be addressed in detail in next Chapter.[14] Column S4 shows that firms' region-

[14]The negative sign of the interaction term provides a clear direction and magnitude how it affects the latent profit; while the general marginal effect of the interaction term is not necessarily negative.

specific investing experience works in the same direction as country-specific experience. The region-specific experience is equal to 1 if the firm has invested in any country within the same region as defined in previous section, otherwise 0. This result confirms early finding in the literature of management that multinational firms prefer to expanding from countries with similar economic and cultural environment.

From column S5 to S7, I apply the "fixed effect" approach to control for unobserved firm heterogeneity by exploring the panel structure conditional on the sufficient statistics as Chamberlain (1980) suggested. However, in my model, Chamberlain's approach yields an inconsistent, but only a lower bound for the coefficient (β_1) of experience variable as discussed in the previous section. Compared results from column S2 to S4, the experience coefficient is underestimated biased towards zero, the same as shown in Chintagunta et al. (2001). Nevertheless, the lower bound of β_1 still significantly positive, with magnitude around 0.3 in column S5. This estimation result clearly identify the presence of state dependence in firms' FDI location choice behavior; that is, firms are more likely to keep investing in the same country where they have invested before. In column S6, I add the same interaction term as in column S3. This interaction term also keeps significantly negative coefficient after controlling for unobserved firm heterogeneity. The last column S7 provides a robustness check for firms regional experience and confirms that it also plays the same role as the of country-specific investing experience in the FDI location decision with unobserved firm heterogeneity taken into account.

Before I present the estimation results under "random effect" of F_{ij} approach in another table, I first briefly discuss other determinants of FDI. As in Table 2.5, market size[15] has a significantly positive effect for firms FDI location decision, which support the theory of horizontal FDI that firms are making such kind of investment to maximize profit in the local market. This positive effect of market size is very robust across all specifications. Besides market size, unemployment rate (unempt) also has a significantly positive effect to attract horizontal FDI. The possible explanation could be that when a country is experiencing high unemployment rate, local government would introduce a set

[15] Growth rate is one period lagged value and all other variables are in current value.

Table 2.5: Results of conditional logit model

	S1	S2	S3	S4	S5	S6	S7
gdp	1.0502**	0.9544*	1.0265*	1.1726**	6.5286***	6.6428***	6.5855***
	(0.5163)	(0.5520)	(0.5375)	(0.6131)	(0.0083)	(0.0095)	(0.0093)
gdppc	-0.0172	-0.0007	0.0006	-0.0172	-0.0435***	-0.1084***	-0.0733***
	(0.0460)	(0.0483)	(0.0467)	(0.0446)	(0.0057)	(0.0068)	(0.0065)
growth rate	0.0671*	0.0717**	0.0714**	0.0781**	-0.0147	-0.0222	-0.0444
	(0.0343)	(0.0356)	(0.0356)	(0.0306)	(0.0343)	(0.0341)	(0.0338)
labor cost	0.0076	0.0071	0.0092	0.0160	0.0844***	0.0988***	0.0968***
	(0.0171)	(0.0159)	(0.0157)	(0.0156)	(0.0124)	(0.0125)	(0.0097)
tax	-0.1653	-0.7569	-0.8039	-0.9520	-0.3956***	-0.4508***	-0.3067**
	(1.5121)	(1.5085)	(1.5497)	(1.4464)	(0.1148)	(0.0453)	(1.2780)
unempt	0.0390***	0.0386***	0.0365***	0.0427***	0.0392***	0.0515***	0.0403***
	(0.0117)	(0.0117)	(0.0118)	(0.0139)	(0.0067)	(0.0068)	(0.0065)
CI	0.0474	0.0492	0.1139*	0.1993***	-0.0754	-0.0615	-0.0771
	(0.0638)	(0.0592)	(0.0618)	(0.0773)	(0.0483)	(0.0542)	(0.0413)
experience-country	2.4009***	2.3390***	4.0740***		0.3153***	2.0644***	
	(0.2823)	(0.2695)	(0.4642)		(0.1115)	(0.1856)	
experience-region				4.4256***			1.3931***
				(0.5371)			(0.3263)
experience-country*CI			-0.2460***			-0.2424***	
			(0.0420)			(0.0276)	
experience-region*CI				-0.3349***			-0.1853***
				(0.0463)			(0.0408)
C FE	✓	✓	✓	✓	✓	✓	✓
C*sector FE		✓	✓	✓	✓	✓	✓
C*size FE		✓	✓	✓	✓	✓	✓
Unobserved FE					✓	✓	✓
$-L$	-7543.68	-7534.02	-7399.13	-7655.64	-173.43	-173.37	-173.07

Notes: CI denotes corruption perceived index; ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively.

of various investment promotion policies in order to reduce unemployment. Regarding the variable average company profit tax rate (tax), it has a negative impact on FDI as expected in all specifications. But it only shows the significance when I start to control for unobserved firm heterogeneity. Corruption perceived index (CI) shows a positive sign before controlling for unobserved heterogeneity, though not very significant. It is in line with conventional wisdom that countries with lower investment risk are relatively attractive to horizontal FDI. However this effect become mixed after introducing unobserved heterogeneity. The same mixed effect occurs to another determinant variable-economic growth rate. It has the significantly positive effect as expected before specification S5 and turn insignificantly negative afterwards. This could be related to the way of controlling for unobserved heterogeneity, under which the "fixed effect" approach could cause some underestimation as discussed in previous section. Labor cost and GDP per capita both show unexpected signs, because high labor cost would deter firms from engaging in FDI there given everything else being equal. One possible explanation would be that the labor cost is not a good measure for the real wage cost in non-manufacturing industry since the variable labor cost used here is from International Comparisons of Hourly Compensation Costs in Manufacturing industry issued by bureau of labor statistics, United States. Another potential interpretation is that the variable GDP per capita (gdppc) seems to capture the general wage cost across countries, which shows a negative effect to attract FDI; while labor cost in manufacturing industry is more likely to capture the similarity in economic endowments between Germany and other countries. The positive coefficient of labor cost implies high labor cost indicates the labor endowments of that country may be more similar to the structure in Germany, since Germany also has the high labor cost in general. Then in theory these countries are relatively more attractive to horizontal FDI from Germany.

The estimation results of "random effect" approach are presented in Table 2.6. The unobserved firm heterogeneity (ω_{ij}) in equation (2.2) is assumed to follow independent multivariate normal distribution, i.e., the covariance matrix is diagonal.[16] The associated

[16]It is reasonable to assume firms' idiosyncratic preference of each location is independent since I have controlled for their preference towards most of the observed location characteristics as well as location

Table 2.6: Results of mixed conditional logit model

	S1	S2	S3	S4	S5	S6	S7
Mean							
gdp	0.1198**	0.1556**	0.1603***	0.0710	1.0503**	1.0475**	1.0024**
	(0.0504)	(0.0606)	(0.0597)	(0.1737)	(0.5163)	(0.5084)	(0.4904)
gdppc	-0.0134	-0.0286**	-0.0305***	-0.0357**	-0.0172	-0.0155	-0.0147
	(0.0098)	(0.0112)	(0.0112)	(0.0143)	(0.0460)	(0.0438)	(0.0384)
growth rate	0.0747***	0.0767***	0.0791***	0.0772***	0.0671*	0.0685**	0.0704**
	(0.0204)	(0.0209)	(0.0219)	(0.0203)	(0.0343)	(0.0340)	(0.0292)
labor cost	-0.0130	-0.0048	-0.0034	-0.0016	0.0076	0.0095	0.0154
	(0.0125)	(0.0123)	(0.0120)	(0.0160)	(0.0168)	(0.0125)	(0.0161)
tax	-2.4000***	-3.0526***	-3.0149***	-3.2250***	-0.7699	-0.7080	-0.9004
	(0.6530)	(0.7036)	(0.7002)	(0.9753)	(1.5121)	(1.5355)	(1.3508)
unempt	0.0298**	0.0328***	0.0304**	0.0291*	0.0390***	0.0371***	0.0426***
	(0.0119)	(0.0120)	(0.0118)	(0.0151)	(0.0117)	(0.0119)	(0.0144)
CI	0.1070	0.1114	0.1748**	0.3287***	0.0474	0.1042	0.2017**
	(0.0701)	(0.0710)	(0.0734)	(0.0841)	(0.0638)	(0.0673)	(0.0783)
experience-country	2.4282***	2.3869***	4.0643***		2.4009***	3.9562***	
	(0.2866)	(0.2859)	(0.5213)		(0.2823)	(0.4583)	
experience-region				4.3907***			4.2812***
				(0.0486)			(0.5484)
experience-country*CI			-0.2347***			-0.2206***	
			(0.0459)			(0.0399)	
experience-region*CI				-0.3023***			-0.2932***
				(0.0486)			(0.0483)
C FE	✓	✓	✓	✓	✓	✓	✓
C*sector FE		✓	✓	✓			
SD							
C FE	✓	✓	✓	✓			
gdp					0.0025	0.0025	0.0025
$-L$	-7549.85	-7484.30	-7452.69	-7740.97	-7985.68	-7513.97	-7726.99

Notes: CI denotes corruption perceived index; ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively.

estimation results from column S1 to S4 show that firms' previous investing experience indeed has a positive effect for the FDI location decision in the current period even I model firm heterogeneity a different way. Moreover, firms' regional investing experience have the same effect under this setting and both of the significance is very robust no matter whether I include observed firm heterogeneity measured by sector dummy or not.

From column S5 to S7 in Table 2.6, I use firms' heterogenous ability in exploring the market to capture their heterogeneity, i.e., random coefficient for market size (gdp). The intuition for this random coefficient is that firms' persistent FDI location choice maybe due to their persistently different ability in making profit from specific markets. That is to say, if we observe one firm keeps investing in one location, it might be that this firm is able to make better use of the market in that location than other firms. Again, firms' experience in both country and region level also has the significantly positive effect as before under this specification. In addition, the mean attractiveness of large market size is also consistently high.

Finally, rather than use a dummy whether firm has operated in this location or region before, I try another measure to represent country-specific and regional experience: the time span of firm that has been present in this country or the region before I observe another new FDI in the data. Given the first subsidiary of firm i set up in location j on period t_0, then variable $span_{ijt}$ is defined as:

$$span_{ijt} = \begin{cases} t - t_{0ij}, & \text{if } t \geq t_{0ij}; \\ 0, & \text{otherwise.} \end{cases}$$

Where t_{0ij} denotes the first period (year) I observe firm i setting up a subsidiary in location j. It would be reasonable to assume that firms' knowledge/experience about specific country is positively correlated with their time length of being present there. Therefore, $span_{ijt}$ is a relatively more informative measure than dummy variable since it captures how long firms have been operating in each individual country. This new

fixed effect (C FE). In addition, there is one technical restriction to run mixed conditional logit model in STATA that the number of random coefficient should be less than 20 due to computational expense. Therefore, I treat countries, like Austria, Belgium, Luxembourg, Netherland, Switzerland and United Kingdom as the base and then left 18 country dummies were considered to have random coefficients.

measure yields a consistent and robust result as in Table 2.7, very similar to previous two tables. To sum up, the empirical evidence establishes that experience in both the specific country and the region has a positive effect on firms' subsequent FDI location decisions due to "true state dependence".

Table 2.7: Results of (mixed) conditional logit model

	S1	S2	S3	S4	S5	S6	S7
Mean							
gdp	1.1148**	0.9962	0.8332	0.1773***	0.2084***	1.1149**	0.9650*
	(0.5138)	(0.6496)	(0.7343)	(0.0555)	(0.0635)	(0.5139)	(0.5643)
gdppc	-0.0119	-0.0045	0.0029	-0.0125	-0.0312***	-0.0118	-0.0052
	(0.0438)	(0.0504)	(0.0518)	(0.0093)	(0.0112)	(0.0438)	(0.0444)
growth rate	0.0823**	0.0875**	0.0883**	0.0902***	0.0937***	0.0823**	0.0831**
	(0.0327)	(0.0350)	(0.0362)	(0.0229)	(0.0270)	(0.0327)	(0.0335)
labor cost	0.0184	0.0179	0.0187	-0.0047	0.0055	0.0184	0.0189
	(0.0160)	(0.0150)	(0.0148)	(0.0122)	(0.0130)	(0.0160)	(0.0156)
tax	-0.5606	-0.6030	-0.5621	-2.3884***	-3.3837***	-0.5610	-0.4957
	(1.4375)	(1.4984)	(1.5514)	(0.6788)	(0.7560)	(1.4373)	(1.4734)
unempt	0.0436***	0.0429***	0.0426***	0.0329**	0.0332*	0.0436***	0.0433***
	(0.0141)	(0.0147)	(0.0146)	(0.0137)	(0.0176)	(0.0141)	(0.0139)
CI	0.0388	0.0411	0.0689	0.1297*	0.1617**	0.0388	0.0620
	(0.0619)	(0.0584)	(0.0608)	(0.0716)	(0.0789)	(0.0619)	(0.0649)
span-country	0.2993***	0.2915***	0.5344***	0.3075***	0.7870***	0.2993***	0.4924***
	(0.5235)	(0.0496)	(0.0844)		(0.1800)	(0.0525)	(0.0863)
span-country*CI			-0.0347***		-0.0640***		-0.0277***
			(0.0082)		(0.0189)	(0.0399)	(0.0089)
C FE	✓	✓	✓	✓	✓	✓	✓
C*sector FE		✓	✓		✓		
C*size FE		✓	✓				
SD							
C FE				✓	✓		
gdp						0.0097	0.0098
$-L$	-7985.69	-7851.94	-7833.05	-8000.05	-7891.71	-7985.68	-7973.14

Notes: CI denotes corruption perceived index; ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively.

2.5 Conclusion

In this paper, I use micro-firm level data to study the determinants of horizontal FDI as well as its location patterns. The main contribution of this paper to the literature is to demonstrate that multinational firms' investment history does matter for their current FDI location choice, in particular, firms are more likely to invest in the same country where they have invested before. In addition to accounting for various country characteristics, I explicitly model a rich set of observed and unobserved firm heterogeneity persistent over time as well as idiosyncratic location preference in order to clearly identify the presence of "true state dependence" in firms' FDI location patterns. The observed firm heterogeneity includes size and sector classifications, which are commonly used in the literature. With respect to persistent unobserved heterogeneity, both "fixed effect" and "random effect" approaches yield the consistent and robust estimation results that firms' previous country-specific and and region-specific investment experience both have a significantly positive effect in their current FDI location decision.

Regarding other determinants of horizontal FDI, I find that large market size, high economic growth and high unemployment rate are the most attractive to multinational firms. The first two factors are consistent with previous findings in the literature; while the unemployment rate most likely captures the positive effect from investment promotion policy.

These findings have an important policy implication for host countries to attract FDI: the short-term FDI promotion schemes may have a long-run positive effect. Take the commonly used tax concession policy for example, even though it is designed to last only for the first two or three years, some new FDI are attracted to enter this country in the short run. However, since there exists state dependence, or persistence in firms' FDI location choice as I find in the data, those firms may still prefer to investing in the same country even without any FDI promotion policy in the future.

While this paper focuses on how heterogenous firms choose locations to engage in FDI, it can be extended in two main directions. First, as most of the empirical paper in the literature, I assume firms make a static decision to invest in one country, even though some variable like economic growth may capture long-run benefit in a reduced-form way. Based on my finding that firms' investment history matters for their current decision, the best way would be explicitly model how firms make the dynamic decision to choose the country sequentially to engage in FDI. Second, I will try to provide an explanation about why we observe such persistence pattern in firms' FDI location choice and seek to investigate which factors drive the state dependence observed in the data. These two possible extensions will be addressed in the next Chapter.

Chapter 3

Quantifying Adaptation Costs in Sequential FDI Location Choices: Evidence from German Firms

3.1 Introduction

Foreign direct investment (FDI) has been constantly attracting a substantial amount of attention from both economists and policy makers. Global FDI flows reach the peak, around $2 trillion in 2007, and the value-added activity (gross product) of foreign affiliates worldwide account for 11% of global GDP in that year.[1] On one hand, a vast literature on FDI, (Carr et al. (2001), etc) has empirically sought to explain what fundamental factors make host countries attractive to multinational corporations (henceforth MNCs). On the other hand, most papers downplay the determinants driving the dynamics of MNCs' FDI process. However, earlier papers in the management literature (Johanson and Vahlne (1977), etc) have demonstrated that MNCs' international expansion is "a process rooted in uncertainty reduction" through the accumulation of experience. They also find the interesting sequential FDI pattern: firms in the initial stage of foreign expansion exhibit a strong preference for markets which are culturally and economically similar to the home country; as firms gain international experience, they start to consider investing on a much

[1]Data source: World Investment Report 2012.

wider range of locations.[2]

The lack of knowledge about foreign culture, social norms and legal systems can be a major impediment to MNCs' global expansion. Often MNCs are able to acquire this country-specific knowledge only by operating in that local market. That is why the presence of subsidiaries in a foreign market will increase the multinational firm's propensity to make subsequent investment in that market. In this paper, I introduce adaptation costs for first-time foreign entrants to capture the fact that it is costly for MNCs to adapt to different institutions and economic environments. This costly adaptation process does not occur if firms decide to invest in the same country again in the future.

In the Chapter 2, I have empirically established that MNCs are more likely to invest in the same country where they have invested before. Given the state dependence in MNCs' FDI location decisions, I propose a dynamic discrete choice model, in which firms sequentially choose a location for FDI to maximize the expected discounted profit with adaptation costs taken into account. Building upon the methodology developed by Berry, Levinsohn and Pakes (1995, 2004) and Gowrisankaran and Rysman (2012), each location in this model is characterized by a bundle of both observed and unobserved attributes. The profit earned by MNCs is approximated as a function of both location and firm characteristics. An obvious concern is that the observed persistence in firms' location choices may not be due to adaptation costs reduction by prior experience. Rather, firms simply have different preferences over locations for unknown reasons that are unrelated to their past decision history. Thus, heterogeneous preferences across firms necessarily need to be taken into account in order to isolate the spurious state dependence (Heckman 1981).

Based on the same plant-level panel data on German non-manufacturing firms' FDI behavior as in the Chapter 2, I go further to quantify the magnitude of adaptation costs using the newly developed methodology from empirical industrial organization. The empirical results imply that adaptation costs are substantial and varying across locations, ranging from 0.5%−29% of the average expected discounted profit[3]. With these estimation

[2] Location and country are used interchangeably throughout this paper.

[3] "average" is referring to the mean value across all observed German non-manufacturing firms over

results, I simulate several policy experiments involving FDI promotion schemes. These counterfactuals suggest that firms with distinct experience states respond systematically differently to different FDI promotion policies. In general, the expected discounted profit across all firms increases by around 15%, on average over time when adaptation costs are eliminated. Most importantly, this increment in profit is largely due to better matching between locations and firms. The information that adaptation costs have a significant influence on firms' FDI location choices is particularly important for both host and home countries. For home countries, the results provide some insight into the benefit of the policy on subsidizing firms when entering foreign markets. For host countries, the right promotion policy will effectively attract FDI which may boost local economic growth.

The main contribution of this paper to the FDI literature is twofold. First, by using detailed plant-level panel data coupled with the newly developed methodology from industrial organization, I am able to quantify the distinct magnitudes of adaptation costs across different locations. Adaptation costs are critical to policy evaluation, but their magnitude have not yet been estimated. These costs can be identified only through their nonlinear effects on dynamic entry patterns generated by firms in different states, e.g. whether entering for the first time. Moreover, with the estimated adaptation costs as well as MNCs' preference towards location characteristics, I can investigate the impact of a series of counterfactual FDI promotion policies on MNCs' new FDI patterns. Second, I explicitly model how MNCs make the dynamic decision to choose the location to engage in FDI. Based on the finding from Chapter 2 that firms' investment history does matter for their current FDI location decision, the best strategy to include the intertemporal links in firms' decision process is to employ a dynamic model. When the forward-looking MNCs make decision about which location to engage in FDI today, they automatically take into account the option value of their today's choices. For instance, we often observe some MNCs decide to enter a emerging market even though it is not profitable at that time, the forward-looking behavior indicates that the future profit in that market must be high enough to compensate the current loss. While in a static model, this investment

the entire sample period.

behavior can only be rationalized by unobserved favorable profit shock.

The remainder of the paper is organized as follows. First, I relate this paper to the existing literature. Next, I describe the dataset and show empirical evidence of persistence in German firms' sequential FDI pattern in following section. Then I present the details of the model to be estimated and discuss how it takes into account findings from the data. After that I lay out the estimation strategy and identification argument. Finally, the estimation results and counterfactual analyses are discussed and then I conclude.

3.2 Related Literature

My paper complements the earlier literature on how MNCs expand globally (e.g., Johanson and Vahlne (1977), Davidson (1980)). They demonstrate that international expansion is a process rooted in uncertainty reduction through the accumulation of relevant types of experience. Barkema et al., (1996) and Shaver et al., (1997), Delios and Henisz (2003) provide empirical evidence for the importance of organizational learning in firms' internationalization and find that country-specific experience can help firms reduce entry barriers like cultural distance and political hazard for the subsequent FDI in the same or related countries. These papers establish the persistence in firms' sequential FDI patterns, but they ignore unobserved firm heterogeneity that could also generate the same observed persistence as the role of experience. In this paper, I explicitly account for unobserved firm heterogeneity to show that the observed state dependence in the data is indeed driven by the presence of adaptation costs.

Most papers in the empirical literature on FDI focus on a static setting to study the important determinants for different types of FDI. Carr et al., (2001) show that similarity in market size and economic endowments between countries are important for horizontal FDI, while labor cost is relatively more important for vertical FDI. More recent papers start to investigate the role of firm heterogeneity in productivity for FDI, e.g. Aw and Lee (2008), Yeaple (2009), Chen and Moore (2010). They find that firms are sorted to choose the location for FDI: more productive firms invest in a larger number of foreign countries and can also access countries with less attractive attributes; while less productive firms only concentrate on a smaller set of countries with better location attributes.

In the related trade literature, several recent papers study the pattern of sequential exporting, such as Roberts and Tybout (1997), Das, Roberts, and Tybout (2007), and Morales, Sheu, and Zahler (2011). They empirically identify a similar persistence pattern in the dynamic exporting behavior. In particular, Roberts and Tybout (1997) first infer the presence of sunk entry costs from the persistence in exporting patterns in Colombian manufacturing firms. Das et al., (2007) structurally estimate the sunk entry costs using the same dataset and find these costs to be at least $\$344,000$ (in 1986 U.S. dollars). These estimated sunk costs are interpreted as the average costs to break in a new market. In this paper, I go further and allow for these costs to differ across different locations. Consequently, firms not only make decisions about whether to invest or not, but also decide which location to choose conditional on engaging in FDI. This is in contrast to the binary decision (whether to export or not) setting in Das et al., (2007).

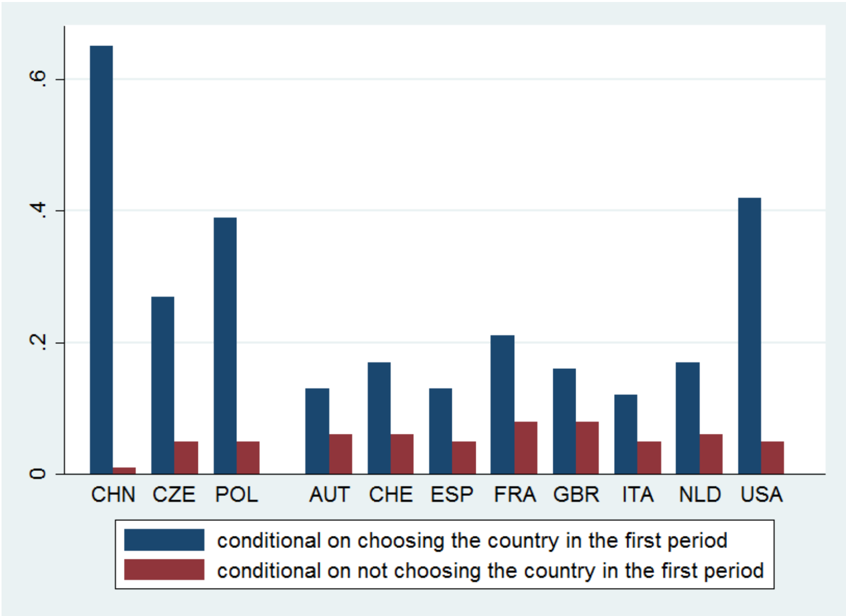
As for the methodology, I heavily rely on the newly developed estimation strategy from empirical industrial organization, involving Berry, Levinsohn and Pakes (1995, 2004), Shcherbakov (2009), Nosal (2011), and Gowrisankaran and Rysman (2012). These approaches allow me to disentangle adaptation costs from the firm heterogeneity as well as both observed and unobserved location attributes. Additionally, I can study the nonlinear effect of adaptation costs on firms' sequential investment decisions.

3.3 Data

In this paper, I use the same dataset as in Chapter 2-Microdatabase Direct Investment (MIDI) provided by Deutsche Bundesbank (German Central Bank) from 2002 to 2009. The detailed description can be found in Chapter 2, thus it is skipped in this Chapter. In addition, following the same procedure in Chapter 2, I also restrict the sample to be composed of non-manufacturing firms. FDI in this paper is again defined as a new affiliate over the reporting threshold set up by German parent firms. In particular, the shareholding by the parent firm must be larger than 50%, i.e., an absolute majority shareholder for the investment to be considered as an FDI.

Given the same sample as well as the same definition of FDI, Figure (3.1) replicates the same persistence pattern in FDI location choice as I find in Chapter 2.

Figure 3.1: Probability of choosing a location given different initial conditions



As we can see in Figure (3.1), the average probability across all firms of choosing a location to engage in FDI conditional on having invested in that location before is much higher than conditional on not investing there before. The difference between these two conditional probabilities are prevalent across all locations. Take the extreme case China (CHN) for example, if one firm invested in China before, then the probability of this firm choosing China to invest is 60 times higher than other firms that did not invest in China before.

3.3.1 Empirical evidence of adaptation costs

In Chapter 2, I have already addressed the issue of firm heterogeneity that may cause the "spurious state dependence" in a comprehensive way. In this paper, I summarize the estimation results from previous Chapter.

Table (3.1) contains summary of estimation results with controlling for observed firm heterogeneity, unobserved heterogeneity under "fixed effect" approach and unobserved

Table 3.1: Summary of estimation results of (mixed) conditional logit model

	S1	S2	S3	S4	S5	S6	S7
Mean							
experience-country	2.4009***	2.3390***	0.3153***	2.0644***		4.0643***	
	(0.2823)	(0.2695)	(0.1115)	(0.1856)		(0.5213)	
experience-region					1.3931***		4.3907***
					(0.3263)		(0.0486)
experience-country*CI				-0.2424***		-0.2347***	
				(0.0276)		(0.0459)	
experience-region*CI					-0.1853***		-0.3023***
					(0.0408)		(0.0486)
gdp	+**	+*	+***	+***	+***	+***	+
gdppc	-	-	-***	-***	-***	-***	-**
growth rate	+*	+**	-	-	-	+***	+***
labor cost	+	+	+***	+***	+***	-	-
tax	-	-	-***	-***	-**	-***	-***
unempt	+***	+***	+***	+***	+***	+**	+*
CI	+	+	-	-	-	+**	+***
C-FE	✓	✓	✓	✓	✓	✓	✓
C-FE*sector		✓	✓	✓	✓	✓	✓
C-FE*size		✓	✓	✓	✓		
Unobserved FE			✓	✓	✓	✓	✓
SD							
C-FE						✓	✓
$-L$	-7543.68	-7534.02	-173.43	-173.37	-173.07	-7452.69	-7740.97

CI: corruption perceived index; C-FE: country fixed effect.

Notes: ***, ** and * denote significance at the 1%, 5% and 10%, respectively.

heterogeneity under "random effect" approach respectively. It is clear that firms' both country-specific and region-specific experience have a significant positive effect across all specifications.[4] The "true state dependence" in firms' FDI location patterns provides convincing evidence for the presence of adaptation costs, i.e., MNCs have to incur huge costs by adapting to different business practices, and legal system, etc when entering a foreign country for the first time. The valuable experience from prior operations enables firms to more easily adjust themselves to local market environment if they invest in the same country.

3.4 Model

Based on the previous empirical evidence, I present a structural model in this section, accounting for the experience effect. In brief, the timing of decision process follows: in the beginning of each period, firms observe all the relevant information that affects their profit. The information includes observed and unobserved location characteristics, which are exogenous and common to all firms, firms' own FDI history involving which countries they have invested before as well as the number of subsidiaries in each country, and firms' idiosyncratic profit shock. After forming a expectation of the relevant future information, firms decide whether to engage in FDI, and also need to choose the location for FDI conditional on investing to maximize expected discounted profit. There is no strategic interaction between firms and profit flow is realized in the end of each period.

[4] Regions are defined in the same way as in Chapter 2.

Firm i's profit flow in period t, π_{it} is given by:

$$\pi_{it} = \begin{cases} -\eta_j \mathbb{1}_{(n_{ijt}=0)} + f_i(\vec{X}_{jt}, n_{ijt}) + \varepsilon_{ijt} + \sum_{j=1} n_{ijt} f_i(\vec{X}_{jt}, n_{ijt}) & \text{if } a_{it} = j \\ \\ \sum_{j=1} n_{ijt} f_i(\vec{X}_{jt}, n_{ijt}) + \varepsilon_{i0t} & \text{if } a_{it} = 0 \end{cases}$$

where

$$f_i(\vec{X}_{jt}, n_{ijt}) = \sum_k (\bar{\alpha}_k + \sum_r \tilde{\alpha}_{kr} z_{ir}) x_{jkt} + \xi_{jt} - \alpha_I I_t - \alpha n_{ijt}$$
$$= \sum_k \bar{\alpha}_k x_{jkt} + \xi_{jt} - \alpha_I I_t + \sum_k \sum_r \tilde{\alpha}_{kr} z_{ir} x_{jkt} - \alpha n_{ijt}$$

In general, π_{it} is composed of two parts: profit from the new subsidiary ($a_{it} = j$) set up this period, $(-\eta_j \mathbb{1}_{(n_{ijt}=0)} + f_i(\vec{X}_{jt}, n_{ijt}) + \varepsilon_{ijt})$ and profit from existing subsidiaries $(\sum_{j=1} n_{ijt} f_i(\vec{X}_{jt}, n_{ijt}))$. η_j denotes adaptation costs in location j and is assumed to be one kind of fixed cost constant over time[5] but different across countries. $n_{ijt} = \{0, 1, \ldots, \bar{n}\}$[6] is the total number of affiliates belonging to firm i at location j before period t;[7] $\mathbb{1}_{(n_{ijt}=0)}$ is an indicator function, equal to 1 if $n_{ijt} = 0$, otherwise 0. If firm i sets up a subsidiary in location j for the first time, i.e $n_{ijt} = 0$, then she has to pay the costs η_j to start a new business there. However, she does not need to incur adaptation costs again if firm i has been operating in the same location before this period, i.e., $n_{ijt} > 0$.

The right arrow on top of a variable denotes a vector. Hence $\vec{X}_{jt}$ represents attributes in location j, composed of observed characteristics $\{x_{j1t}, \ldots, x_{jkt}\}$ and unobserved characteristics ξ_{jt}. The observed characteristics involves market size, GDP per capita, profit tax rate; labor cost, investment risk, an indicator whether the location is in the Euro zone and geographical distance[8]; while unobserved characteristics ξ_{jt} contains all country-specific attributes that are difficult to measure or observe in the data but constant for all firms, like business operating costs, etc. Variable I_t is the interest rate in domestic country (Germany) to capture the common macroeconomic profit shock (or opportunity cost) for

[5] The time-invariant adaptation costs can be interpreted as the average cost over the sample period for each country.

[6] The upper bound $\bar{n}$ for n_{ijt} is 14 for 99% of firms in the data.

[7] It can also be interpreted as in the beginning of period t.

[8] The geographical distance is measured by the logarithm of the distance between the capital of each location to the capital of Germany.

all firms. ε_{ijt} is the random profit shock across firms, locations, and time, following independent identical type 1 extreme value distribution[9]. If firm i does not engage in FDI this period, i.e., $a_{it} = 0$, she only earns profit from existing subsidiaries plus the iid profit shock ε_{i0t} with the same distribution as ε_{ijt}.

Function $f_i(\vec{X}_{jt}, n_{ijt})$ captures the deterministic profit flow generated by each of firm i's subsidiaries at location j in period t, which depends on location attributes as well as firm attributes (z_{ir}).[10] Its parametric form shows how to capture firm heterogeneity in the profit function given the same location attributes. More precisely, firms attributes also consist of observed and unobserved parts, z_{i1}, z_{i2}[11] , both of which interact with location attributes. The observed characteristics is measured by firm's size (z_{i1}) in terms of total assets; while z_{i2} represents the unobserved component, e.g. the productivity of the firm. I assume that there are two types of firms, one of which is with high productivity and the other relatively less productive. The productivity follows a simple Bernoulli distribution with probability λ to be less productive.[12] Another feature of this profit function is that it includes the term n_{ijt} and its coefficient α to capture diminishing returns to total investment, because there always exists an upper bound $\bar{n}$ for every firm observed in the data. The parameter α is expected to be positive, implying that firms are less likely to set up another new affiliate in the location where they already have a large number of subsidiaries.

The expected discounted profit, i.e., the value function of firm i in period t is given by:

$$V_i(\vec{\varepsilon}_{it}, \vec{n}_{it}, \Omega_t) = \max_{a_{it} \in \{0,1,\ldots,J\}} \pi_{it} + \beta E[V_i(\vec{\varepsilon}_{it+1}, \vec{n}_{it+1}, \Omega_{t+1}) | \Omega_t, \vec{n}_{it}] \tag{3.1}$$

Where state variables in firm i's value function $V_i(\vec{\varepsilon}_{it}, \vec{n}_{it}, \Omega_t)$ are the vector of profit

[9] This random profit shock is observed by firms when they make the decision, but unobserved to researchers. In this sense, unobserved is only from researchers' perspective throughout this paper.

[10] Note that the deterministic profit flow generated by existing subsidiaries and new one both equal $f_i(\vec{X}_{jt}, n_{ijt})$, only differing in whether it involves adaptation costs or not. Thus, function $f_i(\vec{X}_{jt}, n_{ijt})$ can be regarded as the average profit flow per unit-subsidiary for parent firm.

[11] The absence of subscript t in z_{i1}, z_{i2} represents that firms' characteristics are persistent over time.

[12] In principal, I could extend the simple two types distribution to multiple types, even to continuous types distribution. However this requires to assume a known distribution and increase the computation burden exceptionally in estimation for continuum case.

shock: $\vec{\varepsilon}_{it} = \{\varepsilon_{i0t}, \ldots, \varepsilon_{iJt}\}$, a vector of the number of subsidiaries in each location: $\vec{n}_{it} = \{n_{i1t}, \ldots, n_{iJt}\}$ and the exogenous information set Ω_t including location attributes in all locations as well as firm characteristics. a_{it} is the choice variable that firm decides whether to invest or not and also which location to invest in conditional on making an investment ($a_{it} \neq 0$) after observing the location attributes and realized profit shock. Due to the arbitrary high dimensional state space,[13] it is computationally infeasible to solve the above Bellman equation. Because it needs to solve Bellman equation infinitely many times to search optimal parameters in the late estimation stage, I need to impose some additional assumptions to transform original problem into a tractable form in the next subsection.

3.4.1 Tractable Specification for Value Function

The curse of dimensionality in the discrete choice model above renders dynamic programming approach intractable. To make the original value function solvable for estimation, I first split the state variable $\vec{n}_{it}$ into two variables $\vec{s}_{it}$ and N_{it}, where $\vec{s}_{it} = \{s_{i1t}, \ldots, s_{iJt}\}$ with

$$s_{i1t} = \mathbb{1}_{(n_{ijt}>0)}, N_{it} = \sum_j n_{ijt}$$

The key information about which countries a firm has invested before this period is well preserved by the vector $\vec{s}_{it}$ and the variation in this variable helps to identify adaptation costs in the model. The state variable N_{it} keep the information of the total number of affiliates in all locations owned by firm i in period t, but loses information about the exact number of affiliates in each location. Therefore, I have to make an assumption that diminishing return occurs at the aggregate level. More precisely, firms that already have a large number of affiliates are less likely to invest in any location, independent of the distribution of $\vec{n}_{it}$. The estimated parameter α will indicate how restrictive this assumption would be. That is, if the estimation results show that α is positive and

[13]The vector profit shock $\vec{\varepsilon}_{it}$ can be analytically integrated out according to the assumption of iid type 1 extreme value distribution. However the total number of all possible combinations of variable $\vec{n}_{it}$ alone will be 14^{24}, more than one billion, let alone information set Ω_t including location attributes in all locations.

significant, then diminishing returns at the aggregate level indeed exist and effectively affect firms' investment behavior. With this assumption, the deterministic profit flow can be written as $f_i(\vec{X}_{jt}, n_{ijt}) = f_i(\vec{X}_{jt}, N_{it})$.

With respect to the profit flow from the existing subsidiaries, it can be approximated by new state variables in the following way

$$\sum_{j=1} n_{ijt} f_i(\vec{X}_{jt}, n_{ijt}) = N_{it}\psi_i(\vec{X}_t, \vec{s}_{it}, N_{it})$$
$$\psi_i(\vec{X}_t, \vec{s}_{it}, N_{it}) = \frac{\sum_{j\in\{j:s_{ijt}=1\}} f_i(\vec{X}_{jt}, N_{it})}{\sum_{j\in\{j:s_{ijt}=1\}} 1}$$

where $\psi_i(\vec{X}_t, \vec{s}_{it}, N_{it})$ represents the mean profit flow from existing subsidiaries across all countries firm i has invested before period t. The total profit flow from existing subsidiaries thus is equal to the product of N_{it} and $\psi_i(\vec{X}_t, \vec{s}_{it}, N_{it})$.

Given the new state variable $\vec{s}_{it}, N_{it}$, I define the location specific per-period profit flow π_{ijt} as follows

$$\pi_{ijt} = \bar{\pi}_{jt} - \eta_j \mathbb{1}_{(s_{ijt}=0)} + \sum_k \sum_r \tilde{\alpha}_{kr} z_{ir} x_{jkt} - \alpha N_{it} + \varepsilon_{ijt} \tag{3.2}$$

$$\bar{\pi}_{jt} = \sum_k \bar{\alpha}_k x_{jkt} + \xi_{jt} - \alpha_I I_t \tag{3.3}$$

where $\bar{\pi}_{jt}$ denotes the mean profit constant across all firms. That is to say, the profit flow (π_{ijt}) from the new subsidiary set up by firm i at location j in period t can be decomposed into two parts, one part is $\bar{\pi}_{jt}$, constant for any firm; the other part varies according to the firm's experience (s_{ijt}), characteristics (z_{ir}), the total number of subsidiaries (N_{it}) as well as the idiosyncratic profit shock ε_{ijt}. If firm i does not engage in FDI in period t, then she will gain π_{i0t}, where $\pi_{i0t} = \varepsilon_{i0t}$.

With respect to the experience vector $\vec{s}_{it}$, I go a bit further to focus on regional experience rather than country-specific experience. I group all 24 locations into 6 regions as defined above, which implies firms do not need to keep track of experience in each country, instead only on the regional level, because the previous empirical evidence establishes that the regional experience and country-specific experience have a very similar positive effect to reduce adaptation costs. Consequently, the cardinality of $\vec{s}_{it}$ is reduced from 2^{24} to 2^6. On the one hand, this grouping method reduces the state space dramatically, but one other hand it explicitly assumes that adaptation costs are constant across countries within regions and the mean profit flow can only be approximated as average profit flow of the regions where firms have invested before.

To deal with the exogenous information set Ω_t, I borrow the concept of the logit inclusive value (δ_{it}) by collapsing the high dimensional vector into a scalar.[14] Thanks to the iid type I extreme value distribution of ε_{ijt}, the inclusive value has the following closed form solution:

$$\begin{aligned}\delta_{it}(\vec{s}_{it}, N_{it}, \Omega_t) &= E\max\{\delta_{i1t} + \varepsilon_{i1t}, \dots, \delta_{iJt} + \varepsilon_{iJt}\} \\ &= \ln\big(\sum_{j=1,\dots,J} exp(\delta_{ijt}(\vec{s}_{it}, N_{it}, \Omega_t))\big)\end{aligned}$$

where

$$\delta_{ijt}(\vec{s}_{it}, N_{it}, \Omega_t) \equiv \pi_{ijt} + N_{it}\psi(\vec{s}_{it}, N_{it}, \Omega_t) - \varepsilon_{ijt} + \beta E[V_i(\vec{\varepsilon}_{it+1}, \vec{s}_{it+1}, N_{it+1}, \Omega_{t+1})|\Omega_t, \vec{s}_{it}, N_{it}]$$

δ_{ijt} denotes the choice specific value function, i.e., the value firm i can reach if she is forced to set up a new subsidiary in location j this period. Intuitively, the inclusive value $\delta_{it}(\vec{s}_{it}, N_{it}, \Omega_t)$ captures the expected value of firm's best location choice for FDI among all available locations. It provides a summary of the location attributes and selection of all possible countries to enter, taking into account adaptation costs and the infinite horizon future value. Instead of keeping track of all detailed location attributes in every country, firms could equivalently focus on the sufficient summary statistics δ_{it} under certain additional assumption. The evolution of δ_{it} can only capture the overall pattern , that is to say, all else equal, the logit inclusive value increases given market size in all countries increases. However, if e.g. the location attributes evolve differently, the change of the logit inclusive value does not provide any information about which country

[14] The detailed discussion of logit inclusive value can be found in Nevo (2006) and Gowrisankaran and Rysman (2012).

evolves better.[15] The required additional assumption is called inclusive value sufficiency (IVS) assumption:

$$P(\delta_{it+1}|\Omega_t, a_{it}) = P(\delta_{it+1}|\delta_{it}, a_{it})$$

IVS assumption implies that given the same action choice, the current inclusive value provides all relevant information about the marginal distribution of the inclusive value in the next period. In consequence, firms do not need to form expectation of each variable in information set Ω_t, but can simply focus on the scalar variable-the inclusive value instead. With this IVS assumption and all the reduction of state space, the value function can be equivalently written as:

$$V_i(\vec{\varepsilon}_{it}, \vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it}) = \max_{a_{it}\in\{0,1,\dots,J\}} \sum_j \pi_{ijt}\mathbb{1}_{(a_{it}=j)} + N_{it}\psi_{it} + \beta E[EV_i(\vec{s}_{it+1}, , N_{it+1}, \psi_{it+1}, \delta_{it+1})|\vec{s}_{it}, , N_{it}, \psi_{it}, \delta_{it}] \tag{3.4}$$

where the expected value function $EV_i(\vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it})$ is given by

$$EV_i(\vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it}) = \int V_i(\vec{\varepsilon}_{it}, \vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it})dF(\vec{\varepsilon}_{it})$$
$$= \ln\left[N_{it}\psi_{it} + \beta E[EV_i(N_{it+1}, \vec{s}_{it+1}, \psi_{it+1}, \delta_{it+1}|\delta_{it}, \psi_{it}, \vec{s}_{it}, N_{it})] + \exp(\delta_{it})\right]$$

Given the forward looking behavior, firms need to form expectations about all state variables respectively. First, the evolution of experience state $\vec{s}_{it}$ follows,

$$\vec{s}_{it+1} = \begin{cases} s_{ijt+1} = 1, s_{i-jt+1} = s_{i-jt}, & \text{if } a_{it} = j; \\ \vec{s}_{it}, & \text{if } a_{it} = 0. \end{cases}$$

If firm i chooses location j for FDI in period t, i.e., $a_{it} = j$, then $s_{ijt+1} = 1$ and all others

[15] This main criticism of using logit inclusive value is that it washes away any different evolution patterns across countries. One possible extension is to introduce two different inclusive values for instance, one for developing countries and the other one for developed countries. These two logit inclusive values at least capture different evolution patterns between developing and developed countries. However, the tradeoff is that the computation burden will increase heavily due to expanding the state space. In brief, the logit inclusive value is simply an approximation of how firms make their forecast for dynamic decisions.

stay the same.

Second, the evolution of the total number of all subsidiaries N_{it} follows:

$$N_{it+1} = N_{it} + \mathbb{1}_{(a_{it} \neq 0)}$$

Since I only focus on the location choice for FDI and abstract from the decision of setting the optimal number of subsidiaries, I assume firms always choose one unit subsidiary conditional on investing this period.[16]

Third, the evolution of the mean profit flow from existing subsidiaries ψ_{it} follows:

$$\psi_{it+1} = \gamma_{0i} + \gamma_{1i}\psi_{it} + \gamma_{2i}\sum_{j} f_i(\vec{X}_{jt}, N_{it})\mathbb{1}_{(a_{it}=j)} + \upsilon_{it} \tag{3.5}$$

where the belief shock υ_{it} is assumed to follow a normal distribution $N(0, \sigma_{1i}^2)$. The evolution of ψ_{it} measures the change in the profit portfolio from existing subsidiaries. If a firm chooses the outside option in current period, then the change in ψ_{it} only reflects the exogenous variation of location attributes in the countries where the firm has invested in before. This exogenous change in the mean profit flow from existing subsidiaries is captured by the parameter γ_{1i}. However, if a firm engages in FDI in one location with better economic development this period, the mean profit flow will be shifted upwards compared with that the firm does not make any investment or choose other locations. This difference is captured by γ_{2i}. The mean value of belief shock equals 0, implying that firms have rational expectations. All parameters in this equation vary across firms to account for firm heterogeneity.

Finally, the evolution of the inclusive value δ_{it} follows:

$$\delta_{it+1} = \begin{cases} \rho_{0i}^{out} + \rho_{1i}^{out}\delta_{it} + \nu_{it}^{out} & \text{if } a_{it} = 0; \\ \rho_{0i}^{old} + \rho_{1i}^{old}\delta_{it} + \nu_{it}^{old} & \text{if } a_{it} = j,\ s_{ijt} = 1; \\ \rho_{0i}^{new} + \rho_{1i}^{new}\delta_{it} + \nu_{it}^{new} & \text{if } a_{it} = j,\ s_{ijt} = 0; \end{cases} \tag{3.6}$$

[16] I actually observe the distribution of the number of new subsidiaries firm set up over time, on average around 62% firms choose one subsidiary. This one unit subsidiary can be interpreted as the expected number of new subsidiaries, around 1.9. In consequence, the profit flow and N_{it} are all referring to the unit subsidiary.

where I assume that the shock ν_{it} on belief also follows a normal distribution $N(0, \sigma_{2i}^2)$. I omit the superscript of ν_{it} to save notation, but the associated distribution still varies according to different types of action as in equation (3.6). The evolution of the logit inclusive value explicitly depends on firm i's current choice a_{it} as well as firms' characteristics according to equation (3.6). If firm i currently chooses the outside option, the evolution of δ_{it} only captures the information on exogenous changes, such as variation in location attributes. If firm i chooses an old location to engage in FDI this period, the evolution of δ_{it} contains information not only on exogenous changes, but also endogenous change in N_{it}, e.g. $N_{it+1} = N_{it} + 1$. Finally, if firm i currently chooses a new location j for FDI, then δ_{it+1} should reflect the reduction of adaptation cost in location j, which implies the option value in the next period should be higher compared with currently choosing old locations or outside option given everything else being equal. Therefore parameter ρ_{1i} in equation (3.6) will vary among different types of actions: choosing the outside option, choosing an old location or a new location to invest. Similarly as in equation (3.5), firms have also rational expectations regarding the evolution of the inclusive value. The associated belief parameters also differ across firms.

Let the value of choosing the outside option δ_{i0t} be:

$$\delta_{i0t} = N_{it}\psi_{it} + \beta E[EV_i(\vec{s}_{it+1}, N_{it+1}, \psi_{it+1}, \delta_{it+1})|\vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it}]$$

After integrating out the unobserved idiosyncratic profit shock, the optimal policy function P_{ijt}, the probability of firm i choosing location j to engage in FDI in period t conditional on her attributes is given by

$$\begin{aligned} P_{ijt}(\vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it}|z_{i1}, z_{i2}) = \int \mathbb{1}\big(\delta_{ijt} + \varepsilon_{ijt} > \delta_{ikt} + \varepsilon_{ikt}, \forall k \neq j, \big) dF(\vec{\varepsilon}_{it}) \\ = \frac{exp(\delta_{ijt})}{exp(\delta_{i0t}) + exp(\delta_{it})} \end{aligned}$$

With all the above reductions of the state space, I am eventually able to estimate a tractable dynamic discrete choice model. The details of estimation are presented in the next section.

3.5 Estimation and Identification

3.5.1 The Estimator

The most important parameters to be estimated are the adaptation costs η_j in each location jm and the preference coefficients α_{ik}, including the mean coefficients on the location attributes, $\bar{\alpha}_k$ as well as the random components varying with firms' attributes, $\tilde{\alpha}_{kr}$,[17] and λ governing the distribution of firms' productivity. There are also nuisance parameters, such as, γ_i and ρ_i for firms' belief on the evolution of inclusive logit value and the mean profit flow from existing subsidiaries respectively; σ_{1i} and σ_{2i} are the corresponding variance of belief shocks[18].

3.5.2 The Estimation Procedure

The estimation method in this paper closely follows Nosal (2011) and Gowrisankaran and Rysman (2012), which involves three levels of optimization. The basic idea is to nest solving a dynamic programming problem inside the location share inversion of Berry, Levinsohn and Pakes (1995). The inner loop solves firms' DP problem in equation (3.4) for each firm type and computes the predicted aggregate location share. The middle loop updates the mean profit flow $\bar{\pi}$ until predicted location shares match the observed location shares. The outer loop search over the parameter space to maximize the likelihood function.

Inner Loop

I set the annual discount factor β equal to 0.9 in order to solve the DP problem in the inner loop.[19] To obtain the fixed point for the Bellman equation (3.4), I need to discretize the continuous logit inclusive value (δ_{it}) as well as ψ_{it}. The state space dimension for variable

[17] Allowing all of location attributes to interact with firms' characteristics would provide a very flexible form of firm heterogeneity, but restricting the interaction to one location attributes helps make the estimation tractable. Thus I only interact the most important variable market size for horizontal FDI with both observed and unobserved firms' characteristics to capture firm heterogeneity.

[18] The belief shocks for the evolution of inclusive value actually vary across action types as shown in equation (3.6).

[19] The discount factor β is generally not identified in the class of dynamic discrete choice models. Thus I pick the value 0.9 that is commonly used in the literature.

δ_{it} are divided into 30 grid points and 20 grids for ψ_{it}. $N_{it} = \{0, \ldots, 9\}$[20] Analogously, observed firm's characteristics size z_{i1} is discretized into 3 types-small size, medium size and large size based on the quantile from the data.[21] Thus there are 6 types of firms in all: three observed types (size) multiplying two unobserved types (productivity). Together with state variable $\vec{s}_{it}$ and N_{it}, there are $64\times30\times20\times10$, 384000 grid points in the Bellman equation to solve for each firm type. The value function $V_i(\vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it})$ is then defined discretely on each grid point and its value is approximated by linear interpolation when the arguments fall between the grid points.

The inner loop finds the joint fixed points of several equations. It finds the value function which is the fixed point of the Bellman equation. It finds the choice specific value functions δ_{ijt},for all j and the logit inclusive value δ_{it} that satisfy their recursive definitions respectively. Finally, it finds the firm's belief parameters γ_i, ρ_i, σ_i that are stable during the iteration. To start the inner loop, some initial value guess for above variables are necessary. Since the expectation of the value function is part of the expression δ_{it} and ψ_{it}, the integration along the dimension of these state variables is achieved by simulation.[22] Once the expected value function has been computed, I can easily get δ_{ijt} for each location j and then use it to update the logit inclusive value δ_{it}. The δ_{it} are then regressed on the δ_{it-1} to obtain a new ρ_i, σ_{2i} and to regress ψ_{it} on ψ_{it-1} to get new γ_i and σ_{1i}. Since both δ_{it} and ψ_{it} are functions of endogenous state variable $(\vec{s}_{it}, N_{it})$ as well as exogenous location attributes, I can pick the realized value in the different state $(\vec{s}_{it}, N_{it})$ to nonparametrically identify these belief parameters for corresponding action choices.

After joint convergence has been achieved, I can obtain the conditional choice probabilities P_{ijt}. The P_{ijt} for all j and t, is used to predict the location share of firms choosing

[20] As discussed in the previous section, N_{it} is associated with number of unit subsidiary and the upper bound 9 covers 97% of firm observations.

[21] Regarding the measure of firms size, I use the average of parent firms' total assets during 2002-2009 to categorize them into 3 groups and assume discretized sizes are constant over the sample period in order to capture observed persistent heterogeneity. To deal with the issue that size can change into different groups, I compared the mean value of total assets with values in the first year (2002) and the last year (2009) and find that 90% of firms are located within the same group, because most of these firms are quite mature and stable in terms of scale and additionally, the variation of N_{it} on the subsidiary level compensates the restrictive grouping method.

[22] The integration over state variable $\vec{\varepsilon}_{it}$ is relatively easy due to its iid type 1 extreme value distribution, leading to a closed form solution of its expectation. However, I have to randomly draw the orthogonal belief shocks ν_{it} and υ_{it} to obtain the integration over state variable δ_{it} and ψ_{it} respectively.

to enter in each country and then pass it into the middle loop estimation.

Middle Loop

The middle loop is an application of the Berry, Levinsohn and Pakes (1995) inversion. They have proved that there is a one-to-one mapping between the average profit flow $\bar{\pi}_{jt}$ and location shares χ_{jt}. For the ease of computation, I divide the main parameters into two mutually exclusive sets of parameters, linear parameters, $\Theta_1 = \{\bar{\alpha}_1, \ldots, \bar{\alpha}_k\}$ and nonlinear parameters, $\Theta_2 = \{\alpha, \tilde{\alpha}_1, \tilde{\alpha}_2, \lambda, \eta_j, j = 1, \ldots, 24\}$; let

$$\chi_{jt} = \hat{\chi}_{jt}(\bar{\pi}_{jt}, \forall j|\Theta), \forall j$$

where χ_{jt} denotes the predicted share of firm choosing location j in period t for FDI. It is is a function of the average profit flow $\bar{\pi}_{jt}$in all locations as well as parameter Θ, which is passed in from the outer loop. To solve the above system of equations, Berry, Levinsohn and Pakes (1995) provide a computational device to aid in concentrating out the $\bar{\pi}_{jt}$.

$$\bar{\pi}^{'}_{jt} = \bar{\pi}_{jt} + \ln \chi_{jt} - \ln \hat{\chi}_{jt}(\bar{\pi}_{jt}, \forall j|\Theta), \forall j$$

where $\bar{\pi}^{'}_{jt}$ is the updated average profit flow which is guaranteed to converge due to the contraction mapping.[23] Given the new average profit flow, I then update the predicted location shares via the inner loop. This means the convergence of middle loop is actually joint convergence of middle and inner loop. Gowrisankaran and Rysman (2012) suggest to iterate inner and middle loop interchangeably until convergence in both stages to save computation time. Once the average profit flow has converged, the linear parameters Θ_1 can be represented as a function of the nonlinear parameters Θ_2.

[23] In the static case, it is proved to be a contraction mapping. However, it is not necessarily the case in our dynamic setting, but the convergence to one of multiple fixed points is not a problem.

Outer Loop

The outer loop search over the set of nonlinear parameters (Θ_2) to maximize the likelihood function:

$$\hat{\Theta}_2 = arg\max_{\theta_2}\{\sum_i \sum_t \sum_j 1_{(a_{it}=j)} \log(\hat{P}_{ijt}(\vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it}))\} \tag{3.7}$$

When convergence is reached in both middle and inner loop given Θ_2, I can obtain the predicted probability $\hat{P}_{ijt}(\vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it})$ after integrating out unobserved firm attributes z_{i2}. To construct the objective function, I take every firm i's state variable in the first period of the sample as exogenous given. This expression might generate the typical "initial condition problem", because firms' states in the initial period of the sample are induced by past investment decisions out of the sample. There are several ways to deal with "initial condition problem" in the literature. In principle, the initial state of every firm can be treated as the steady state derived from optimal policy function. Alternatively, the probability of observing the initial state from the data can be approximated by a reduced-form nonlinear model as Heckman (1981) suggested.[24]

During the optimization, the predicted conditional choice probability $\hat{P}_{ijt}$ needs to be computed at any given parameters vectors. The algorithm will terminate when the outer loop reaches the maximum value and inner and middle loop jointly converge at the same time.[25]

3.5.3 Identification

Given the exogenous discount factor and the parametric form of profit function, different sources of variation in the data help to identify different sets of parameters in the dynamic discrete choice model. The key to identify the adaptation cost η_j is the FDI made by firms entering the country for the first time, because only these firms have to incur the adaptation costs. Given everything else being equal, I should observe that the share of new entrants is monotonically decreasing with adaptation costs. That implies one country

[24] In the current version of this paper, I unfortunately do not address the "initial condition problem".

[25] I use the Nelder-Mead non-derivative "simplex" search method to get the relatively more robust results for nonlinear parameters.

with high adaptation costs is less attractive to new entrants than another country with low adaptation costs; while the decision made by existing entrants that have prior experience in both locations should be independent of the costs. Moreover, the share of experienced firms choosing to invest in existing locations should be larger than the share of new entrants as long as the reduction in adaptation cost can provide a high compensation for the negative effect of diminishing return.

As in Berry, Levinsohn and Pakes (1995, 2004), I identify the mean coefficient on the location attributes, $\bar{\alpha}_k$ as well as the random component varying with firms' attributes, $\tilde{\alpha}_{kr}$, $r = 1, 2$ and λ by combining the micro and macro moment conditions from the share of firms that have invested more than once in the same location.[26] To be more precisely in this paper, the coefficient ($\tilde{\alpha}_{k1}$) for the interaction between firm size and market size can be identified by the variation of locations chosen by the firms belonging to different size group. Regarding the coefficient ($\tilde{\alpha}_{k2}$) for the interaction between unobserved firm attributes and location attributes, it can be identified by keeping track of the same firms' location choices over time. As unobserved firm types follow a Bernoulli distribution, the relatively more productive firm can make better use of market size through that parameter. Therefore, the variation in locations chosen by the same firm helps to identify the magnitude of that parameter. Suppose the mean coefficients are simply zero, then less productive firms with equal probability choose each location, however productive firm will always choose the location with largest market size over time. Given the identification of $\tilde{\alpha}_{k2}$, the variation in the aggregated share of firms entering each location over time helps to identify the distribution (λ) of these two different types of firms. However if $\tilde{\alpha}_{k2}$ is close to 0, then the two type distribution will be degenerated into one type, i.e., no unobserved firm heterogeneity.

[26]Since the two unobserved types of firms in this paper is a special case of general random coefficients i n discrete choice model, the identification argument form Berry, Levinsohn and Pakes (1995, 2004) also works here.

3.6 Results and Counterfactuals

Since the discount factor β is not identified, I present estimation results for both a dynamic model with $\beta = 0.9$ and a myopic model with $\beta = 0$.[27] The results of counterfactual analysis below are based on the dynamic model because forward looking behavior is the fundamental assumption in this model.

3.6.1 Estimated Parameters

Table 3.2: Estimation results

Θ_1	Model		Θ_2	Model	
	myopic($\beta = 0$)	dynamic($\beta = 0.9$)		myopic($\beta = 0$)	dynamic($\beta = 0.9$)
gdp	2.10	2.61	η_{WE}	1.16	-0.01
	(0.0215)	(0.0765)		(0.1922)	(0.0089)
gdppc	0.24	0.46	η_{SE}	1.58	0.96
	(0.0085)	(0.0083)		(0.3449)	(0.1791)
eruo zone	0.91	0.32	η_{NE}	2.10	0.07
	(0.0144)	(0.0480)		(0.4310)	(0.1446)
CI	0.58	-0.60	η_{EE}	1.83	1.39
	(0.0901)	(0.1188)		(0.2371)	(0.0600)
labor cost	0.33	-0.39	η_{NA}	1.95	0.11
	(0.0155)	(0.0228)		(0.3185)	(0.0790)
unemployment	-2.14	0.07	η_{EA}	3.23	4.12
	(0.6721)	(0.0499)		(1.4068)	(0.1695)
tax rate	-4.88	-7.36	α	-0.14	2.47
	(0.5712)	(0.4415)		(0.0492)	(0.1097)
interest rate	-14.85	-17.48	$\tilde{\alpha}_1$	0.005	0.0005
	(0.1483)	(0.5281)		(0.0374)	(0.050)
distance	-7.23	-2.80	$\tilde{\alpha}_2$	1.49	2.64
	(0.0229)	(0.0688)		(0.6182)	(0.1086)
			λ	0.64	0.91
				(0.4614)	(0.1003)

Notes: CI denotes corruption perceived index; Standard errors are reported in the parentheses.

The coefficients (Θ_1) on the location attributes in both myopic and dynamic models have a mix of expected and unexpected signs in Table (3.2). The coefficient on market size (gdp) is significantly positive as expected and consistent with the incentives of horizontal FDI to serve a local market. Tax rate (tax) on profit, GDP per capita (gdppc), the domestic interest rate, the distance and the Euro zone dummy have expected sign in both models. The significantly negative coefficient of the interest rate in Germany indicates

[27] As $\beta = 0$ indicates, myopic agents only care about current payoff and put zero weight on future payoff in their lift-time payoff.

a high opportunity costs for German firms investing abroad. The unexpected positive coefficient on labor cost in the myopic model can have two possible reasons. One possibility is that labor cost is highly correlated with GDP per capita (gdppc), which reflects the endowment similarity between host countries and Germany. Since Germany is a highly developed country, horizontal FDI made by German firms is more likely to be attracted to countries with similar economic development. Another possibility is that the variable labor cost is measured by hourly compensation costs in manufacturing industry which does not really capture the labor cost in Non-manufacturing industry. The coefficient of CI in dynamic model is unexpectedly negative because investment risk is detrimental to FDI. These unexpected sign in both models requires more robustness checks.[28]

The important parameters Θ_2 include adaptation costs in each region as well as random coefficients. All adaptation costs except η_{WE} in both models have the correct sign and are statistically significant.[29] Moreover, the magnitude and ordering are also similar in both models. Adaptation costs are the lowest in Western Europe η_{WE}; while Eastern Asia η_{EA} has the highest adaptation costs for German firms. Adaptation costs in Eastern Europe η_{EE} is the second highest, but with a relatively small scale, only one third of the highest costs. North America η_{NA} and Northern Europe η_{NE} are relatively close to each other from German firms' perspective. Regarding the ordering of Adaptation costs except the highest and lowest, myopic and dynamic model present us a mixed picture. Together with the negative adaptation cost η_{WE} in dynamic model, an extensive robustness check needs to be done in the future research. Since these estimated values themselves do not provide us with any information about how large they are economically, I am going to quantify the magnitude of these adaptation costs in terms of firms' expected discounted profit. I find the these adaptation costs ranging from 0.5% to 29% of German firms' expected discounted profit on average, defined as the average across all firms over the whole sample

[28]Please recall that due to intensive computational burden, I don't directly search the parameters in Θ_1 via MLE; instead I use recovered $\bar{\pi}_{jt}$ in the middle loop optimization to estimate Θ_1. It is obvious that the total number of observations of $\bar{\pi}_{jt}$ is the product between the number of locations and the length of periods. Since the choice set is fixed, i.e., 24 locations, only the time length can add identification power.

[29]Recall that country-specific adaptation costs within the same region are assumed to be the same for the ease of computation.

period.[30]

With respect to the parameter α, the significantly positive sign in the dynamic model indicates that there exists diminishing returns on FDI at the aggregate level. That is to say, firms with a large number of affiliates abroad are less likely to set up one more new subsidiary in any country due to this diminishing marginal return which decreases the option value generated by previous operating experience in some countries.[31] Coefficient $\tilde{\alpha}_1$ is positive but insignificant, implying that firms with different size do not have strong heterogeneous preference towards different market size. In other words, there are no obvious differences between the profits earned by firms of different capital size. However, $\tilde{\alpha}_2$ is significantly positive as expected. This means that even if all these firms engage in FDI, which in general are regarded the most productive firms compared with those only serve domestic markets, there is still some heterogeneity in productivity among these multinational firms. This result is in line with the new findings in recent papers by Yeaple (2009) and Chen and Moore (2010) that firm heterogeneity affects the structure of MNC activity. Parameter λ shows that only 36% and 9% firms are relatively more productive than others in the myopic and the dynamic models respectively.

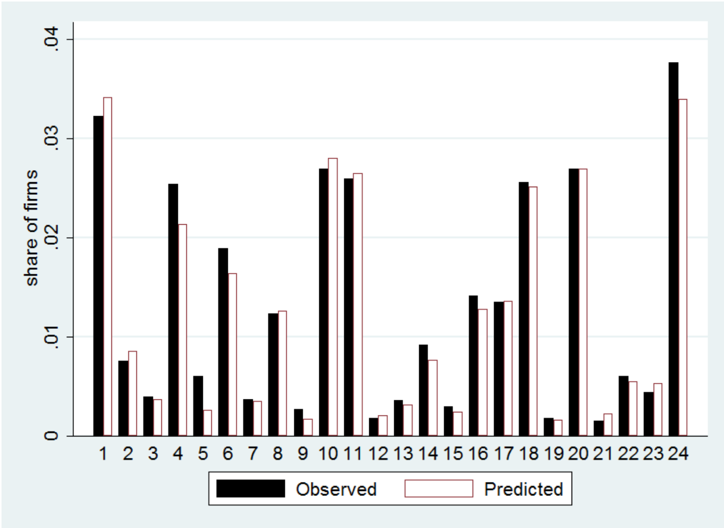
3.6.2 Model Fit

Before conducting any counterfactual analysis, I present how these estimated parameters fit the data. I first randomly draw iid profit shocks (ε_{ijt}) from type 1 extreme value distribution, and then given the estimated parameters from the dynamic model together with exogenous location attributes, firms need to re-optimize each period, i.e., choosing the best location to engage in FDI to maximize expected discounted profit. After observing new FDI choice made by every firm, I can compute the percentage of firms engaging in FDI in each location. By integrating these predicted probabilities over all firms and sample periods, I get the predicted share of each location as in figure (3.2). In general, the estimated parameters fit the data quite well for all 24 locations. Thus, I will use the

[30] Due to negative value of the lowest adaptation cost, I use the second lowest adaptation cost as a lower bound for the range, and the upper bound is from the highest adaptation cost in Eastern Asia.

[31] The significantly negative sign of α in the myopic model also needs more robustness checks.

Figure 3.2: Model fit



same estimated parameters as well as the same draw of random shock for counterfactual analysis in the next subsection.

3.6.3 Counterfactual Analysis

Table 3.3: National regulatory changes 2002-2009

Item	2002	2003	2004	2005	2006	2007	2008	2009
Number of countries that introduced changes	43	59	80	77	74	49	41	45
Number of regulatory changes	94	126	166	145	132	80	69	89
Liberalization/promotion	79	114	144	119	107	59	51	61
Regulation/restriction	12	12	20	25	25	19	16	24

Source: UNCTAD, Investment Policy Monitor database

Table 3.3 suggests which policy experiments might be useful for consideration. It is clear that the overall policy trend favors continuous liberalization and promotion of foreign investment. Thus I will simulate FDI promotion policy, which will be associated with the reduction of adaptation costs in my model. However the ways to reduce different components of adaptation cost are not unified. For components involving business environment and government regulations, host country could adopt international standard business practice such as use English as common working language and meanwhile reduce regulation to facilitate foreign investment. Other components like social culture and legal system are quite stable and very difficult to change; however, both home and host countries could provide various subsidies to help firms enter different local markets.

In brief, I compute two counterfactuals. In the first one, adaptation costs are reduced by the same proportion of the original scale in all countries. In the second one, adaptation costs are cut by the same amount in all countries. In each counterfactual, firms are allowed to re-optimize in reaction to the exogenous change in adaptation costs. Thus firms' location choices for FDI could change in a given period. Additionally, firms also have different beliefs about the evolution of state variables, leading to a different solution to the Bellman equation in the counterfactuals. There are two reasons why we should be interested in these different scenarios. Firstly, it should not be surprising that countries with the highest adaptation costs would benefit the most in terms of new market share by scaling down the same proportion simply because of its largest magnitude of reduction. Admittedly, it is true given everything else, like location attributes, being equal, but then I can compare the second scenario to see whether it is really the case if adaptation costs in all countries could be cut by the same amount. Secondly, two distinguished groups of firms defined as whether they engage in FDI for the first time or not, will make systematically different responses under above two counterfactuals. Under the second scenario, firms that engage in FDI for the first time will not change their location choices because this policy just introduces a constant shift for all alternatives, which would not affect firms' decision in a standard discrete choice model. However, it does affect the investment behavior for firms that already have operating experience because it changes the relative benefit of experience from specific countries. In the fist counterfactual, firms from both groups could both change location choices due to re-optimize in reaction to this policy change. Regarding the timing of FDI, both counterfactuals make investment relatively more attractive in contrast to not investing. To sum up, both counterfactuals affect firms' decision when to invest, but induce different behaviors with respect to where to invest for different groups of firms.

I use the parameters estimated from the dynamic model as true parameters to derive the benchmark expected maximum profit (henceforth EMP) for every firm based on each individual observation in the data. For each counterfactual, I then compute the EMP change that is essentially a compensating variation that would induce an equivalent change in the EMP, taking into account firms' re-optimizing investment behavior. McFadden (1999) outlines the methodology in standard discrete choice models, but the random coefficients and dynamic part needs one more step of calculation. To deal with random coefficients, I first calculate the compensating variation separately by each firm type then integrating them over the distribution of types. The dynamics are accounted for by including the continuation value in the profit for every firm. In addition to the change in EMP, the actual and counterfactual shares of firms entering each location are reported.

Counterfactual (1): scale down adaptation costs

In this counterfactual, adaptation costs in every location are permanently set to half of the original scale and zero, respectively, i.e., $\eta_j^{new} = \frac{1}{2}\eta_j$ and $\eta_j^{new} = 0$, for all j. Even though no policy would actually eliminate adaptation cost completely to zero, there are some ways like subsidies from both host and home country could reduce adaptation cost substantially. In every year during the sample period, firms face the new zero adaptation costs in every location, and they decide whether or not to engage in FDI and then choose a location conditional on investing to maximize expected discounted profit.

Figure 3.3: Change of share in regions if scaling down adaptation costs

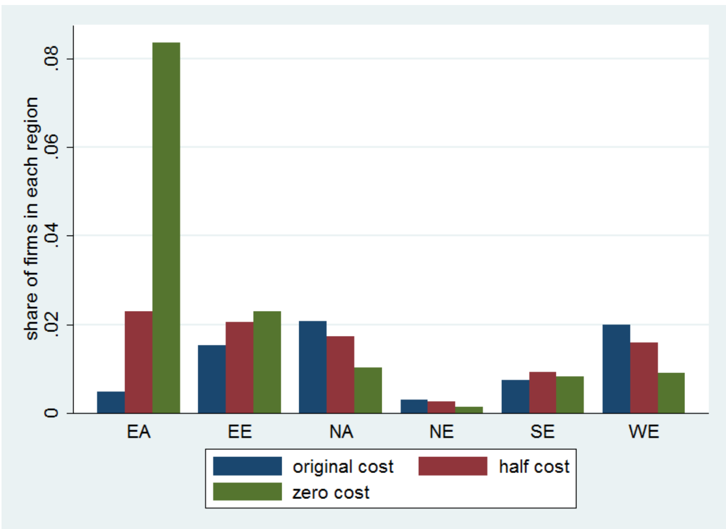
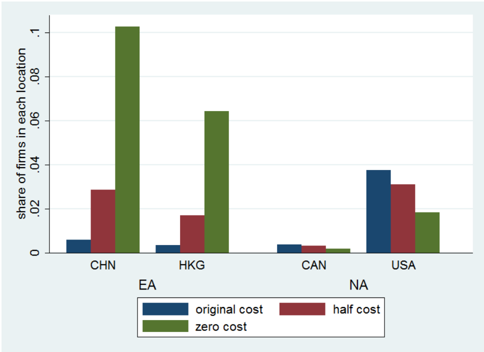
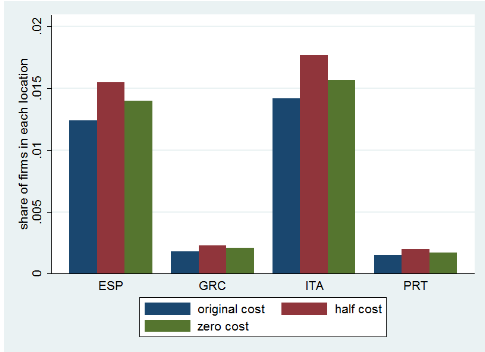


Figure 3.4: Change of share in countries if scaling down adaptation costs

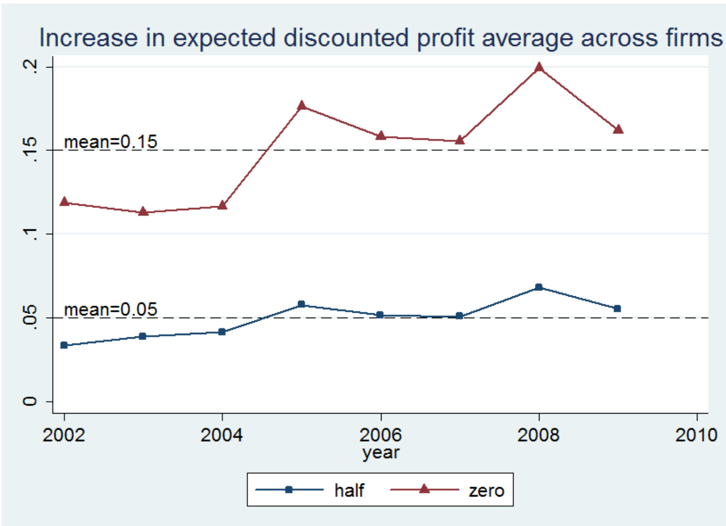


(a) China and North America



(b) Southern Europe

Figure 3.5: Change of firms' profit if scaling down adaptation costs

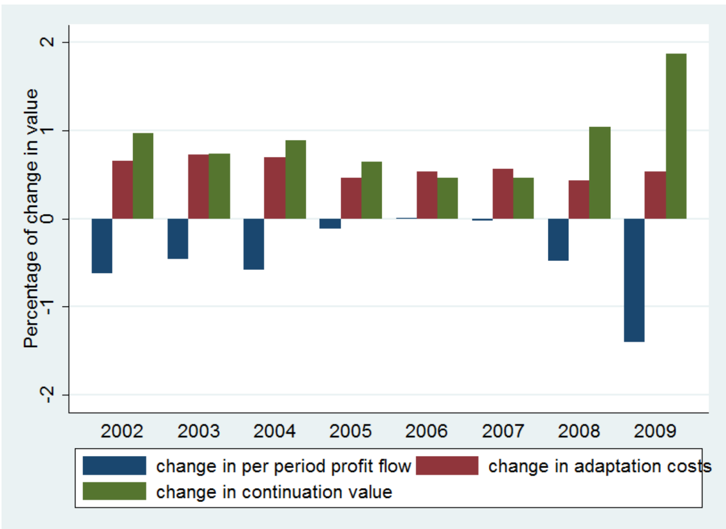


The actual and counterfactual aggregated share of firms investing in each region and country are reported in figure (3.3) and (3.4) respectively. Countries in Eastern Asia (EA), Eastern and Southern Europe (EE,SE) are more attractive compared with countries in North America (NA), Northern and Western Europe (NE,WE) if adaptation costs in all countries are scaled down to half or zero. For example, in zero adaptation cost scenario, the share of firms entering Eastern Asia region multiplies 17 times on average over the sample period. As expected, Eastern Asia benefits the most from complete elimination of the largest adaptation costs there. However, in addition to the reduction of adaptation costs per se, it is the new matching process that contributes to that highest share. Under the counterfactuals with zero adaptation cost everywhere, I observe around 22% firms change the location choice every period, which implies most of the variation in the counterfactual share comes from the new firm-location matching. What drives the matching patterns in this counterfactual entirely depends on location attributes. Since China has a huge market and its economic future is bright, it attracts most firms investing there, outweighing the USA and becomes the top 1 location for FDI in the absence of adaptation costs. On the contrary, the share of countries in North America falls by almost half with zero adaptation cost because more attractive countries like China may provide a better match for those firms originally investing in this region.

Figure (3.4b) shows that there are still some variations in country performance within the same region. In comparison with Greece and Portugal, Spain and Italy have a relative better economic situation, therefore would attract more firms entering their markets. This figure again reinforces the role of matching channels in the counterfactuals. Concerning the significant drop of share in Western Europe (WE), it is due to the negative adaptation costs. This means the adaptation costs for all countries in that region are actually increasing if I scale down all adaptation costs to half or zero. Consequently, I see a sharp drop in the share of firms entering that region.

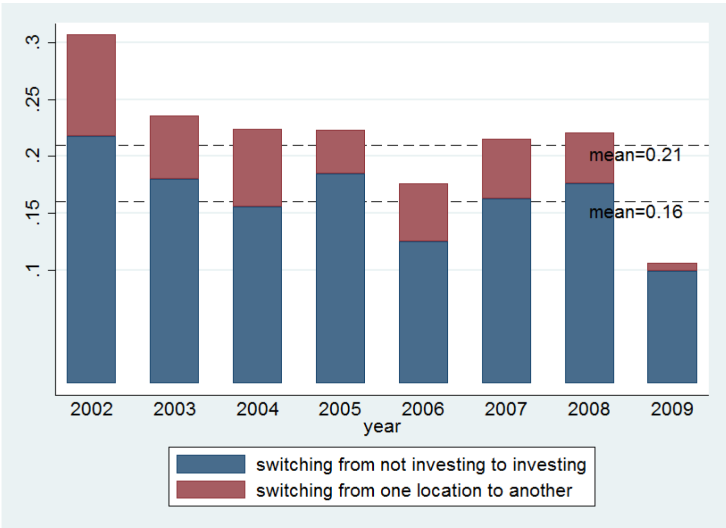
With respect to the change of profits in this counterfactual, figure (3.5) shows that the increase of firms' expected discounted profits is growing over time. This increment in profit is also contributed by the same two sources; one is from the reduction of adaptation costs per se; the other one is even more important and comes from the new matching between location and firms. Moreover, as market size exogenously growing over time, the gap between the profit increase is also expanding under two different levels of reduction in adaptation costs. During the sample period, the increase of expected discounted profits average across all firms is around 5% on average over time if adaptation costs in all countries are scaled down to half; while the increase is around 15% if all adaptation costs are completely eliminated. Figure (3.6) shows the decomposition of the changes in firms' expected discounted profit in the counterfactual with zero adaptation cost. In the absence of adaptation costs, the first period's profit flow on average across firms decreases in most periods after switching to a new location, but the total profit still increases due to the large compensation generated by the continuation value from switching as well as the elimination of adaptation costs. On average over time, the elimination of adaptation costs contributes around 57% of the increase in expected discounted profit, while an even larger contribution is from increment in the continuation value, around 88%.

Figure 3.6: Decomposition of firms' profit change with zero adaptation cost



Finally, Figure (3.7) shows how firms would switch locations when adaptation costs were subsidized to be zero in every location. The general switching pattern is that on average 21% of firms in each period would change their original location choice in the absence of adaptation costs. This average changing behavior can be further decomposed into two types of location change, one of which is switching from not investing to investing; the other one of which is switching from one location to another. As in Figure (3.7), most of the switching patterns are driven by the first type that on average 16% of firms in each period would be motivated to start to invest again if adaptation costs were completely subsidized. Combined with Figure (3.6), the profit flow would be likely to be negative in the period when firms switch from not investing to investing, but the future profit (continuation value) is large enough to compensate the one period's loss, which implies a consistent forward-looking behavior in the dynamic process of firms' FDI location choice.

Figure 3.7: Location switching patterns with zero adaptation cost



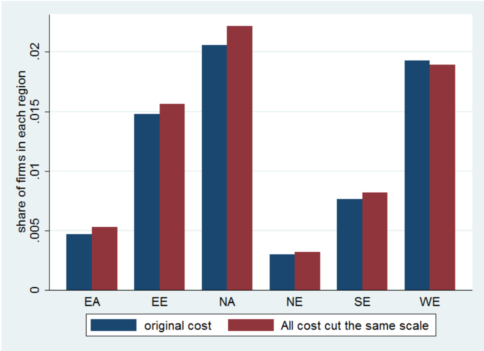
Counterfactual (2): all adaptation costs cut by the same amount

In this counterfactual, adaptation costs in every location are all cut by 0.07, i.e., $\eta_j^{new} = \eta_j - 0.07$, for all j.[32] In this setting, it is equivalent to normalize the adaptation costs in Northern Europe to be 0, that is $\eta_{NE}^{new} = 0$.

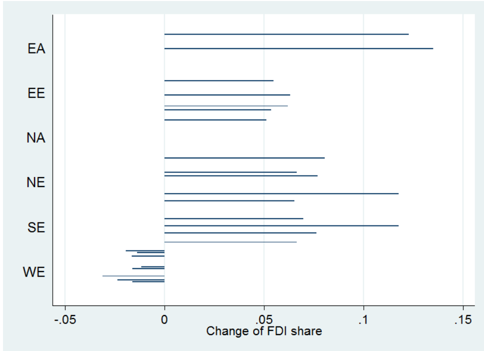
Figure (3.8) shows how firms respond to this policy scenario and the predicted aggregate share of firms choosing each region and country respectively. In this scenario, all other regions are attracting firms that originally invested in Western Europe (WE)

[32] Adaptation costs in Northern Europe represent the lowest bound, which is equal to 0.07. I use this lowest bound as a shift to adaptation costs in all countries.

Figure 3.8: Counterfactual: all adaptation costs are cut by the same amount



(a) Region



(b) Country

to their local markets, and additionally attract firms that originally did not choose to engage in FDI. Firms, fleeing from WE, can find a better location for FDI to maximize the expected discounted profits there, which indicates that countries within WE are the least attractive. The right sub-figure in figure (3.8) presents the growth rate of aggregate shares of firms investing in each location if adaptation costs everywhere are cut by the same amount. Not surprisingly, every country in WE has a negative growth rate in its share because of the available better location matches outside WE. Again countries within Eastern Asia (EA) enjoy the highest growth rate in its share, even though the level of its share is still second lowest. This observation on one hand reflects the attractive market potential in EA, on the other hand, it also indicates that location attributes there are still not good enough to compensate its highest adaptation costs.

All the switching between locations are generated by firms with certain experience in some locations. The reduction of adaptation costs by the same amount makes the benefit of state dependence shrink. In a sense, it encourages firms to invest in a broader range of countries than before. In addition, due to a better match from a larger set of host countries. The increase of expected discounted profits on average across all firms and over time is around 7% in this counterfactual.

3.7 Conclusion

In this paper, I develop a dynamic structural model that characterizes firms' decision of when and where to engage in FDI. It embodies uncertainty, observed and unobserved firm heterogeneity in the profit function of FDI, and adaptation costs for firms breaking into a new foreign market. Based on plant-level data on FDI made by German non-manufacturing firms, I estimate the model using a newly developed methodology from empirical industrial organization. After recovering adaptation costs in every location from firms' profit function, I use them to conduct counterfactual FDI promotion policy analysis.

The main focus throughout this paper is on adaptation costs, which strongly affect firms' sequential FDI location choice pattern. The estimation results suggest that these adaptation costs are substantial, varying across locations and ranging from $0.5\% - 29\%$ of the expected discounted profit on average across all German firms. Consequently, firms do not engage in FDI in new locations unless the expected discounted profit there is large enough to compensate adaptation costs. They also tend to invest even when current net profit is negative, thus avoiding the adaptation costs of starting the new business in foreign markets when economic condition improves. In this sense, history and expectations are very important for firms to engage in FDI sequentially.

The policy experiments of FDI promotion suggest that reduction of adaptation costs contributes the increase of firms' profits through two channels. The first channel is the exact amount of increase directly from the cost reduction; while more importantly, the second channel provides a better matching process between firms and location, which generates most of the increment in firms' profits. Moreover, firms with distinct experience state will respond differently to different FDI promotion policies.

3.8 Appendix

3.8.1 Fixed Effect

To control for the unobserved firm-country fixed effect F_{ij}, I apply the conditional likelihood approach developed by Chamberlain (1980). Chamberlain has proved that the sum of decisions over periods is a sufficient statistic for the individual fixed effect in the binary choice model. It is an easy extension to show that $\sum_t d_{ijt}$ for $j = 1, \ldots, J$ is a sufficient statistics for F_{ij} in the multiple choice model, where $d_{ijt} = 1$ is equivalent to $a_{it} = j$, denoting firm i choose location j in period t.

Proof. To prove that $\sum_t d_{ijt}$ is is a sufficient statistics for F_{ij} for $j = 1, \ldots, J$. I can write the logit model in terms of these parameters alone without loss of generality:

$$P(d_{ijt} = 1) = \frac{exp(F_{ij})}{1 + \sum_k exp(F_{ik})}$$

Then the log-likelihood function is given by

$$\begin{aligned} L &= \sum_i \sum_t \sum_j d_{ijt} \ln P(d_{ijt}) \\ &= \sum_i \sum_j \sum_t d_{ijt}[F_{ij} - \ln(1 + \sum_k exp(F_{ik}))] \\ &= \sum_i \sum_j F_{ij}(\sum_t d_{ijt}) - \sum_i \sum_j \ln(1 + \sum_k exp(F_{ik}))(\sum_t d_{ijt}) \end{aligned}$$

Notice that this likelihood function is now in a form that depends only on the parameter F_{ij} and the statistic $\sum_t d_{ijt}$. Thus, by the factorization theorem, $\sum_t d_{ijt}$ is is a sufficient statistic for family of probability distribution in F_{ij} for $j = 1, \ldots, J$. □

Hence, I can construct the conditional likelihood function that does not depend on the incidental parameters F_{ij}. Moreover, the conditional maximum likelihood estimator has been proved to be consistent given some mild restrictions on the F_{ij}. To illustrate Chamberlain's technique in my model I begin with the simplest case $T = 2$, i.e. two periods observation indexed by t_1 and t_2.

$$P[d_{it_1} = (1,0,\dots,0), d_{it_2} = (1,0,\dots,0)|X, \sum_t d_{it} = (2,0,\dots,0)] = 1,$$

$$P[d_{it_1} = (1,0,\dots,0), d_{it_2} = (0,1,\dots,0)|X, \sum_t d_{it} = (1,1,\dots,0)]$$
$$= \frac{P[d_{it_1} = (1,0,\dots,0), d_{it_2} = (0,1,\dots,0)|X]}{P[d_{it_1} = (1,0,\dots,0), d_{it_2} = (0,1,\dots,0)|X] + P[d_{it_1} = (0,1,\dots,0), d_{it_2} = (1,0,\dots,0)|X]}$$
$$= \frac{exp(\alpha X_{1t_1} + \alpha X_{2t_2})}{exp(\alpha X_{1t_1} + \alpha X_{2t_2}) + exp(\alpha X_{1t_2} + \alpha X_{2t_1})}$$

where $d_{it_1} = (d_{i1t_1}, \dots, d_{iJt_1})$, with $d_{ijt_1} = \{0,1\}$,

$$P[d_{it_1} = (1,0,\dots,0), d_{it_2} = (0,1,\dots,0)|X] = \frac{exp(F_{i1} + \alpha X_{1t_1})}{1 + \sum_j exp(F_{ij} + \alpha X_{jt_1})} \frac{exp(F_{i2} + \alpha X_{1t_2})}{1 + \sum_j exp(F_{ij} + \alpha X_{jt_2})}$$

When conditioning on $\sum_t d_{it}$, the sequences $\sum_t d_{it} = (2,0,\dots,0)$ does not contribute to the likelihood value because it is independent of the all the parameters. The advantage of this conditional probability is that it cancels out incidental parameters F_{ij}.[33] Then I can obtain conditional log-likelihood function:

$$L = \sum_i \ln[\frac{exp(\alpha \sum_t \sum_j X_{jt} d_{ijt})}{\sum_{D \in B_i} exp(\alpha \sum_t \sum_j X_{jt} D_t)}]$$

Where

$$B_i = \{\mathbf{D} = (D_1, \dots, D_T)|D_t = (d_{i1t}, \dots, d_{iJt}), \sum_t D_t = \sum_t d_{it}\}$$

B_i is the set that contains all the permutations generating the same investment sequences as sufficient statistics $\sum_t d_{it}$, and it varies across different firm i. This conditional likelihood function shows a transformation of the original data that preserves the information necessary to estimate α. Since L in the above expression is a normal logit likelihood function without any incidental parameters, standard maximum likelihood estimation programs will yield consistent estimators.

[33] Actually, it drops out all the time invariant terms, i.e. controlling for all the observed and unobserved fixed effect.

Chapter 4

Joint Audit and Audit Market Competition

4.1 Introduction

Audit market concentration has been a serious concern for the regulators and many market participants. It is a well established fact that the Big Four auditors(Ernst & Young, KPMG, PricewaterhouseCoopers and Deloitte) provide services to the bulk of public listed firms in most major economies (Francis et al., 2012; Sikka 2013). For instance, in the UK market, 90% of FTSE 350 index firms are audited by the Big Four and in the US they collect more than 90% of total audit fees.[1] Moreover, even for the small cap companies, it seems that the Big Four are gaining their business (Morningstar Professional Services Rankings Guide, 2012). As indicated by the Morningstar research report, Deloitte and Ernst & Young are the top two earners in terms of total profit. The regulators concern that the high concentration has created high entry barriers for the audit service market and companies purely favor the Big Four because of their dominance (Government Accountability Office, 2008; European Commission, 2010). They also concern that the lack of choice in the audit market, especially for financial service industry, might have

[1]FTSE 350 is a share index of the 350 companies listed on the London Stock Exchange with the highest market capitalization.

serious ramifications if one of the Big Four fail like Arthur Andersen did.[2] After the financial crisis, these Big Four audit firms came under scrutiny after giving banks a clean bill of health just before they were rescued by taxpayers in the financial crisis (Jones, 2013). Both the EU and the US regulators are discussing possible regulatory changes. When the Public Company Accounting Oversight Board (PCAOB) in the US is still exploring the possibility of introducing audit reforms, the European parliament is acting more swiftly. Several proposals have been made by Michel Barnier (the EU internal market commissioner) and the parliament members are currently drafting the proposed law. These efforts made by regulators suggest that they are rather determined to improve competitiveness of the audit market.

Among some major reform proposals put forward by the EU regulators, joint audit policy has been more controversial in the policy debate (Jones, 2013). Since the European Commission green paper (2010) expresses the concern over audit market concentration, the mandatory joint audit was first seriously considered by the European Commission (henceforth EC) in 2010, then in the early 2011 proposal, the joint audit policy become only "encouraged" due to the concern for seeking enough support in the EU states.[3] However, later the European Union lawmakers beef up the reform by adding the joint audit to a draft EU law designed to improve the performance of audit firms. The most recent development indicates that the lawmakers again drop the initial proposal when facing unpopular support from corporate clients and the Big Four (Baker and Jones, 2011; CFO UK, 2012; Jones, 2012; Jones, 2013). Apparently the EU regulators have gone back and forth on this particular policy debate. This implies that joint audit is quite controversial since this proposed reform is obviously welcomed by the mid-tier firms and the Big-four quickly lobby against such policy (UK parliament report, 2011; Jones, 2012). And the existing evidences are not conclusive enough to support regulators to push the reform forward.

Hence our paper intends to provide a thorough investigation on how joint audit policy

[2] This might be possible since some of the Big Four have been investigated by US regulators for criminal wrongdoings either in the US or abroad

[3] Some European states allow voluntary joint audits, for example Sweden.

would affect the audit market. In particular, we focus on the public quoted firms in the UK. The main reason we look into the UK audit market is because of its importance in the European setting. As for now the EU lawmaker is still waiting for the outcome of UK inquiry into competition in the audit market. And the pending views from the Competition Commission would significantly affect the reform of the audit market in the EU (Crump, 2012). Moreover, the current draft law on audit market reform is sponsored by one British conservative member in the European parliament. In his own words, "The views presented by the Competition Commission will be one of a number of factors considered when designing the future of the audit market in the EU," said Sajjad Karim, the British lawmaker who is leading the reform.

Joint audit policy was adopted or is still in use for certain EU state. Denmark had mandatory joint audit until 2005, then the requirement was abolished. In fact, France is the only EU country that currently implements this policy for the listed companies. France and Denmark are reported to have the least concentrated audit markets in Europe (London Economics, 2006). So it is the EC's intention to use such policy to shape up the audit market competition and allow the medium or small size audit firms to participate in large audits (EC 2010). As mentioned above, such policy is not welcomed by all parties. The advocates of joint audit argue that the potential benefits could include: less concentrated market, the audited evidences are better assured by two professional firms, and the audited report has to be co-signed by both firms, then it is less likely for both firms to collude with the client. However joint audits may suffer from a potential free rider problem (Deng et al., 2012) and it is possible that there could be chance of miscommunication between two firms [4]. The potential rising audit fee is the other argument brought against the joint audit.

What actual effects this policy reform might bring are still unclear. The up to date empirical research provides mixed evidences on the impact of joint audit on audit fees and audit quality (Francis et al., 2009; Andre et al., 2012; Ratzinger et al., 2013). They find limited support to suggest that joint audits lead to increased audit quality, but some

[4]There are concern raised by audit committee chairs in the UK (Jones, 2012)

support to suggest that joint audits lead to additional costs. However, these studies mainly focus on the correlations between audit fees (or audit quality) and firm characteristics.

Our research goes beyond and addresses these issues with a demand and supply market framework. We model client firms' preference for audit service based on the clients' own attributes as well as the auditors' attributes. So our estimation quantifies clients' heterogeneous preference over each individual audit firm. Moreover, since most of the existing empirical researches are conducted on the French or Danish cases, these current evidences could not answer the question "how would the joint audit shape up an audit market like UK which doesn't have this policy". Thus our research intends to contribute to such question and additionally address it especially from a social welfare perspective.

In order to answer the question how such possible regulation would change audit market dynamics, we mainly address these issues threefold. First, we describe and identify the demand fundamentals for the current audit market in both France and the UK with a key assumption that public listed companies have the similar preference for the characteristics of auditors in both markets conditional on their attributes. Audit market is defined as a service-providing market with differentiated products, in which client firms picking the auditor in order to maximize their own utility. In the utility maximization, clients consider their own needs for certain attributes of audit firms, such as industry expertise, etc. They also consider their willingness to pay for audit fees and how well the audit firms match with their own attributes like size and so on. Then based on this structure, we can identify these demand fundamentals by using data on listed firms in France and in the meantime, use the UK data to validate the key assumption about the preference of listed firms for auditor choice. Hence when the counterfactual joint audit policy were introduced, the client firms in the UK would react accordingly based on their current preferences. Thus we can simulate how audit market participants would re-optimize auditor choice under different policy scenarios.

Second, under the joint audit policy there is the question that what if clients is not picking individual auditors but consider certain combinations of pairs. It is suggested that certain pairs are preferred by the French firms (Francis et al., 2009). Considering an example, if a client currently has one big four auditor, when joint audit policy kicked in, the client would choose another auditor only based on cost effectiveness, then it is more likely for him to choose a medium or small auditor. And if the argument for better audit quality holds, it is also possible that different auditor pairs might have different synergies. Certain type of pairs might work better than the other, for instance, the Big Four may share the similar work structure and comparable techniques. Hence, in this step we need to measure the pair effect of having two auditors from the French market. Here we incorporate all possible pair combinations (considering big, medium and small auditors) into the clients' utility to choose auditors. In addition, we interact these pairs combinations with clients' attributes to capture the fact that clients also have heterogenous preference for pairs. In other words, synergies between pair combinations might work differently for different clients. we then take such pair effect into the UK joint audit policy simulation, assuming that similar clients in the UK would react in a similar fashion.

Last but not least, we derive counterfactuals on the change of market structure and clients' social welfare by introducing joint audit policy into the UK market. In general, we allow the UK clients to choose another auditor while keeping the original one. This scenario is more likely because the persistence of auditor client relationship. While not changing their preferences on other attributes, we include heterogenous preferences for certain pair combinations. By using the demand estimations from the first two steps, we can address the policy implications of joint audit quantitatively.

The simulation results suggest that the big four auditors would benefit very differently in terms of their market share. The market leader PwC in the UK would experience a significant increase in the market share not only measured by the number of clients, but also would have a fairly increase around 6.5% in the market share measured by clients' assets. The second largest auditor Deloitte would have slight decrease in the market share on the number of clients; but a slight increase in the share measured by clients' assets. The other two big four auditors would have significant percentage decrease in their market shares in terms of both the number of clients and clients' assets. With respect to the medium auditors (in this case Grant Thornton and BDO), the change in their market share shows a different pattern. Both auditors would have a sharp drop, more than 40% in their market share measured by the number of clients; however, their share in the client assets would change in a completely contrasting way. BDO would stay more or less the same, but Grant Thornton would quintuple the original share. The means that although medium auditors lose some clients, they would be able to compete for the big clients under the counterfactual joint audit policy.

Given the simulation results, it seems no surprise that Grant Thornton and BDO are the main voice to support such reform in the audit market since the medium auditors would gain some market share. However, it is surprising that the Big Four would necessarily lose their clients because of joint audit and some of the Big Four would even gain market share, for instance PWC as the market leader would benefit from such policy. The small auditors would not benefit too much compared with the medium auditors. The total market share audited by all small auditors in terms of both the number of clients and client assets would increase very marginally, around 2%. Our counterfactuals also show that if such policy were introduced in the UK, the total consumer surplus would decrease by 7.2 million GBP on average over time. The decomposition of the change in consumer surplus shows that the consumer surplus difference between one single auditor and two single auditors would decrease by 220 million GBP; while the pair effect associated with joint audit would increase the consumer surplus by 212.8 million. It is not surprised that the consumer surplus would decrease if clients were forced to choose another auditor, generally the second-best in the market, but the pair effect would compensate this loss even though not high enough.

The remainder of the paper is organized as follows. First, we relate this paper to the existing literature. Next, we present the details of the structural model to be estimated and discuss some specific issues related to the audit market. Then we describe the dataset used for empirical analysis. Finally, the estimation results and counterfactual analyses are discussed and then followed by the conclusion.

4.2 Literature Review

Our paper relates to studies that describe audit market competition and draw implication on pricing or differences on audit quality from the analysis (Simunic, 1980; Francis et al., 2005; Hay et al., 2006 etc.). This strand of literature describes the audit market as imperfectly competitive and heterogeneous clients pay different price for audit service based on their client attributes. We do not intend to deviate from previous literature and we use the client attributes that previous studies describe as important in the audit pricing (Hay et al., 2006). But our estimation approach offers a broader theme of possible evaluations on the client and auditor relationship. The issues related to demand, supply and strategic responses of market participants are all able to be included under such framework. This allows us to provide more thorough evidences on possible policy effects than the previous papers. The existing studies also show that the Big Four or industry specialist may earn a fee premium. Such evidences on the fee premiums are more prominent for US studies (Numan and Wilekens, 2012). In our setting, the typical auditor attributes includes industry expertise proxies follow what these studies use and these variables are used to capture how certain clients prefer certain audit firms.

As our model describes how listed firms choose auditors, our research is related to the studies on auditor choice. Prior research argues the selection of an auditor could be due to either cost or quality considerations, or both (Knechel et al., 2008). Cost is often associated with audit fees and the quality perspective is often manifested as the Big Four or certain groups of auditors provide better quality. Of course the evidences are rather mixed regarding different types of clients (Francis, 2004). The general consensus is that the characteristics of client firms are affecting their choice of auditors. And what we usually observe from previous studies is that the client attributes associated with audit fees often affect the auditor choice (Craswell et al., 1995; Hope et al., 2012 etc.). Moreover, the existing studies on auditor choice have been more focus on the choices between the Big Four and the non-Big Four. Our research has more detailed choice sets and the choice set contains each individual big four auditor, each individual mid-tier firms and other small small auditors (outside options). The specifications allow us to investigate change in the market structure for more relevant individual audit firms if certain policy were introduced.

Our study models the audit service market as client choosing better matched auditors, and the matching is conditioned on both the client attributes and auditor attributes. So the framework of audit market competition no only captures the cost consideration of client firms but also the quality perspective. It would be ideal to have data on more directly observable auditor attributes such as hourly rate on audit work, but since our model describes the market more from the client firms' point of view, it is adequate to measure the demand fundamentals based on publicly observable attributes.

The demand estimation we use in this paper is well developed in the industrial organization literature (Berry et al., 1995; Berry et al., 2004). But there is very little empirical research on the service related market. There are some unique features about the audit service market. For instance since the listed firms are obligated to have their financial reports audited, there is a minimum amount of service required. In a typical differentiated product market, the price for the same product does not vary across clients. But for the audit market the price (audit fees) differs across clients and only available for those actual choice observed in the data. Hence by addressing these issues in the demand estimation, our paper also contributes to the IO research which investigates demand fundamentals in the general service market. Gerakos and Syverson (2013) is one contemporaneous study that applies similar approach to investigate possible market impact when one of the Big Four fails or mandatory rotation in auditor choice were introduced in the US. But we focus on different policy issues and address the issue of audit fees in a more careful way.

4.3 Demand Model

To model how clients choose audit firms, we apply the random utility maximization approach, rooted in McFadden's choice theory (Mcfadden, 1973). In order to accommodate the auditor choice for clients from both France and the UK in a common framework, we need to specify the fundamental assumption throughout this paper that public listed firms with same characteristics have the same preference for auditors. In principle, the client firms from France and the UK are all listed firms and are subject to the similar sets of accounting and audit regulations, although these two markets are somewhat different and there are different audit policy implementation. In our study we assume the underlying heterogenous preference for auditors are the same across markets for clients in both markets conditional on their characteristics. This assumption does not seem as strong as it appears but we provide empirical validation that it does hold empirically.

We also assume the utilities enjoyed by clients for individual auditor attributes are addable. In another word, we assume the clients' utility of having particular auditor is additive which is similar to the idea that the utility customer enjoy from each unit of merchandise are additive. Of course, the audit service is a product different from typical merchandise, but since the service is difficult to separate into units and the directly measurable variable such as working hours is not observable for us. In the model, each firm chooses his auditor based on the expected utilities from having each of the auditors. So the client firm's preference represents the relative level of client utility for audit service. Then these utilities are defined as the effects of firm specific attributes, auditor specific attributes and match specific attributes. And the match specific attributes refers to the heterogeneous preference of certain client choosing certain auditor.

4.3.1 Choice of single audit

The publicly listed firms (clients) are mandated to purchase the audit service each period (year) to maximize their utility. The audit firms in the choice set for every client includes the Big Four (Ernst & Young, Deloitte, KPMG and PWC), two medium audit firms (BDO and Grant Thornton) and all other audit firms grouped as small auditors. The deterministic part of the utility of client i at period t choosing one of the top 6 (the Big Four plus 2 medium auditors) audit firms $j = 1, \dots, 6$ in both France and the UK is given by:

$$V_{ijt} = \alpha_0 X_{ijt} + \alpha_1 \chi_{ijt} + \sum_{k=1}^{6} (\beta_{1k}\delta_k + \beta_{2k}\delta_k\tau_{it}) - \alpha_2 p_{ijt} + \xi_{jt} \tag{4.1}$$

The deterministic component V_{ijt} of utility is approximated as a function of observed auditors' attributes as well as clients' characteristics. Variable X_{ijt} denotes audit firm j's attributes: industry expertise or industry specialist, defined in the same industry as client

i.[5] This is to capture the fact that clients in certain industries may have systematically different preference for specific auditors. Variable χ_{ijt} denotes the tenure between client i audit firm j in period t. p_{ijt} is the audit fee that client i pays to auditor j in period t, which will be discussed in length in the next subsection. Parameter α_2 captures the marginal willingness to pay a dollar of audit fees; ξ_{jt} denotes unobserved (to researchers) auditor j's attributes, e.g., reputation and quality. [6]

Variable δ_k is the dummy variable for the top 6 audit firms; parameter β_{1k} captures the auditor fixed effect that represents the mean utility for all clients choosing auditor k. It is well known that clients with different size prefer different audit firms: big clients may prefer the Big Four while small clients may prefer the non-Big Four.[7] Hence, we use the interaction between clients' size measured by logarithm of total assets with auditor fixed effect to capture this heterogeneous preference. In principle, we could interact all clients' characteristics with auditors' attributes to allow for a very flexible form of heterogeneous preference, but this requires more variation from the data to identify all the parameters. At this stage, we just use the general notation τ_{it} for clients' attributes and we will specify the exact interaction term in the estimation stage.[8] If client i chooses outside option, i.e, a small auditor (non-top 6 audit firms), we represent the utility as V_{i0t} and normalize it to be zero:

$$V_{i0t} = 0$$

It is a standard approach to normalize the deterministic component of utility of choosing outside option as 0 because utility is invariant to monotone transformations.

[5] Industry classification is based on Famma-French criterion.

[6] "unobserved" term refers some auditors' attributes difficult to measure or observe in the data from researchers' perspective. From clients' perspective, in the model they can observe every attribute when making the decision to choose the auditor.

[7] the Big Four audit almost all the FTSE 100 companies, and 240 of the companies in the FTSE 250 (the House of Lords Economic Affairs Committee, 2011).

[8] Please find the complete description of variables in the Appendix

4.3.2 Choice of joint audit

As joint audit is mandatory in French audit market,[9] clients are obligated to choose a pair, i.e, two different auditors at the same time. Similar to the UK market, the set of single audit firms in France also consists of the top 6 auditors and the small auditors. Therefore, the choice set for clients in France is composed of all possible pairs of auditors. The total number of all possible pairs in the choice set equals to 22.[10] The deterministic utility of client i in period t choosing a single auditor j follows the same specification as equation (4.1) in both markets. However, the utility of choosing a pair of auditors is not simply the sum of individual utility of choosing two single auditors, because the cooperation process between different auditors may vary vastly due to concerns about the reputation, technology platform, auditor liability and so on. Compared with single audit, the unique feature of joint audit hinges on the pair effect that varies across different combinations of auditor types. The most straightforward way to capture pair effect between two audit firms is to define them pair-wise, i.e. for each possible combination of auditor i and j. It is, however, almost impossible to estimate these pair-wise combination effect due to difficulties in computation and identification. Motivated by the observation in the data that clients are interested in certain combinations of different groups of audit firms as in table (4.2), we think group-wise combination are perfect candidate for measuring the interesting pair effect without loss of generality as well as for identification.

Auditors are categorized into three mutually exclusive groups, G_L, G_M, G_S; the group G_L denotes the Big Four; G_M includes two medium audit firms, i.e. top5 and top6, and small auditors in group G_S. In addition, we assume the pair effect is the same for all

[9] According to French commercial law, statutory joint audit is required when firms register in France and issue the report on consolidated financial statement.

[10] Since small auditors also include many small audit firms, around 200 in France, clients can choose two small different auditors as a pair. Thus the total number of possible combination of pairs is given by $\binom{7}{2} + 1 = 22$.

audit firms in the same group:

$$\Gamma(j,k) = \begin{cases} \Gamma_{LL}, & \text{if } (j,k) \in (G_L \times G_L); \\ \Gamma_{LM}, & \text{if } (j,k) \in (G_L \times G_M) \bigcup (G_M \times G_L); \\ \Gamma_{LS}, & \text{if } (j,k) \in (G_L \times G_L) \bigcup (G_S \times G_L); \\ \Gamma_{MM}, & \text{if } (j,k) \in (G_M \times G_M); \\ \Gamma_{MS}, & \text{if } (j,k) \in (G_M \times G_S) \bigcup (G_S \times G_M); \\ \Gamma_{SS}, & \text{if } (j,k) \in (G_S \times G_S); \end{cases}$$

Moreover, the pair effect of joint audit for each client firm i takes the following parametric form:

$$\Gamma_{it}(j,k) = \sum_l \gamma_{0l}\Gamma_l + \sum_l \sum_r \gamma_l \Gamma_l \tau_{rit}, l = LL, LM, LS, MM, MS, SS$$

where γ_{0l} is the constant utility of choosing each specific combination of auditors; τ_{rit} denotes the client i's attributes that affect the choice of pair, such as size, sector, complexity of financial statement, etc. This term captures clients' heterogeneous preference towards different specific groups as in table (4.2), for instance, large clients may prefer two big four auditors as a pair and then one big four combined with one medium auditor; in contrast small clients may prefer one big four coupled with one small auditor or two small auditors as a pair.

The overall utility of client i of choosing a pair of auditor j and k in period t is given by:

$$u^i_{jkt} = V_{ijt} + V_{ikt} + \Gamma_{it}(j,k) + \epsilon^i_{jkt} \tag{4.2}$$

Where ϵ^i_{jkt} is the idiosyncratic preference for a pair of auditors. This is the random component of overall utility, which follows iid type 1 extreme value distribution. Since we have assumed the utility of choosing one single small auditor is normalized to be 0, we also normalize the deterministic part of the utility of choosing a pair of small auditors to be 0 for consistency. As a result, the overall utility of choosing a pair of small auditors equals to

$$u^i_{00t} = \epsilon^i_{00t}$$

The term ϵ^{i}_{00t} denotes random preference shock of choosing this specific pair. It also follows iid type 1 extreme value distribution. Given the distribution of the random shock, the probability of client i of choosing a pair of auditor j and k in period t conditional on client's and auditors' attributes equals to

$$Pr^{it}_{(j,k)} = \frac{exp[V_{ikt} + V_{ijt} + \Gamma_{it}(j,k)]}{1 + \sum_{l_2=l_1+1}^{6} \sum_{l_1=0}^{5} exp[V_{il_1t} + V_{il_2t} + \Gamma_{it}(l_1, l_2)]} \tag{4.3}$$

$Pr^{it}_{(j,k)}$ represents the probability of client i choosing a pair between auditor i and j in period t. As the function form indicates, this probability is monotonically increasing with the utility derived from each single auditor as well as the pair effect between these two auditors.

4.3.3 Audit Fees

As mentioned before, we only observe audit fees for real matches between clients and auditors. Following their approach, we also estimate what audit fees a client would have expected to pay had it hired an audit firm other than the one we observed in the data. A large body of literature has demonstrated that audit fees are associated with measures of client size, client risk, and client complexity as well as auditors' characteristics (Hay et al., 2006). Size measured in clients' total asset generally accounts for a large proportion of the variation in audit fees (Hay et al., 2006). In particular, we use the logarithm of clients' total asset to capture the economy of scale in common practice. Complexity is measured by the number of product segment, number of foreign subsidiaries as well as the number of operating business sectors. We use leverage ratio and current ratio to capture the clients' risk. Loss indicator (equal 1 if loss occurs) and return on assets (ROA) are used to capture clients' profitability. We also control for price to book ratio, growth in sales in prior year, dummy variable to capture whether the firm was a client of the auditor in prior year, industry fixed effect using Fama-French 12-industry classification[11] and time

[11] We run robustness checks for different industry classifications and the results remain similar.

fixed effect in equation (4.4).

$$\ln(p_{ijt}) = \bar{p}_{jt} + \sum_{r} \beta_r \tau_{rit} + \mu_{jt} \tag{4.4}$$

Where $\bar{p}_{jt}$ is the basic audit fee charged by each auditor j in period t and is constant across all clients; the premium of audit fee varies across clients and it is assumed to be a linear function of client's characteristics, equal to $\sum_r \beta_r \tau_{rit}$. Variable μ_{jt} denotes the iid normally distributed error term.[12] This equation (4.4) implicitly assumes that the rule of setting audit fees is a common knowledge between auditors and clients. Hence clients know exactly how much audit fees they would expect to pay if they decide to switch another auditor.

4.3.4 Endogenous Audit Fees

An obvious concern in the demand estimation are the endogenous audit fees. If we leave some unobserved or unmeasured auditor attributes, e.g., audit quality and reputation, etc, in the error term, the audit fees charged by each auditor will be correlated with the error term, i.e., $cov(p_{ijt}, \xi_{jt}) \neq 0$. As Gerakos and Syverson (2013) suggested, we can use the supply shock among audit firms as an instrument variable to correct the upward biased coefficient of audit fees. Intuitively, if there were mergers and acquisitions (M&A) between clients, the supply structure of audit firms would be changed because one of transaction parties in M&A has to drop the original audit firm. In the next period, the dropped audit firms will use attractive audit fees to compete for new clients to compensate the client loss from M&A. Therefore, the supply shock induced by M&A between clients in the previous period is correlated with audit fees (p_{ijt}) but uncorrelated with demand shifts (ξ_{jt}).

However, differing from their procedure of dealing with endogenous audit fees, we use the same instrumental variable, but apply the control function approach which is

[12] In the current version of this paper, we only use OLS to predict the audit fees. The estimation results show such approach is sufficient. We will try other regression methods e.g., random forest as Gerakos and Syverson (2013) suggested for robustness check later.

more appropriate in the discrete choice model.[13] The general idea underlying the control function correction is to find a proxy for the unobserved auditor attributes ξ_{jt}. With the proxy of ξ_{jt} in the demand estimation, the variation of endogenous audit fee will be independent of the error term. And the estimation result of the standard approach would become consistent (Petrin and Tain, 2010). More precisely, we first use z_{t-1}, the ratio of three-digit SIC industry assets merged in the prior year to instrument the variation of audit fees in the current period as in equation (4.5).

$$\ln(p_{ijt}) = \bar{p}_{jt} + \rho_0 z_{t-1} + \sum_r \rho_r \tau_{rit} + \mu_{jt} \tag{4.5}$$

μ_{jt} and ξ_{jt} are independent of z_{t-1} and τ_{rit}, but are not independent of each other. The key idea of control function approach is that we can use μ_{jt} as the proxy variable for audit quality ξ_{jt} and then obtain the consistent estimators of demand preference condition on it. After the first stage regression of equation (4.5), residual $\hat{\mu}_{jt}$ enters the demand estimation in the second stage. The general control function approach allows for a flexible function form of $\hat{\mu}_{jt}$ in the second-stage estimation

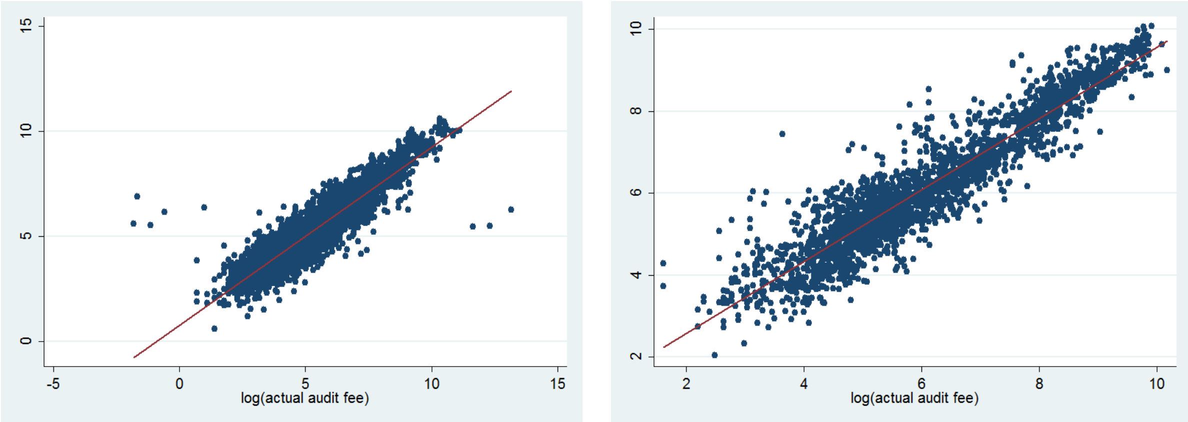
$$\xi_{jt} = h(\hat{\mu}_{jt}) \tag{4.6}$$

Where $h(\hat{\mu}_{jt})$ denotes the control function. The simplest form of $h(\hat{\mu}_{jt})$ would be linear function. Alternatively, a high order polynomial approximation can be used for robustness check. It is worth mentioning that the unobserved auditor attributes ξ_{jt} which does not vary across clients implicitly make it feasible to apply the control function approach. Otherwise, we can not obtain the proxy for audit quality from the unobserved audit fees.

Similar to Gerakos and Syverson (2013), Table (4.1) shows that the coefficients of "Scaled merged assets" are significantly negative, which means the bigger mergers and acquisitions of clients are associated with lower audit fees in the following period. The industry wise shock presented by M&A activities creates a downward price effect. Hence

[13] Please find the details of control function approach in Petrin and Train (2010). In the literature of health economics, the way used by Gerakos and Syverson (2013) is the two-stage predictor substitution (2SPS) while the control function approach is called two-stage residual inclusion (2SRI). 2SRI is generally statistically consistent for nonlinear models, but 2SPS is not by Terza, Basu and Rathouz (2008).

Figure 4.1: Actual v.s. predicted audit fees



(a) UK (b) France

we can use supply shock from M&A between clients as a valid instrument for audit fees.

Figure (4.1) illustrates the plots of actual versus predicted audit fees in both France and UK. As shown in Figure 1, the predicted audit fees fit well with the actual audit fees. The correlation between predicted audit fees and actual audit fees is larger than 0.93 in both countries. As we can see the average magnitude of audit fees in the UK is relatively higher and a larger proportion of them distributed between the range of above 5 and 10 compared with the distribution in France,[14] this is mainly driven a larger proportion of big clients measured in assets in the UK.

4.4 Data

The sample in our study consists of the listed firms in the UK and France with available data. The sample period lasts from 2005 to 2012. Our data are from commercial databases and publicly available financial reports of listed firms. The data on client attributes and auditor-client matches are from Amadeus database. We pull the audit fees from Datastream for the UK firms and hand-collected data from annual reports for French firms. And we also obtain the mergers and acquisitions data from SDC database.

As in Figure (4.2). The average market share measured by number of clients over the sample period is around 13% among the big four auditors in the UK. For small auditors,

[14]The scale in Figure (4.1) equals to the logarithm of thousand audit fees in the local currency, i.e., GBP for the UK and EURO for France.

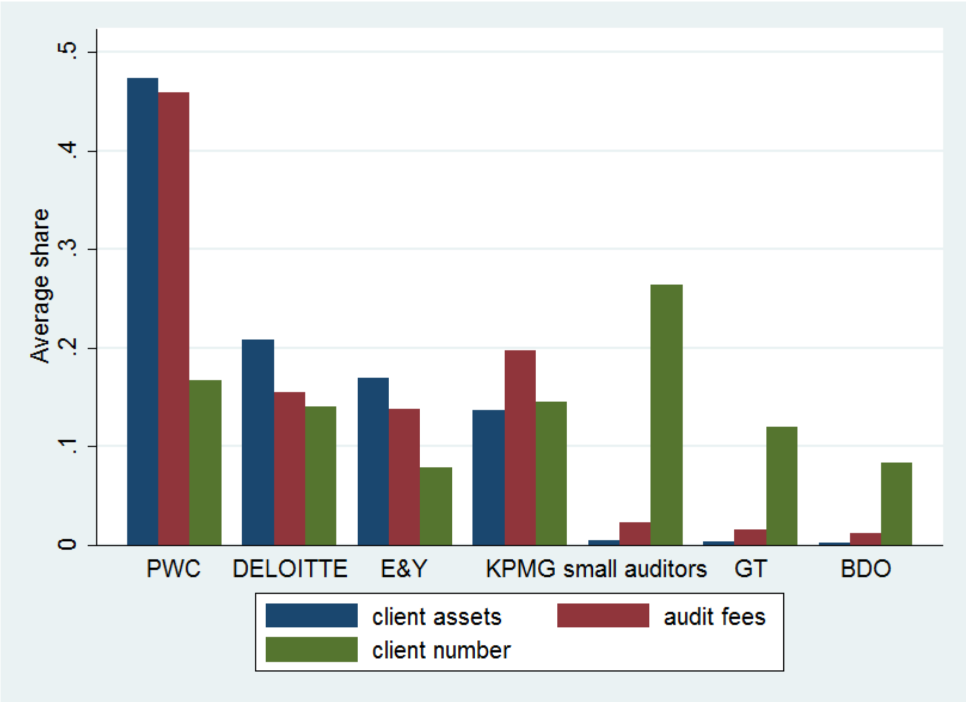
Table 4.1: M&A between clients as supply shock during 2005-2012

ln(Audit Fees)	UK	France
scaled merged assets	-0.1653**	-0.2878***
	(0.0560)	(0.000)
ln(assets)	0.5072***	0.6842***
	(0.0219)	(0.0144)
leverage ratio	0.0004***	-0.0026
	(0.0001)	(0.0021)
No. geographical subsidiary	0.0945***	0.0251**
	(0.0105)	(0.0083)
No. product segment	0.0865***	0.0283
	(0.0136)	(0.0196)
current ratio	-0.0042*	-0.1140***
	(0.0019)	(0.0306)
price to book value	0.0007	0.0073**
	(0.0007)	(0.0026)
sale growth	0.0001***	0.0053
	(0.0000)	(0.0516)
tenure	-0.0002	0.0220
	(0.0181)	(0.0362)
cross listed	0.2774***	0.2429***
	(0.0679)	(0.0510)
receivable to assets	0.1217*	1.0523***
	(0.0649)	(0.1589)
ROA	-0.0024***	-0.0089*
	(0.0007)	(0.0041)
loss dummy	0.0611	0.1041*
	(0.0523)	(0.0524)
location	0.1395***	0.1766***
	(0.0323)	(0.0433)
constant	-0.5976*	-3.0979***
	(0.2727)	(0.2704)
Auditor fixed effect	Yes	Yes
Year effect	Yes	Yes
Industry fixed effect	Yes	Yes
Observations	6159	2392
Adjusted R^2	0.6612	0.9122

Notes: Standard errors clustered at industry-level; ***, ** and * denote significance at the 1%, 5% and 10%.

Scaled merge assets denotes the ratio of occurred M&A assets over total assets in three-digit SIC industry; cross listed is a dummy, equal to 1 if the client is cross listed in the US stock market; location is an indicator whether the headquarter of the client is located in the capital of the country.

Figure 4.2: Market shares of audit firms in the UK during 2005-2012



their total market share in terms of the number of clients is up to 27% average over time, but there are between 200 and 300 small auditors in all each period, resulting in the market share per small auditor almost trivial in the market. However, if we use the total assets audited by each auditor to measure the market share, the audit market is extremely dominant as expected in the UK. The sum of this average share by the Big Four is more than 90%, and the market leader PwC alone even reach 47%, almost half of the total markets. In this case, two medium auditors- BDO and Grant Thornton also becomes marginal in the market. With respect to the share of audit fees, it is positively correlated with total assets audited by each auditor. In brief, big clients are audited by the Big Four and other clients are shared by medium and small auditors.

As joint audit is mandatory implemented in France, Table (4.2) describes the distribution of different pairs between audit firm during the sample period. L denotes one of the big four audit firms; M denotes one of the medium audit firms; S denotes one of the small audit firms in France. LL denotes a pair composed of any two of the big four audit firms. LM denotes a pair composed of one of the big four audit firms and one of the

Table 4.2: The distribution of different pairs in France 2005-2012

Group	Percentage	Client Size		
	(%)	I	II	III
LL	24.18	9.48	18.60	44.64
LM	18.17	7.73	17.39	29.43
LS	37.99	48.88	44.20	20.70
MM	0.99	1.75	1.21	0.50
MS	7.89	12.72	10.39	0
SS	10.77	19.45	8.21	4.74

medium audit firms. LS denotes a pair composed of one of the big four audit firms and one of the small audit firms. Client size I, II, and III denotes small, medium and large clients measured in total assets respectively.[15] It is clear that clients have heterogenous preference for specific pairs under the joint audit policy as in Table (4.2). Although LS is the most prevalent pair in general, around 38%, large clients strongly prefer LL and LM to LS; while it remains the most preferred pair for medium and small clients.

Table (4.3) shows the basic statistic summary of audit fees charged by each auditor in both countries. The audit fee in the UK is in GBP and EURO for France. We first divide both original fees by thousand and then take the logarithm, leading to the observations in (4.3). It is not surprise that the big four auditors on average charge a higher audit fee than the medium auditors and similarly the medium auditors charge higher fees than the small auditors, because the audit fees mainly depends on the workload which is measured by clients size. Compared with France, the audit fees charged by small auditors, medium auditors and the Big Four in the UK on average is relatively higher, but with a smaller variance. The composition of two medium auditors is different across the two countries. The fifth largest auditor in the UK is Grant Thornton (GT) and Mazars in France; while the sixth auditor auditor is BDO in the UK and GT in France. Therefore, the label for the medium auditors can be commonly regarded as a ranking or recognition of auditor reputation in both countries.

[15]The total assets are discretized by the 3-quantiles. So the top 33% are large client, the medium 33% are medium client and the bottom 33% are small clients.

Table 4.3: Summary of audit fees in France and the UK during 2005-2012

	mean	**sd**	**median**	**min**	**max**	**N**
UK						
PwC	12.871	0.231	12.720	12.674	13.367	1314
E&Y	11.666	0.165	11.576	11.478	11.970	629
Deloitte	11.796	0.127	11.760	11.636	12.087	1111
KPMG	11.984	0.577	11.789	11.630	13.430	1148
GT	9.473	0.306	9.445	8.964	9.914	936
BDO	9.234	0.148	9.192	9.012	9.436	657
Small auditors	6.224	0.211	6.155	5.982	6.593	2114
overall	10.015	2.568	11.571	5.982	13.430	7909
France						
PwC	11.047	0.614	11.131	8.082	11.359	232
E&Y	11.754	0.750	11.882	8.050	12.122	389
Deloitte	11.161	0.556	11.209	7.979	11.516	350
KPMG	11.076	0.659	11.232	8.038	11.396	299
Mazars	10.867	0.780	11.175	7.762	11.406	254
GT	8.758	0.911	8.589	4.771	9.724	87
Small auditors	5.709	0.663	5.402	4.558	6.585	819
overall	9.280	2.688	11.069	4.558	12.122	2430

4.5 Demand Estimation Results

4.5.1 Demand Estimation Results in French market

The demand model is estimated in two steps as required by the control function approach. We first regress the endogenous variable (audit fee) on other observed clients' characteristics and the instruments. The corresponding estimation results have been reported in Table (4.1). The residuals of first-stage regression are used to compute the control function, which enters the discrete choice model as an extra variable in the second step. Then we implement bootstrap to correct the standard error for the two-step estimators (Petrin and Train, 2010).

The Table (4.4) presents the general preference of public listed firms in France estimated by conditional logit approach. The fist column in this table does not use control function to address the endogenous audit fee; while the other two does and allows for a different form of control function. As expected, control function approach helps to correct the biased coefficient of willingness to pay the audit fees in column 1, from -0.4 to -0.5 as

Table 4.4: Demand Estimation in France

industry leader	3.8726***	3.9275***	3.9301***
	(0.5141)	(0.5901)	(0.5892)
industry specialist	4.1172***	4.1692***	4.1675***
	(0.4647)	(0.5167)	(0.5145)
tenure	22.6902***	25.2915***	24.5025***
	(3.3080)	(5.0744)	(5.3123)
ln(audit fee)	-0.4422*	-0.5003**	-0.5029**
	(0.2338)	(0.2356)	(0.2374)
PwC	-5.9710***	-5.8629***	-5.6679***
	(2.2027)	(2.1539)	(2.1270)
E&Y	-6.0621**	-6.0183***	-5.8149**
	(2.3802)	(2.3253)	(2.3095)
Deloitte	-5.1957***	-4.8842***	-4.6111***
	(1.1578)	(1.0613)	(1.1060)
KPMG	-5.0292***	-5.1046***	-4.9487***
	(0.8568)	(0.7737)	(0.7794)
Mazars	-6.3394***	-6.1735***	-5.8366***
	(1.8086)	(1.7496)	(1.7375)
GT	-4.4489**	-4.7346**	-4.8091**
	(1.9534)	(1.9274)	(1.8683)
PwC*ln(assets)	0.4122**	0.4013**	0.3819**
	(0.1758)	(0.1707)	(0.1707)
E&Y*ln(assets)	0.4248**	0.4181**	0.3981**
	(0.1781)	(0.1734)	(0.1728)
Deloitte*ln(assets)	0.3434***	0.3185***	0.2948***
	(0.0949)	(0.0865)	(0.0902)
KPMG*ln(assets)	0.3403***	0.3428***	0.3256***
	(0.0663)	(0.0620)	(0.0607)
Mazars*ln(assets)	0.4676***	0.4610***	0.4396***
	(0.1316)	(0.1287)	(0.1278)
GT*ln(assets)	0.2211	0.2515	0.2581
	(0.1703)	(0.1694)	(0.1643)
audit quality		1.9334*	2.6697**
		(1.0020)	(1.0917)
audit quality square			-5.5148
			(4.5755)
pair1*geography	0.1587	0.1691	0.1764
	(0.1564)	(0.1536)	(0.1557)
pair2*geography	0.1133	0.1220	0.1145
	(0.1143)	(0.1092)	(0.1083)
pair3*geography	0.1896	0.1906	0.1889
	(0.1226)	(0.1228)	(0.1223)
pair4*geography	0.0242	0.0273	0.0235
	(0.1306)	(0.1275)	(0.1310)
pair5*geography	-0.1502	-0.1501	-0.1494
	(0.1176)	(0.1149)	(0.1139)
pair1*receivable	-3.9552***	-3.7406***	-3.4206***
	(0.8560)	(0.7670)	(0.7621)
pair2*receivable	-3.1668***	-3.6651***	-3.4826***
	(1.1258)	(1.0466)	(0.9643)
pair3*receivable	-0.2766	0.1385	0.1863
	(0.6323)	(0.6902)	(0.6884)
pair4*receivable	-1.1392	-0.9592	-0.8633
	(4.2133)	(4.1956)	(3.9982)
pair5*receivable	0.2172	0.7096	0.8033
	(1.4453)	(1.4997)	(1.3760)
pair1*location	1.7254***	1.7663***	1.7866***
	(0.4076)	(0.3750)	(0.3669)
pair2*location	0.2319	0.1530	0.0870
	(0.4269)	(0.3896)	(0.3809)
pair3*location	0.9708**	0.9327*	0.9145*
	(0.4799)	(0.4937)	(0.5005)
pair4*location	-15.7572***	-17.4911***	-16.2370***
	(0.6808)	(0.6870)	(0.6735)
pair5*location	0.3731	0.4281	0.5002
	(0.5297)	(0.5639)	(0.5300)
R^2	.8907496	.8911568	.8913553
$-L$	-401.2769	-399.7815	-399.0523

Standard errors clustered at industry-level.

Notes: ***, ** and * denote significance at the 1%, 5% and 10%.

in column 2. In column three, we add a higher order term in the control function, but the estimated parameter is fairly close to that in column 2 and the square term is not significant. Thus we stick to the estimation results in column 2 and use them for counterfactual analysis in next subsection. The proxy for audit quality also have the expected positive sign and significant, which implies control function provides a good approximation for audit quality.

The estimation results above audit quality display clients' preference for individual auditor. Variables involving tenure, industry leader and industry specialist are all significant and have the expected sign.[16] That is to say, given everything else being equal, public listed firms prefer low audit fees, more likely to choose the auditor that already has a long run tenure, and also prefer auditors that are industry leader and specialist. The interaction terms between clients' size and auditors' fixed effect capture clients' heterogeneous preference. The positive sign of these interaction terms suggest that client with larger size prefer the top 6 auditors compared with small auditors. The magnitudes of interactions on the top 5 auditors are similar but much larger than the interactions on one medium auditor (GT), which is consistent with the observations in the data and findings in the literature. Bigger clients would strongly prefer the big four auditors over the medium auditors, although the big 6 auditors are preferred by big clients in general.

Pair1 to pair6 are the pair dummies, representing big-big, big-medium, big-small, medium-medium, medium-small and small-small pairs between audit firms respectively.[17] The small-small pair (pair6) is used as the base due to the normalization and collinearity. The pair dummies are interacting with clients' attributes including financial complexity measured by the number of foreign subsidiaries, risk measured by ratio of receivable over total assets and willingness for high reputation measured by variable location. Location equals to 1 if this client is located in the capital of the country, otherwise 0. Normally speaking, clients would like to pay for large premium for auditors with high reputation if

[16]Detailed definition of these variables are present in the Appendix. ln(assets) is the natural logarithm of the client's total assets.

[17]The estimation results in Table (4.4) do not contain pair fixed effect because of the identification issue. The pair dummies are just certain linear combination of individual dummies, thus we can not identify pair fixed effect alone when controlling for clients' preference for each individual auditor.

their headquarter is located in the capital of a country. As in Table (4.4), the interaction between clients' financial complexity shows a positive sign, though not significant. This means as the financial structure of public listed firms become complicated, these clients would prefer pairs with at least one auditor from the Big Four or medium-tier. It is consistent with the perception that the Big Four are more capable of handling financially complicated clients. The coefficient of interaction between receivable and pair fixed effect reach significantly negative sign in such pairs as big-big and big-medium. This suggests if clients are getting risky in terms of high ratio of receivable over assets, they will be less likely to choose pairs with at least one auditor from the Big Four. This indicates the Big Four might avoid risky clients because their better risk management. Regarding the interaction between location and pair dummies, we find significantly mixed sign for distinguished pairs. Clients with headquarter in the capital most prefer big-big pair and then big-small pair, but strongly dislike medium-medium pair compared with small-small pair base. These results confirm that there is the pair effect under the joint audit policy. Moreover clients have heterogenous preference for specific pair combinations in addition to its original preference for single auditor.

4.5.2 Model fit for the UK market

In the beginning of this paper, we have imposed a basic assumption that publicly listed firms have the same preference for auditors conditional on their attributes in both countries. We will validate this assumption in this section in spite of a variety of country differences. Although Table (4.4) presents the results of the demand estimation in the French market, it also contains the information on the public listed firms' preference for individual auditors. These parameters for variable tenure, industry leader, industry specialist, individual auditor dummy and its interaction with assets are assumed to be the same as clients in the UK. We use them to predict the single auditor choice for clients in the UK and compare the prediction with actual choice observed in the data.

$$u_{ijt}^{uk} = V_{ijt} + \epsilon_{ijt}$$

The utility of client i choosing single auditor j at period t in the UK market is represented by u_{ijt}^{uk}. It composes of two parts: the deterministic part V_{ijt}, exactly the same formula as in France and the random part ϵ_{ijt} also iid extreme value type 1 distribution. Similarly, the utility of choosing small auditors is normalized as:

$$u_{i0t}^{uk} = \epsilon_{i0t}$$

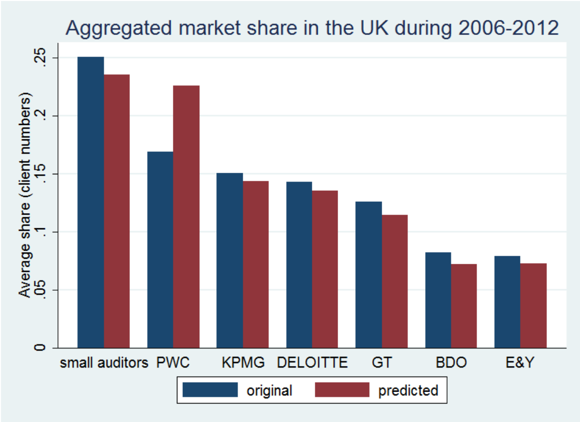
Given this utility specification as well as the preference for single auditors derived from French market, we can compute the probability Pr_{ijt}^{uk} of client i choosing each single auditor j in each year t in the UK conditional on her attributes τ_{it}. We then use the corresponding highest probability as the predicted choice to compare with actual choice in the data.

Table 4.5: Model fit in the UK market during 2006-2012

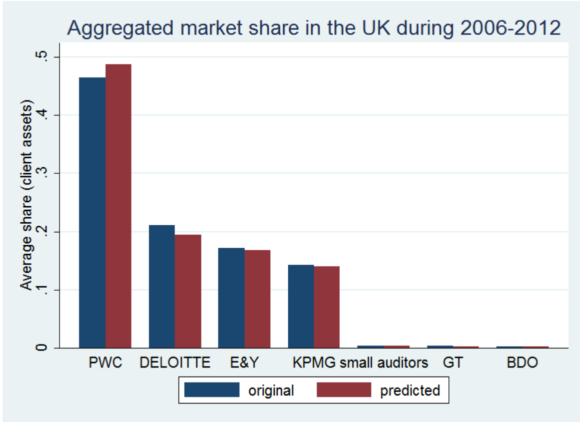
Actual Choice	PwC	E&Y	Deloitte	KPMG	GT	BDO	Small auditor
PwC	93.8%	0.5%	1.4%	1.2%	0.9%	0.1%	2.1%
E&Y	10.2%	84.2%	1.4%	1.2%	0.6%	0.4%	1.8%
Deloitte	8.3%	1.0%	86.7%	0.7%	0.8%	0.3%	2.2%
KPMG	5.1%	0.9%	1.1%	89.2%	0.8%	0.9%	2.2%
GT	8.1%	0.5%	1.6%	1.3%	83.5%	0.5%	4.5%
BDO	7.1%	0.6%	1.4%	1.6%	0.8%	82.2%	6.3%
Small auditor	9.9%	0.3%	1.2%	1.0%	0.9%	0.4%	86.3%

Table (4.5) shows how the preference parameters derived from French market fit the observations in the UK. The row denotes clients' actual choice and the column displays the predicted choice according to highest probability. The first row should be interpreted as conditional on the actual choice of clients choosing PwC during 2006 to 2012, the model predicts that 93.8% of these clients choose PwC, which coincides with the actual choice. And 0.5% of these clients are predicted to choose Deloitte and so on. Therefore, the numbers on the diagonal of Table (4.5) indicate the fitness of the preference parameters. On average 86.5% of the predicted choice is consistent with actual choice in the UK, in particular for clients that actually choose PwC, the correctness of prediction reaches 93.8%. Regarding clients that choose other auditors in the UK, on average around 85% predictions coincides with actual choice and PwC seems to be the second best choice

Figure 4.3: Model fit of the UK during 2006-2012



(a) Clients' number



(b) Clients' asset

among these clients. Please note that we only use UK sample from 2006 to 2012, because the variable tenure has more than 67% missing values in the first year 2005. Consequently, we drop the first year 2005's observations in the UK and all the following counterfactual analyses are also based on the sample from 2006 to 2012. Figure (4.3) presents another measure how the preference parameters fits the UK data in terms of aggregated market share. In general, the predicted market shares by both number of clients and assets of clients fit the actual share quite well. However, the aggregated share of PwC is overpredicted in both measures, because PwC is systematically over predicted for clients that choose other auditors as shown in Table (4.5).

4.6 Counterfactual Policy Analysis

The European Commission has been concerned about the high concentration of Big Four in the UK market. France is the only country that implements mandatory joint audit policy and has the least concentrated audit market in Europe. The debate on this policy has been controversial. In this section, we would like to provide a guideline for policy makers on the potential impact of joint audit policy in the UK market.

4.6.1 Change of Market Share under Joint Audit in the UK

If public listed firms in the UK were mandated to choose two audit firms under joint audit policy, two possible cases would arise. One case would be that clients still keep the current auditor and choose a second auditor. The other one would be that clients drop current auditor and choose two new auditors for the joint audit service. The first case is more likely to be expected to be true because the variable tenure in the demand estimation shows that clients prefer to establish a long-run relationship with audit firms. Therefore, we simulate how the audit market structure evolves in the UK under the joint audit policy, in which clients keep the original auditor and meanwhile choose a second audit firm. Then the utility of client i in period t keeping original auditor j_0 and adding another auditor j_1 in the UK would be

$$u^{uk}_{ij_0j_1t} = V_{ij_0t} + V_{ij_1t} + \Gamma_{it}(j_0, j_1) + \epsilon_{ij_0j_1t}$$

$\Gamma_{it}(j_0, j_1)$ represents the pair effect between auditor j_0 and j_1, and it varies across client i according to their individual characteristics. It is worth mentioning how we calculate the predicted audit fee under joint audit policy. The two auditors are supposed to share workload and charge each individual audit fee associated with the separated workload. The criteria for dividing workload in the counterfactual is derived from the observed ratio in the French market. In addition to the shared workload, the individual auditor's attributes also multiply the associated workload ratio to enter V_{ij_0t} and V_{ij_1t} under joint audit policy. Take individual fixed effect for example, client i can enjoy all the utility from the fixed effect (reputation) of auditor j_0 under single audit, while under the joint audit, client i's utility from original auditor j_0 would get a discount because auditor j_0 now only provide part of auditing service under the joint audit. Given the each individual predicted audit fee $\hat{p}_{ij_0t}$ and $\hat{p}_{ij_1t}$, clients' preference parameters, auditors' attributes, and the random draw of the idiosyncratic preference shock, we can compute each client's optimal choice in every period. As a result, we calculate every auditor's new market share in the counterfactual joint policy.

Figure 4.4: Change of market shares in the UK (clients' number)

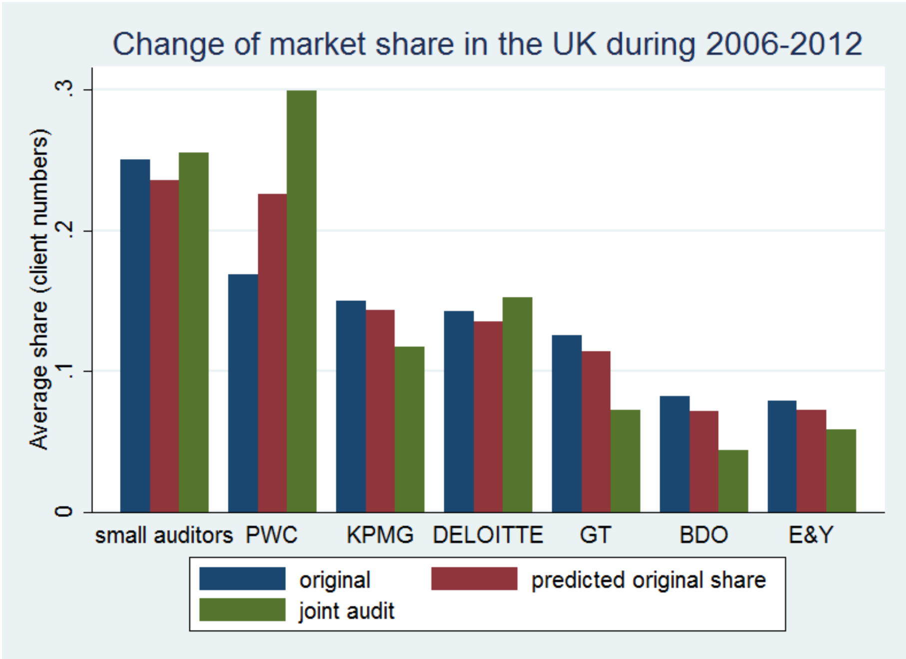
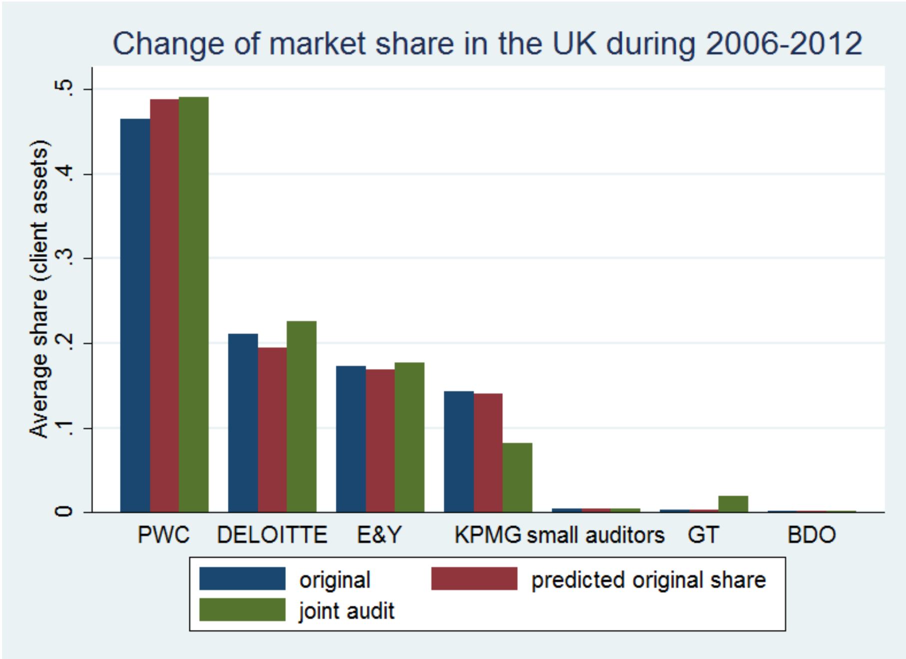


Figure (4.4) and Figure (4.5) show the market share would change for each auditor in terms of number of clients and assets of clients respectively under the joint audit policy. In both figures, we plot the actual share observed in the data, the predicted share under single audit as in previous subsection and the share under joint audit. The predicted share under single audit provides another necessary benchmark to compare with counterfactual scenario because it helps to provide a robustness change in the market share with controlling for the prediction error.

As shown in the figures, the market share of the Big Four changes quite differently from each other in both on the number of their clients and the asset size of their clients. The market leader PwC would have a significant increase in the share of clients' number, from originally around 17% increase to around 30% under joint audit policy. Its market share of clients' assets would also have a fair increase, from around 46% to around 49% average over time. The direct follower after Pwc in the UK market is Deloitte, who also experience a slight increase in both market share measures. The third Big Four- Ernst & Young's share of clients' assets would almost stay the same as in single audit, while its

Figure 4.5: Change of market shares in the UK (clients' asset)



share of clients' number would have a fair decrease. The forth Big Four KPMG in the UK market would experience a sharp drop in the share for both number of clients and assets of clients. The market share loss for KPMG and Ernst & Young seems to justify their incentive to lobby against this potential reform.

It is interesting to see how two medium size auditors change their market share since they are quite in favor of this policy reform (Jones 2012). As both figures shows, Grant Thornton (GT) and BDO would loss at least one third of their original share in number of clients, but would benefit substantially for their share in clients' assets. This market share measure for Grant Thornton would quintuple, from originally 0.36% to around 1.8%. That means Grant Thornton would be able to compete for some big clients even though loss some small clients under the joint audit. However for BDO, this measure would increase very tiny, almost the same as before. With respect to small auditors, they would get relatively bigger share of the pie than before. Although on average, their market shares still are trivial to the bigger auditors as in Figure (4.4) and Figure (4.5). Their total market shares would increase by 2% on average over time in terms of both clients' number and assets.

Table 4.6: Patterns of chosen pair under the joint audit policy in the UK

	Big4	Medium	Small
PWC	0.8904	0.0891	0.0205
E&Y	0.9699	0.0115	0.0186
Deloitte	0.9628	0.0177	0.0195
KPMG	0.9612	0.0181	0.0208
GT	0.9564	0.0043	0.0393
BDO	0.9294	0.0072	0.0634
Small auditors	0.0295	0.0138	0.9567

Table (4.6) shows detailed pattern how clients choose pairs conditional on keeping their current auditors in the UK. Around 89% clients that originally choose PwC would choose another Big Four to form a pair; and then 9% of them would choose one medium firm as pair, leaving the left 2% to choose a small auditor in a pair. For these clients that originally choose other Big Four auditors, more than 96% of them would choose another Big Four (most likely PwC) as a pair choice under joint audit, and the remaining has a relatively higher probability to choose small auditors compared with medium auditors. The same pattern holds for clients that choose medium auditors originally. Regarding small auditors' clients, they would like to choose another small auditor as a pair because they do not benefit too much by choosing one Big Four or medium auditor indicated by their preference parameter. The simulation results seems unexpected to some extent because we observe that big-small pair is the generally most prevalent pair choice in France. But Table (4.2) also shows that clients with large amount of assets have a strict preference ordering as big-big big-medium, big-small pair and other pairs. Since clients' size in the UK market on average is larger than in France, it is not surprising that most of them would choose big-big pair if joint audit policy were introduced in the UK. Combined this table with previous figure, we can well explain which channel drives the change of the aggregated market share. PwC would lose a few big clients, but harvest more relatively small clients from other Big Four and medium auditors under joint audit policy. That's why its share in number of clients increases much higher than the share in assets of clients. The medium auditor-Grant Thornton would successfully compete for some big clients that originally choose PwC, leading to the soar of its share in clients' assets. Small auditors would benefit very marginally due to distribution of clients' size in the UK. In all, the concentration of the Big Four auditors measured by the sum of their market shares in clients' assets drop slightly because of the rise of the medium auditors under joint audit.

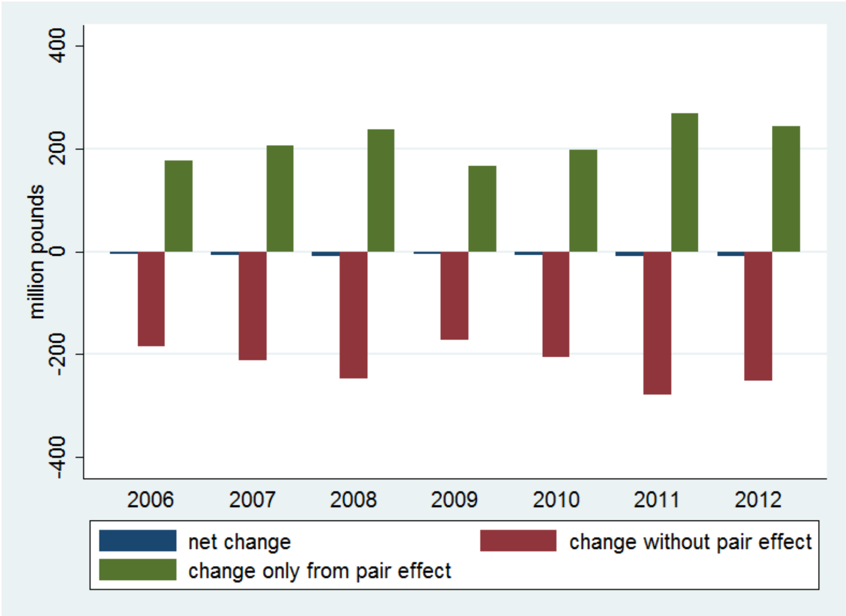
These results show the possible introduction of joint audit would significantly shape up the current market structure in the UK. However, we would like to point out that these results do not include the possible strategic price response from auditors, especially the Big Four. Since the Big Four have much more market power in the current market, they would react to the policy change by setting a new optimal price in order to compete for more clients. The same applies to other auditors, but they might be in a disadvantaged position to compete. In addition, there would not be any entry or exit of audit firms during the sample period if the joint audit were implemented in the UK. Basically, we focus on the short-run effect of joint audit policy. The general equilibrium model with audit firms entering or exit would enable us to investigate the long-run policy effect, but it is much more complicated and beyond the scope of current version. Hence our current results should be interpreted with caution.

4.6.2 Change of welfare under Joint Audit in the UK

The welfare in this paper is equal to consumer surplus of all clients in the UK since we do not model cost function in the supply side so far. To estimate the change of consumer surplus, we apply the approach developed by McFadden (1999): calculate the expected change in consumer surplus for each client as the expected dollar transfer required to make that client indifferent between choosing original auditor in the single aduit and choosing new auditors arising under the counterfactuals. Then we sum the change across all clients to obtain the expected total change in consumer surplus.

As a similar procedure to compute the change of market share in the counterfactual, we can simulate each client's optimal auditor choice and then compute the difference of the maximum utility for each client derived in both original and counterfactual world. As shown in Figure (4.6), clients would slightly be worse off over the sample period in the counterfactual joint audit policy, ranging from 9 million GBP to 5.6 million GBP.

Figure 4.6: Change of welfare in the UK after joint audit policy



The estimated average change in welfare over time would decrease by 7.2 million GBP. However, the decomposition of consumer surplus change suggests clients instead would benefit from pure pair effect from joint audit to a great extent, on average around 212 million GBP better off, but unfortunately this compensation is not high enough to balance out the loss from being forced to choose another auditor in the pair.

To mitigate the concern that these simulation results are driven by specific forms, we try alternative utility function forms, i.e., adding several more clients' attributes to interact with pair fixed effect as a robustness check. We first re-estimate all the preference parameters associated with each specification and then use them to simulate the counterfactual results. We find that the change in the direction of above all shares as well as consumer surplus is quite consistent and robust, but the percentage of the change varies across different specifications.

4.7 Conclusion

Within EU, the European Commission green paper (2010) raises the issue of audit market concentration. The UK regulators are also extremely concerned with the concentration of audit market. The report from the House of Lord (2011) indicates that they believe there is a lack of choice in the UK audit market, especially for large client firms. While a series of policy debates are ongoing, some possible reforms are proposed. In this paper we investigate how one of the possible reforms, namely mandatory joint audit would affect the audit market concentration in the UK. The demand estimation approach allows us to identify clients' preferences to substitute among individual auditors. Using observations in French market, we can measure how listed firms perceive services provided by the Big Four, the medium auditors and the small auditors. In the meantime, we are able to identify how different firms choose different pairs of auditors. In the policy experiment, we force the UK clients to choose another auditor while keep their original auditor under joint audit. Given the preference for individual auditors as well as for pair choices derived from French market, we can simulate how listed firms in UK respond to such policy change and the potential evolution of market structure.

Our demand estimation results show that the public listed firms in both countries have heterogenous preference for the big four auditors, mainly varying across clients' size. While considering the audit market in France, the heterogenous preference is also manifested in the pair choice. That is to say, different groups of listed firms do prefer certain types of pairs, e.g., the bigger firms prefer the combination of having two big four auditors as a pair. After recovering preference parameters, we calculate the market share changes of audit firms and the welfare change of client firms in the UK under the counterfactuals. Our results show that the market leader auditor would experience substantial rise in the share of number of clients as well as a fair increase in the share of clients' assets. The second auditor would enjoy a small growth in the both market share measures. However, for other two big four auditors, they would have to incur market share losses: on average 20% decrease of client numbers and on average of over 25% decrease of client sizes. Even though, both medium auditors would loss the share of clients' number substantially, one medium auditor-fifth largest player in the UK market-Grant Thornton would quintuple its share of clients' assets after joint audit policy. Another medium auditor BDO would not benefit much from such reform. And the small auditors would benefit very marginally from the joint audit by expanding their market shares by a very tiny percentage. The concentration of the Big Four under joint audit would drop mildly due to the rise of medium auditors. The counterfactual results also indicate that joint audit would increase clients' consumer surplus to some extent due to the positive pair effect, but this pair benefit is not high enough compensate the welfare loss from being forced to choose another auditor. As a consequence, the net welfare change for clients in UK would be negative and decrease by 7.2 million GBP on average over the sample period.

We would like to point out that although the evidences suggest dramatic changes, we would interpret these counterfactuals with caution. Nevertheless, these estimates are informative about the trade-offs of changing the auditor choice of clients and the cost v.s. benefits of changing audit market structure. Currently we are considering several extensions to further the discussions. For instance, we include modeling of strategic price responses from auditors in the policy simulation and the comparison of audit fee changes.

Appendix

Table 4.7: Variable definition

Industry	Fama-French 12 industry classification
Industry leader	equal 1 if the audit firm has the highest asset market share in each industry, otherwise 0
Industry specialist	equal 1 if the audit firm has a fee market share over 30% in each industry, otherwise 0
Tenure	equal 1 if the public listed firm is a client of the audit firm in the last year, otherwise 0 in the UK
	equal 0 if the public listed firm is not a client of any audit firms in the pair in the last year in France
	equal 1 if the public listed firm is a client of one of the pair of two audit firms in the last year in France
	equal 2 if the public listed firm is a client of both of the two audit firms in the last year in France
Size	the natural logarithm of total assets
No of industrial segments	includes number of business segments and number of geographical segments
Leverage ratio	the ratio of short plus long term debt to total assets
Current ratio	the ratio of current assets to current liabilities
Quick ratio	the ratio of cash and receivable to current liabilities
Receivables	the ratio of receivables to total assets
Foreign sales	the ratio of foreign sales to total assets
Growth in sales	the ratio of sales in current year to sales in previous year
Price to book ratio	the ratio of market value of a firm to its book value
ROA	return to total assets
Loss dummy	equal 1 if profit is negative, otherwise 0
Cross list dummy	equal 1 if firm is crosslisted in US, otherwise 0

Bibliography

Alex, B. and J. Adam (2011). Big four' accountants face audit reform. FT.

André, P., G. Broye, C. Pong, and A. Schatt (2013). Are joint audits associated with higher audit fees? *working paper*.

Arcidiacono, P. and P. B. Ellickson (2011). Practical methods for estimation of dynamic discrete choice models. *Annu. Rev. Econ. 3*(1), 363–394.

Aw, B., Y. and Y. Lee (2008). Firm heterogeneity and location choice of Taiwanese multinationals. *Journal of International Economics 75*, 167–179.

Barkema, H., J. Bell, and J. Pennings (1996). Foreign entry, culture barriers and learning. *Strategic Management Journal 17*, 151–166.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibirum. *Econometrica 63*(4), 841–890.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy 112*(1), 68–105.

Carr, D. L., J. R. Markusen, and K. E. Maskus (2001). Estimating the knowledge-capital model of the multinational enterprise. *The American Economic Review 91*(3), 693–708.

CFO UK (2012). British business and politicians are vehemently against joint audits.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies 47*(1), 225–238.

Chen, X. and M. Moore (2010). Location decision of heterogenous multinationals firms. *Journal of International Economics 80*, 188–199.

Chintagunta, P., E. Kyriazidou, and J. Perktold (2001). Panel data analysis of household brand choices. *Journal of Econometrics 103*(1), 111–153.

Craswell, A. T., J. R. Francis, and S. L. Taylor (1995). Auditor brand name reputations and industry specializations. *Journal of accounting and economics 20*(3), 297–322.

Crump, R. (2012). Competitive pressure, audit quality and industry specialization. AccountancyAge.

Das, S., M. J. Roberts, and J. R. Tybout (2007). Market entry costs, producer heterogeneity, and export dynamics. *Econometrica 75*(3), 837–873.

Datar, S. M., G. A. Feltham, and J. S. Hughes (1991). The role of audits and audit quality in valuing new issues. *Journal of accounting and Economics 14*(1), 3–49.

Davidson, W. (1980). The location of foreign direct investment activity: Country characteristics and experience effects. *Journal of International Business Studies 11*(2), 9–22.

Delios, A. and W. Henisz (2003). Political hazards, experience, and sequential entry strategies: The international expansion of Japanese firms, 1980¨c1998. *Strategic Management Journal 24*, 1153–1164.

Deng, M., T. Lu, D. Simunic, and M. Ye (2012). Do joint audits improve or impair audit quality? *working paper (University of Toronto).*

EC (2010). Audit policy: Lessons from the crisis. green paper. com (2010) 561 final. Brussels, Belgium: European Commission.

Economia (2012). PWC becomes top auditor among small-cap companies. Economia.

Ewert, R. and London Economics (2006). Study on the economic impact of auditors' liability regimes (markt/2005/24/f). final report to ec-dg internal market and services.

Francis, J. R. (2004). What do we know about audit quality? *The British accounting review 36*(4), 345–368.

Francis, J. R., P. N. Michas, and S. E. Seavey (2012). Does audit market concentration harm the quality of audited earnings? evidence from audit markets in 42 countries. *Contemporary Accounting Research*.

Francis, J. R., K. Reichelt, and D. Wang (2005). The pricing of national and city-specific reputations for industry expertise in the us audit market. *The Accounting Review 80*(1), 113–136.

Gerakos, J. J. and C. Syverson (2013). Competition in the audit market: Policy implications. National Bureau of Economic Research.

Glendinning, C. (2002). Partnerships between health and social services: developing a framework for evaluation. *Policy & Politics 30*(1), 115–127.

Government Accountability Office (2008). Audits of public companies: Continued concentration in audit market for large public companies does not call for immediate action.

Gowrisankaran, G. and M. Rysman (2012). Dynamics of consumer demand for new durable goods. *Journal of Political Economy 120*, 1173–1219.

Grossman, G. M., E. Helpman, and A. Szeidl (2006). Optimal integration strategies for the multinational firm. *Journal of International Economics 70*(1), 216–238.

Hay, D. C., W. R. Knechel, and N. Wong (2006). Audit fees: A meta-analysis of the effect of supply and demand attributes. *Contemporary Accounting Research 23*(1), 141–191.

Heckman, J. (1981). Heterogeneity and state dependence. *Studies in Labor Markets*, 91–139.

Helpman, E., M. Melitz, and S. R. Yeaple (2004). Export versus FDI with heterogeneous firms. *American Economic Review 93*(1), 300–316.

Hendel, I. and A. Nevo (2006). Measuring the implications of sales and consumer inventory behavior. *Econometrica 74*(6), 1637–1673.

Honoré, B. E. and E. Kyriazidou (2000). Panel data discrete choice models with lagged dependent variables. *Econometrica 68*(4), 839–874.

Hope, O.-K., J. C. Langli, and W. B. Thomas (2012). Agency conflicts and auditing in private firms. *Accounting, Organizations and Society*.

Johanson, J. and J. Vahlne (1977). The internationalization process of the firm-a model of knowledge development and increasing foreign market commitments. *Journal of International Business Studies 8*(1), 23–32.

Jones, H. (2012). Firms in eu face joint checks in auditing shake-up. REUTERS.

Jones, H. (2013). EU states in breakthrough over accounting reform. Reuters.

Jones, J. M. and J. T. Landwehr (1988). Removing heterogeneity bias from logit model estimation. *Marketing Science 7*(1), 41–59.

Knechel, W. R., L. Niemi, and S. Sundgren (2008). Determinants of auditor choice: Evidence from a small client market. *International Journal of Auditing 12*(1), 65–88.

Manski, C. F., D. McFadden, et al. (1981). *Structural analysis of discrete data with econometric applications*. MIT Press Cambridge, MA.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

McFadden, D. (1999). Computing willingness to pay in random utility models. *Trade, Theory and Econometrics: Essays in honour of John S. Chipman 15*, 253–274.

Morales, E., G. Sheu, and A. Zahler (2011). Gravity and extended gravity: Estimating a structural model of export entry. *working paper (Havard University)*.

Nosal, K. (2011). Estimating switching costs for medicare advantage plans. *working paper (University of Mannheim)*.

Numan, W. and M. Willekens (2012). Competitive pressure, audit quality and industry specialization. Working paper Katholieke Universiteit Leuven.

Petrin, A. and K. Train (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research 47*(1), 3–13.

Ratzinger-Sakel, N. V., S. Audousset-Coulier, J. Kettunen, and C. Lesage (2013). Joint audit: Issues and challenges for researchers and policy-makers. *Accounting in Europe 10*(2), 175–199.

Roberts, M. and J. Tybout (1997). The decision to export in Colombia: an empricial model of entry with sunk costs. *American Eonomic Review 87*, 545–564.

Shaver, J., W. Mitchell, and B. Yeung (1997). The effect of own-firm and other-firm experience on foreign direct investment survival in the United States, 1987-92. *Strategic Management Journal 18*(10), 811–824.

Shcherbakov, O. (2009). Measuring consumer switching costs in the television industry. *working paper (University of Mannheim)*.

Sikka, P. (2013). The race is on, but the field is limited in the auditing business. The Conversation.

Simunic, D. A. (1980). The pricing of audit services: Theory and evidence. *Journal of accounting research 18*(1), 161–190.

Terza, J. V., A. Basu, and P. J. Rathouz (2008). Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *Journal of health economics 27*(3), 531–543.

Tomiura, E. (2007). Foreign outsourcing, exporting, and FDI: A productivity comparison at the firm level. *Journal of International Economics 72*(1), 113–127.

UK Parliament Report (2011). Government response to report on auditors: Market concentration and their role.

Villas-Boas, J. M. and R. S. Winer (1999). Endogeneity in brand choice models. *Management Science 45*(10), 1324–1338.

Yeaple, S. R. (2009). Firm heterogeneity and the structure of U.S. multinational activity. *Journal of International Economics 78*, 206–215.

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig angefertigt und mich keiner anderen als der in ihr angegebenen Hilfsmittel bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Mannheim, den 28.04.2014

Curriculum Vitae

2014-	Assistant Professor, Wuhan University, China
2009-2014	PhD in Economics, Center for Doctoral Studies in Economics (CDSE) University of Mannheim, Germany
2010-2011	Visiting PhD Student, Yale University
2006-2009	M.A in Economics, Fudan University, China
2002-2006	B.A in Economics, Dalian Maritime University, China
2002	High School Diploma, Jiangyan No. 2 High School, China