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

Aiyong Zhu†

Abstract

When engaging in foreign direct investment (FDI) for the first time, multinational firms often incur large costs. They have 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 if they invest in the same location again. In this context, I develop a structural dynamic model of how firms make sequential decisions of where to invest. Using data on German firms’ FDI from 2002 to 2009, I estimate the model allowing for country-specific adaptation costs and firms’ heterogeneous preference towards location attributes. Estimation results suggest that adaptation costs are statistically and economically significant, ranging from 1.7% – 34.3% of the average expected discounted profit of a firm. If adaptation costs were subsidized completely, I find that firms’ FDI location choices would change dramatically. In addition, the average expected discounted profit would increase by 20.6%, not only due to the reduction of adaptation costs, but more importantly, due to a better matching between firms and locations.

JEL Codes: F14, F23, L23

Keywords: adaptation costs, sequential FDI, location choice

*I would like to thank helpful comments from Volker Nocke, Kathleen Nosal, Philipp Schmidt-Dengler, Alex Shcherbakov, Konrad Stahl, Yuya Takahashi, Naoki Wakamori, and Stefan Weiergraeb. In addition, I gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft and highly appreciate Deutsche Bundesbank for their hospitality as well as the access to the database.

†Department of Economics, Wuhan University. Email: aiyong.zhu@gmail.com
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.\(^1\) A vast literature on FDI, (Carr, Markusen, and Maskus (2001), etc) has empirically sought to explain what fundamental factors make host countries attractive to multinational corporations (henceforth MNCs). 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 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 need to occur if firms decide to invest in the same country again in the future.

Based on plant-level panel data on German MNCs’ outward FDI behavior,\(^3\) I find that there indeed exists a persistent pattern in firms’ FDI location choice. MNCs are more likely to keep investing in the same country where they have invested before. In other words, firms’ investment history matters. An obvious concern is that this

\(^{1}\)Data source: World Investment Report 2012.

\(^{2}\)Location and country are used interchangeably throughout this paper.

\(^{3}\)Theoretical studies often break down FDI into horizontal FDI and vertical FDI due to different incentives. Horizontal FDI is defined as an activity that MNCs 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. Since the bulk of FDI is horizontal rather than vertical in my dataset, I will focus on horizontal FDI in this paper. Details about the definition of FDI can be found in the next section.
observed persistence in firms’ location choices may not be due to reduction in adaptation costs 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).

At the firm level, I first establish that it is the presence of adaptation costs that generates the observed persistence in the data, after controlling for firm heterogeneity. I then quantify the magnitude of adaptation costs using the newly developed methodology from empirical industrial organization. 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. To incorporate 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. The empirical results show that adaptation costs are substantial and varying across locations, ranging from 1.7% to 34.3% of the average expected discounted profit. With these estimation results, I simulate several policy experiments involving FDI promotion schemes. These counterfactuals suggest firms would change location choice dramatically in response to the subsidization of adaptation costs, and additionally, 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 20.6%, on average over time if adaptation costs were completely subsidized. 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 short-run promotion policy will effectively attract long-run FDI due to state dependence, which may greatly boost local economic growth.

My paper complements the earlier literature on how MNCs expand globally (e.g.,

---

4“average” is referring to the mean value across all observed German MNCs in the data and over the entire sample period.
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 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).

The main contribution of this paper to the FDI literature is twofold. First, by using detailed plant-level panel data, I empirically validate the existence of adaptation costs, while allowing for firm heterogeneity in their sequential FDI location choices. Second, the newly developed methodology from industrial organization enables me 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, I can investigate the impact of a series of counterfactual FDI promotion policies on MNCs’ FDI pattern.

The remainder of the paper is organized as follows. First, 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.

2 Data

In this paper, I use the confidential Microdatabase Direct Investment (MIDI) provided by Deutsche Bundesbank (German Central Bank).\(^5\) 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 and individuals 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 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\(^6\) on the affiliate level from 1996 onwards, and since 2002, the informa-

\(^5\) According the confidentiality rule, I am not allowed to show any statistics aggregated less than three German firms.

\(^6\) 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.
tion about the attributes including total assets, turnover, number of employees, etc of reporting enterprises become available. Therefore, the general data description in the next subsection will be based on the sample from 1996 to 2009, while the data used for estimation only covers from 2002 to 2009 due to availability of parent firms’ attributes.

<table>
<thead>
<tr>
<th>Reporting year</th>
<th>Shares of parent firms</th>
<th>Total assets (affiliate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>≥ 20%</td>
<td>&gt; DM 1 million</td>
</tr>
<tr>
<td>1999</td>
<td>10%-50%</td>
<td>&gt; DM 10 million</td>
</tr>
<tr>
<td></td>
<td>≥ 50%</td>
<td>&gt; DM 1 million</td>
</tr>
<tr>
<td>2002</td>
<td>≥ 10%</td>
<td>&gt; Euro 3 million</td>
</tr>
</tbody>
</table>

### 2.1 Definition of 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 function derived 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 there is no additional information to distinguish the purpose of FDI made by manufacturing firms in the dataset, I can only restrict the sample to firms in non-manufacturing industry. Moreover, FDI in this paper is 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.\(^7\) In the remainder of this paper, “German firms” refers to German non-manufacturing firms for short and both are used interchangeably.

\(^7\)Due to data availability, I cannot 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.
2.2 Overview of FDI Pattern

With respect to the potential locations for FDI, I focus on most of the OECD countries plus China and Hong Kong, and group them into different regions, mainly based on the geographical proximity, similarity in culture and economic development. United States (USA), Canada (CAN) grouped as North America; France (FRA), Austria (AUT), Switzerland (CHE), Belgium (BEL), Ireland (IRL), Luxembourg (LUX) Netherlands (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. In spite of this relatively small choice set with only 24 locations, it covers around 81% of the locations chosen by German firms during the sample period.

Given the definition of FDI and choice set used in this paper, I can observe interesting patterns of sequential FDI location decisions in Table 2. The first column of Table 2 shows the overall top 10 locations among all German non-manufacturing firms. 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 year in the sample when firms are observed to engage in FDI for the first time. Top 10 locations are concentrated on developed countries and all but the USA are located within Europe. This observation is consistent with conventional theory that horizontal FDI is positively correlated with similarity in market size and factor 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. A large number of firms, about 60%, only invested once during the whole period 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

---

8I 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.

9There are several reasons that I restrict the choice set to be such a small sample. The main reason is due to data availability.
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: Top 10 locations for different groups of German firms 1996-2009

<table>
<thead>
<tr>
<th>Overall</th>
<th>Single time (t = t_0)</th>
<th>Multiple Times (t &gt; t_0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>USA</td>
<td>FRA</td>
</tr>
<tr>
<td>FRA</td>
<td>AUT</td>
<td>AUT</td>
</tr>
<tr>
<td>AUT</td>
<td>CHE</td>
<td>CHE</td>
</tr>
<tr>
<td>CHE</td>
<td>FRA</td>
<td>NLD</td>
</tr>
<tr>
<td>NLD</td>
<td>POL</td>
<td>GBR</td>
</tr>
<tr>
<td>GBR</td>
<td>NLD</td>
<td>USA</td>
</tr>
<tr>
<td>POL</td>
<td>GBR</td>
<td>ITA</td>
</tr>
<tr>
<td>CZE</td>
<td>CZE</td>
<td>POL</td>
</tr>
<tr>
<td>ITA</td>
<td>HUN</td>
<td>CZE</td>
</tr>
<tr>
<td>ESP</td>
<td>ESP</td>
<td>ESP</td>
</tr>
<tr>
<td></td>
<td>6,408</td>
<td>4,041</td>
</tr>
</tbody>
</table>

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

Two different conditional probabilities in Table 3 shed some light on how German non-manufacturing firms make sequential FDI location choices. Variable \(a_{it} = j\) denotes that firm \(i\) chooses location \(j\) for FDI in period \(t\). 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 \((Pr(a_{it} = j|a_{i0} = j, t > t_0, a_{it} \neq 0))\) represents 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 3, this conditional probability is significantly higher than another conditional probability, \((Pr(a_{it} = j|a_{i0} \neq j, t > t_0, a_{it} \neq 0))\), defined as the probability of choosing the same location, but conditional on firms having not chosen it in the first period \((t_0)\). The contrast between these two conditional probabilities clearly presents an interesting persistence pattern in sequential FDI location choices. In addition, the first conditional probability in less developed countries such as Poland, Czech and China is also very different (higher) in contrast to more developed nations. This difference suggests that the level of persistence seems to vary across countries. In general, location
attractiveness, unobserved firm heterogeneity and adaptation costs are three main sources that could generate the observed persistence pattern in the data. If the attractiveness of economic situation is the only incentive for firms to invest in the same location repeatedly over time, 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 3. Thus the difference indicates that location attractiveness is not the main source for observed persistence and the remaining two channels play a vital role. However it is very hard to disentangle adaptation costs from unobserved firm heterogeneity based on mere descriptive statistics.

Table 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.

### 2.3 Summary of Data

To deliver a clear picture of which sources drive this persistence in German firms sequential FDI decisions, I will provide the empirical evidence in the next subsection.

In the following empirical analysis, 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 lie in the following sectors: wholesale $^{10}$ 23.73%, household-related services 14.24%, real estate 7.92%, retail $^{11}$ 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

---

$^{10}$ The wholesale sector excludes motor vehicles and motorcycles.

$^{11}$ Motor vehicles, motorcycles repair of personal and household goods are excluded in this sector.
during 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. Moreover, I also drop firms that have ever invested in locations out of the predetermined choice set, i.e. 24 locations. Consequently, there are around 1700 firms left in the final sample for estimation.

Table 4 gives summary statistics of the choices made in the final sample. The first column of Table 4 presents the number of locations chosen by German firms every period. In the third column of Table 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>
<thead>
<tr>
<th>locations(#)</th>
<th>Percentage ($%$)</th>
<th>Timing</th>
<th>$t = t_0$</th>
<th>$t &gt; t_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>82.26</td>
<td></td>
<td></td>
<td>90.93</td>
</tr>
<tr>
<td>1</td>
<td>14.62</td>
<td>82.29</td>
<td>7.49</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.98</td>
<td>10.82</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>$\geq 3$</td>
<td>1.14</td>
<td>6.89</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>OBS(firm-year)</td>
<td>12,019</td>
<td>1,146</td>
<td>10, 873</td>
<td></td>
</tr>
</tbody>
</table>

2.4 Empirical evidence of adaptation costs

To control for unobserved firm heterogeneity as well as location attributes, reduced form regressions analysis can help us to discern which channels play a role in explaining the observed persistence pattern in sequential FDI location choices. When firm $i$
sets up a new subsidiary in location $j$ at period $t$, the latent profit (or attractiveness) $\pi^*_{ijt}$ can be approximated by the following equation:

$$\pi^*_{ijt} = F_{ij} + C_j + \beta_1 s_{ijt} + \sum_k \alpha_k x_{jkt} + \varepsilon_{ijt} \quad (1)$$

where $j = \{1, \ldots, 24\}$ denotes locations; $x_{jkt}$, location attributes: market size (measured by real gdp), GDP per capita (gdppc), economic growth (growth rate), labor cost, tax rate, unemployment rate and investment risk which is captured by corruption perception index (henceforth CI)\(^{14}\), Higher CI means less corruption, i.e. less investment risk in this country; $C_j$ denotes location fixed effect, and firm-location fixed effect (firm heterogeneity) is captured by $(F_{ij})$. Variable experience-country $s_{ijt} = \{0, 1\}$, 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. Profit shock $\varepsilon_{ijt}$ follows iid type 1 extreme value distribution across firms locations and time. 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,

$$P(a_{it} = j) = P(\pi^*_{ijt} > \pi^*_{ikt}, \forall k \neq j)$$

For the purpose of comparison, I also estimate several variant forms of equation (1). I replace experience-country ($s_{ijt}$) with the variable experience-region ($s_{irt}$) to see whether regional experience\(^{15}\) also works in the same direction as country specific experience. $s_{irt} = 1$ represents firm $i$ has invested in at least one country within region $r$ before period $t$; otherwise 0. In addition, I interact the CI with both regional and country specific experience. This interaction term is to capture whether investment risk (CI) has different impact on location choices for firms with distinct experience states.

Since the latent profit $\pi^*_{ijt}$ is not observable in the data, I will apply McFadden’s conditional logit model to estimate the parameters in equation (1). Note that the

---

\(^{14}\)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.

\(^{15}\)Region here follows the same definition as in the previous section.
regression results in Table 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 (adaptation costs) rather than heterogeneity matters to that location decision. Accordingly, it is worth mentioning how I control for the firm-location pair effect (firm heterogeneity) $F_{ij}$. The most straightforward way of controlling firm-country effect is to create the interaction between all firm dummies and all country dummies, however this will generate the typical incidental parameters problem in estimation. On the other hand, the coefficient of experience variable will be upward-biased if I completely ignore the fixed effect in the error term. As one way to mitigate the concern, I use the dummies of firm type in terms of sector and size to interact with country dummies to control for that firm-country pair effect. The size of firms measured by total assets are grouped into 3 types: small, medium and large; and the sector are distinguished into financial and non-financial industries among all the non-manufacturing firms. Thus there are 6 types of firms in all for the estimation. This classification implies I need to assume that the firm-country pair effect are constant for firms within each type, i.e., there is no other persistent firm heterogeneity within each type. The estimation results under this approach are presented from column S2 to column S4 in Table 5. Another way to fully control for the fixed effect $F_{ij}$ is to explore the panel structure as Chamberlain (1980) suggested. Chamberlain’s technique allows us to control for all the time invariant term in the conditional logit model, including observed and unobserved fixed effect. However, Honoré and Kyriazidou (2000) pointed out that the approach proposed by Chamberlain to control for firm heterogeneity requires strict exogeneity of the observed covariates. It is clearly violated in the current model that the lagged dependent variable (experience) is included in the covariates. Nevertheless, Chamberlain’s approach can still provide a lower bound for the coefficient ($\beta_1$) of experience variable (Chintagunta, Kyriazdou and Perktold (2001)).16 Thus the corresponding estimation results are listed from column S5-S7. An alternative way to obtain an consistent estimator for parameter $\beta_1$ is to impose additional assumption.

16Chintagunta, Kyriazdou and Perktold (2001) compare the method developed by Honoré and Kyriazidou (2000) with Chamberlain’s approach and find that the estimation results under Chamberlain’s technique are underestimated in both Monte Carlo simulation and real data application. The crucial point here is to show a significantly positive coefficient of firms’ country experience to be in line with the presence of adaptation costs. Therefore, a lower bound of positive $\beta_1$ is enough to support that adaptation costs indeed can account for observed persistence pattern in firms’ sequential location choice of FDI even after controlling for firm heterogeneity as well as location attributes.
about the distribution of $F_{ij}$, i.e. random effect approach, which is also widely used in the literature. The estimation results under random effect approach are presented in the appendix.

Table 5: Results of conditional logit model

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
</tr>
</thead>
<tbody>
<tr>
<td>gdp</td>
<td>1.0502**</td>
<td>0.9544*</td>
<td>1.0265*</td>
<td>1.1726**</td>
<td>6.5286***</td>
<td>6.6428***</td>
<td>6.5855***</td>
</tr>
<tr>
<td></td>
<td>(0.5163)</td>
<td>(0.5520)</td>
<td>(0.5375)</td>
<td>(0.6131)</td>
<td>(0.0083)</td>
<td>(0.0095)</td>
<td>(0.0093)</td>
</tr>
<tr>
<td>gdppc</td>
<td>-0.0172</td>
<td>-0.0007</td>
<td>0.0006</td>
<td>-0.0172</td>
<td>-0.0435***</td>
<td>-0.1084***</td>
<td>-0.0733***</td>
</tr>
<tr>
<td></td>
<td>(0.0460)</td>
<td>(0.0483)</td>
<td>(0.0467)</td>
<td>(0.0446)</td>
<td>(0.0057)</td>
<td>(0.0068)</td>
<td>(0.0065)</td>
</tr>
<tr>
<td>growth rate</td>
<td>0.0671*</td>
<td>0.0717**</td>
<td>0.0714**</td>
<td>0.0781**</td>
<td>-0.0147</td>
<td>-0.0222</td>
<td>-0.0444</td>
</tr>
<tr>
<td></td>
<td>(0.0343)</td>
<td>(0.0356)</td>
<td>(0.0356)</td>
<td>(0.0306)</td>
<td>(0.0343)</td>
<td>(0.0341)</td>
<td>(0.0338)</td>
</tr>
<tr>
<td>labor cost</td>
<td>0.0076</td>
<td>0.0071</td>
<td>0.0092</td>
<td>0.0160</td>
<td>0.0844***</td>
<td>0.0988***</td>
<td>0.0968***</td>
</tr>
<tr>
<td></td>
<td>(0.0171)</td>
<td>(0.0159)</td>
<td>(0.0157)</td>
<td>(0.0156)</td>
<td>(0.0124)</td>
<td>(0.0125)</td>
<td>(0.0097)</td>
</tr>
<tr>
<td>tax rate</td>
<td>-0.1653</td>
<td>-0.7569</td>
<td>-0.8039</td>
<td>-0.9520</td>
<td>-0.3956***</td>
<td>-0.4508***</td>
<td>-0.3067**</td>
</tr>
<tr>
<td></td>
<td>(1.521)</td>
<td>(1.5085)</td>
<td>(1.5497)</td>
<td>(1.4464)</td>
<td>(1.1148)</td>
<td>(0.4534)</td>
<td>(1.2780)</td>
</tr>
<tr>
<td>unemployment</td>
<td>0.0390***</td>
<td>0.0386***</td>
<td>0.0365***</td>
<td>0.0427***</td>
<td>0.0392***</td>
<td>0.0515***</td>
<td>0.0403***</td>
</tr>
<tr>
<td></td>
<td>(0.0117)</td>
<td>(0.0117)</td>
<td>(0.0118)</td>
<td>(0.0139)</td>
<td>(0.0067)</td>
<td>(0.0068)</td>
<td>(0.0065)</td>
</tr>
<tr>
<td>CI</td>
<td>0.0474</td>
<td>0.0492</td>
<td>0.1139*</td>
<td>0.1993***</td>
<td>-0.0754</td>
<td>-0.0615</td>
<td>-0.0771</td>
</tr>
<tr>
<td></td>
<td>(0.0638)</td>
<td>(0.0592)</td>
<td>(0.0618)</td>
<td>(0.0773)</td>
<td>(0.0438)</td>
<td>(0.0542)</td>
<td>(0.0413)</td>
</tr>
<tr>
<td>experience-country</td>
<td>2.4099***</td>
<td>2.3390***</td>
<td>4.0740***</td>
<td>3.153***</td>
<td>2.0644***</td>
<td>4.256***</td>
<td>1.3931***</td>
</tr>
<tr>
<td></td>
<td>(0.2823)</td>
<td>(0.2695)</td>
<td>(0.4642)</td>
<td>(0.1115)</td>
<td>(0.1856)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>experience-region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>experience-country*CI</td>
<td>-0.2460***</td>
<td>-0.2424***</td>
<td>(0.0420)</td>
<td>(0.0276)</td>
<td>(0.0408)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>experience-region*CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C FE*sector FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C FE*size FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Unobserved FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$-L$</td>
<td>-7543.68</td>
<td>-7534.02</td>
<td>-7399.13</td>
<td>-7655.64</td>
<td>-173.43</td>
<td>-173.37</td>
<td>-173.07</td>
</tr>
</tbody>
</table>

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

***, ** and * denote significance at the 1, 5 and 10 percent level, respectively.

As in Table 5, market size\(^{17}\) 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. Another important factor-firms’ previous experience from a specific country is also highly significant and has a positive effect as expected in all specifications. That is to say, conditional on firms making a new investment, firms are more likely to choose the same location for that FDI as where they have invested before. The variable experience-region represents whether firm has experience from any country within the same region, defined in previous section. The significant positive effect from regional experience also confirms early

\(^{17}\)Growth rate is one period lagged value and all other variables are in current value.
literature on the MNCs’ globalization pattern. Due to the similarity in culture and institution system as well as proximity in geography, the regional experience has the same effect as the operating experience from specific country within the same region.

The interaction term between experience and CI is significantly negative, implying that firms with prior operating experience is less sensitive to investment risk compared with firms who engage FDI in this country for the first time. In the meantime, the interaction term also corrects the sign of CI in specification S2-S5, even though not significant. The higher CI is, the less corruption there; thus firms should be more likely to invest there, leading to a positive sign of CI. Again regional experience is also a perfect substitute for specific country experience in this aspect. After controlling for additional unobserved firm-country fixed effect in specification S5-S7, both country specific and regional experience as well as gdp still keep the positive and significant effect, which has the same sign as the results in S1-S4. The significantly negative coefficients of the interaction between experience and CI also indicate that investment risk can be reduced by former presence in the country or that region. Besides the consistent sign, difference in parameter scales arise under two different approaches. The coefficient of experience is much smaller compared with those in S2-S4. This seems to suggest that there may exist unobserved heterogeneity persistent over time, which may confound with the experience effect if we fail to take it into account. But it could also be due to the underestimation bias by applying Chamberlain’s technic (Chintagunta, Kyriazdou and Perktold (2001)). Nevertheless, the positive effect of country specific and regional experience still keeps high significance. It is worth mentioning that all the significance is achieved after controlling for location fixed effect and firm-location pair effect (firm heterogeneity) under different approaches. This result 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.

Rather than use a dummy variable whether firm has operated in the 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 the country or the region before I observe another new FDI in the data. This new measure yields a very similar
and robust result as Table 5. 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 the elimination of adaptation costs.

3 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 necessary information that affects their profit. The information includes observed and unobserved (to researchers) location characteristics, which are exogenous and common to all firms. It also includes firms’ own FDI history involving which countries they have invested before as well as the number of subsidiaries in each country, and firms’ own 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. In addition, I also abstract from the decision to exit production and instead focus on the investment decision of location choice.

Firm $i$’s profit flow in period $t$, $\pi_{it}$ is given by:

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

where

$$
f_i(\tilde{X}_{jt}, n_{ijt}) = \sum_k (\tilde{\alpha}_k + \sum_r \tilde{\alpha}_{kr} z_{irt}) x_{jkt} + \xi_{jt} - \alpha_1 I_t - \alpha n_{ijt}
$$

In general, $\pi_{it}$ is composed of two parts: profit from the new subsidiary set up this period ($-\eta_j + f_i(\tilde{X}_{jt}, n_{ijt}) + \varepsilon_{ijt}$) and profit from existing subsidiaries ($\sum_{j=1}^{\bar{n}} n_{ijt} f_i(\tilde{X}_{jt}, n_{ijt})$). $\eta_j$ denotes adaptation costs in location $j$ and is assumed to be one kind of fixed cost constant over time but different across countries. $n_{ijt} = \{0, 1, \ldots, \bar{n}\}$ is the total number of affiliates belonging to firm $i$ at location $j$ before
period \(t\); \(\bar{n}\) denotes the finite upper bound for the total number of affiliates. Notation \(1_{(n_{ijt}=0)}\) 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 \(\eta_j\) 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 \(\{x_{j1t}, \ldots, x_{jkt}\}\), similar control variables as in equation (1); unobserved (to researchers) characteristics \(\xi_{jt}\) such as business operating costs, which are constant for all firms. Variable \(I_t\) is the interest rate in domestic country (Germany) to capture the common macroeconomic profit shock (or opportunity cost) for all firms. \(\varepsilon_{ijt}\) is the random profit shock across firms, locations, and time, following independent identical type 1 extreme value distribution. If firm \(i\) does not engage in FDI this period, she only earns profit from existing subsidiaries plus the iid profit shock \(\varepsilon_{i0t}\) with the same distribution as \(\varepsilon_{ijt}\).

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 \((z_{irt})\). Its parametric form shows how firm heterogeneity affects the profit given the same location attributes. More precisely, firms attributes also consist of observed and unobserved parts, \(z_{i1t}, z_{i2t}\), both of which interact with location attributes. The observed characteristics \(z_{i1}\) without subscript \(t\) is to model persistent firm heterogeneity and it is measured by firm’s size in terms of total assets; while \(z_{i2t}\) represents the unobserved heterogeneity, e.g. the productivity of the firm. The unobserved component not only differs across firms, but also varies over time which could be interpreted as time-varying productivity shock for every firm. I assume that there are two types of productivity shock, low and high. The productivity shock is assumed to follow a simple Bernoulli distribution (with probability \(\lambda\) to be less productive) iid across firms and over time. Given the simple iid distribution

---

18 It can also be interpreted as in the beginning of period \(t\).
19 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.
20 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.
21 In principle, I could extend the simple two types distribution to multiple types, even to the continuum types distribution. However this requires to assume a known distribution and increase the computation burden exceptionally in estimation for continuum case.
of productivity shock, parameter \( \lambda \) can be regarded as the average share of less productive firms over time. Another feature of this profit function is that it includes the term \( n_{ijt} \) and its coefficient \( \alpha \) 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 \( \alpha \) is expected to be positive, implying that firms are less likely to set up another new affiliate abroad 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:

\[
V_i(\vec{\varepsilon}_{it}, \vec{n}_{it}, \Omega_t) = \max_{a_{it} \in \{0, 1, \ldots, J\}} \pi_{it} + \beta E[V_i(\vec{\varepsilon}_{it+1}, \vec{n}_{it+1}, \Omega_{t+1}) | \Omega_t, \vec{n}_{it}]
\] (2)

Where state variables in firm \( i \)'s value function \( V_i(\vec{\varepsilon}_{it}, \vec{n}_{it}, \Omega_t) \) are the vector of profit 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 \( \Omega_t \) 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,\(^{22}\) 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.

### 3.1 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

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

\(^{22}\)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 \( \Omega_t \) including location attributes in all locations.
The key information about which countries a firm has invested before this period is well preserved by the vector \( \mathbf{s}_{it} \) 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 \( \mathbf{s}_{it} \) across locations. The estimated parameter \( \alpha \) will indicate how restrictive this assumption would be. That is if the estimation results show that \( \alpha \) is positive and 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 \( f_i(X_{jt}, n_{ijt}) = f_i(X_{jt}, N_{it}) \).

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

\[
\sum_{j=1}^{J} n_{ijt} f_i(X_{jt}, n_{ijt}) = N_{it} \psi_i(X_t, s_{it}, N_{it})
\]

\[
\psi_i(X_t, s_{it}, N_{it}) = \frac{\sum_{j \in \{j : s_{ijt} = 1\}} f_i(X_{jt}, N_{it})}{\sum_{j \in \{j : s_{ijt} = 1\}} 1}
\]

where \( \psi_i(X_t, s_{it}, N_{it}) \) 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 between \( N_{it} \) and \( \psi_i(X_t, s_{it}, N_{it}) \).

Given the new state variable \( s_{it}, N_{it} \), I define the location specific per-period profit flow \( \pi_{ijt} \) \( (j = 0, 1, \ldots, J) \) from the new subsidiary as follows:

\[
\pi_{ijt} = f_i(X_{jt}, N_{it}) - \eta_j 1(s_{ijt}=0) + \varepsilon_{ijt}
\]

(3)

\[
f_i(X_{jt}, N_{it}) = \bar{\pi}_{jt} + \sum_k \sum_r \tilde{\alpha}_{kr} z_{irt} x_{jkt} - \alpha N_{it}
\]

(4)

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

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 implicitly 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 \(\Omega_t\), I borrow the concept of the logit inclusive value \((\delta_{it})\) to collapse the high dimensional vector into a scalar.\(^{23}\) Thanks to the iid type I extreme value distribution of \(\varepsilon_{ijt}\), the inclusive value has the following closed form solution:

\[
\delta_{it}(\vec{s}_{it}, N_{it}, \Omega_t) = E \max\{\delta_{i1t} + \varepsilon_{i1t}, \ldots, \delta_{ijt} + \varepsilon_{ijt}\} = \ln \left( \sum_{j=1}^{J} \exp(\delta_{ijt}(\vec{s}_{it}, N_{it}, \Omega_t)) \right)
\]

where

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

\(\delta_{ijt}\) denotes the choice specific value function, i.e., the value firm \(i\) can obtain 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 \(\delta_{it}\) under certain additional assumption. The evolution of \(\delta_{it}\) can only capture the overall

\(^{23}\)The detailed discussion of logit inclusive value can be found in Nevo (2006) and Gowrisankaran and Rysman (2012).
pattern, that is to say, all else being 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 evolves better.\footnote{The main restrictive assumption 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.} 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 next period. In consequence, firms do not need to form expectation of each variable in information set \(\Omega_t\), but can simply focus on the scalar variable-the inclusive value. With this IVS assumption and all the reduction of state space, the value function can be equivalently written as:

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

where the expected value function \(E[V_i(\vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it})]\) is given by

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

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 stay the same.

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

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

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.\(^{25}\)

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

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

(7)

where the belief shock $\nu_{it}$ is assumed to follow a normal distribution $N(0, \sigma_{it}^2)$. The evolution of $\psi_{it}$ measures the change in the profit portfolio from existing subsidiaries. If a firm chooses the outside option in current period, then the change in $\psi_{it}$ 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 $\gamma_{1i}$. 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 $\gamma_{2i}$. 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 $\delta_{it}$ follows:

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

(8)

where I assume that the shock $\nu_{it}$ on belief also follows a normal distribution $N(0, \sigma_{2i}^2)$.

\(^{25}\)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.
I omit the superscript of $\nu_{it}$ to save notation, but the associated distribution still varies according to different types of action as in equation (8). 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 (8). If firm $i$ currently chooses the outside option, the evolution of $\delta_{it}$ only captures the information on exogenous changes, such as variation in location attributes as well as the associated continuation value. If firm $i$ chooses an old location to engage in FDI this period, the evolution of $\delta_{it}$ 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 $\delta_{it+1}$ should reflect the reduction of adaptation cost in location $j$, which implies the best choice next period should be higher compared with currently choosing old locations or outside option given everything else being equal. Therefore parameter $\rho_{1i}$ in equation (8) 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 (7), 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 $\delta_{0it}$ be:

$$\delta_{0it} = N_{it}\psi_{it} + \beta E[V_i(s_{it+1}, N_{it+1}, \psi_{it+1}, \delta_{it+1})|s_{it}, N_{it}, \psi_{it}, \delta_{it}]$$

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

$$P_{ijt}(s_{it}, N_{it}, \psi_{it}, \delta_{it}|z_{i1}, z_{i2}) = \int \mathbb{1}(\delta_{ijt} + \epsilon_{ijt} > \delta_{ikt} + \epsilon_{ikt}, \forall k \neq j, dF(\bar{\epsilon}_{it})$$

$$= \frac{\exp(\delta_{ijt})}{\exp(\delta_{0it}) + \exp(\delta_{it})}$$

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.
4 Estimation and Identification

4.1 The Estimator

The most important parameters to be estimated are the adaptation costs $\eta_j$ and the preference coefficients $\alpha_{ik}$, including the mean coefficients on the location attributes, $\bar{\alpha}_k$ as well as the random components varying with firms’ attributes, $\bar{\alpha}_{kr}$, and $\lambda$ governing the distribution of firms’ productivity. There are also nuisance parameters, such as, $\gamma_i$ and $\rho_i$ for firms’ belief on the evolution of inclusive logit value and the mean profit flow from existing subsidiaries respectively; $\sigma_{1i}$ and $\sigma_{2i}$ are the corresponding variance of belief shocks.

4.2 The Estimation Procedure

The estimation method in this paper closely follows Nosal (2011), Shcherbakov (2009) and Gowrisankaran and Rysman (2012), which involves three levels of optimization. The basic idea is to nest solving a dynamic programming (henceforth DP) problem inside the location share inversion of Berry, Levinsohn and Pakes (1995).

The inner loop solves firms’ DP problem in equation (6) 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.

4.2.1 Inner Loop

I set the annual discount factor $\beta$ equal to 0.9 in order to solve the DP problem in the inner loop. To obtain the fixed point for the Bellman equation (6), I need to discretize the continuous logit inclusive value ($\delta_{it}$) as well as $\psi_{it}$. The state space dimension for variable $\delta_{it}$ are divided into 30 grid points and 20 grids for $\psi_{it}$. $N_{it} = \{0, \ldots, 9\}$. Analogously, observed firm’s characteristics size $z_{i1}$ is discretized into 3

---

26 Allowing all of location attributes to interact with firms’ characteristics would provide a very flexible 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.

27 The discount factor $\beta$ 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.

28 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.
types-small size, medium size and large size based on the quantile from the data.\footnote{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.} 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 $64 \times 30 \times 20 \times 10$, 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 $\delta_{ijt}$, for all $j$ and the logit inclusive value $\delta_{it}$ that satisfy their recursive definitions respectively. Finally, it finds the firm’s belief parameters $\gamma_i$, $\rho_i$, $\sigma_i$ 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 $\delta_{it}$ and $\psi_{it}$, the integration along the dimension of these state variables is achieved by simulation.\footnote{The integration over state variable $\vec{\epsilon}_{it}$ 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 $\nu_{it}$ and $\upsilon_{it}$ to obtain the integration over state variable $\delta_{it}$ and $\psi_{it}$ respectively.} Once the expected value function has been computed, I can easily get $\delta_{ijt}$ for each location $j$ and then use it to update the logit inclusive value $\delta_{it}$. The $\delta_{it}$ are then regressed on the $\delta_{it-1}$ to obtain a new $\rho_i$, $\sigma_{2i}$ and to regress $\psi_{it}$ on $\psi_{it-1}$ to get new $\gamma_i$ and $\sigma_{1i}$. Since both $\delta_{it}$ and $\psi_{it}$ 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 to enter in each country and then pass it into the middle loop estimation.
4.2.2 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 $\chi_{jt}$. For the ease of computation, I divide the main parameters into two mutually exclusive sets of parameters, linear parameter vector $\Theta_1$ that only enters the mean profit flow $\bar{\pi}_{jt}$, and nonlinear parameters, $\Theta_2 = \{\alpha, \tilde{\alpha}_1, \tilde{\alpha}_2, \lambda, \eta_j, j = 1, \ldots, 24\}$;

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

where $\chi_{jt}$ denotes the predicted share of firm choosing location $j$ in period $t$ for FDI. It is a function of the average profit flow $\bar{\pi}_{jt}$ in all locations as well as parameter $\Theta$, 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.\(^{31}\) 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 $\Theta_1$ can be represented as a function of the nonlinear parameters $\Theta_2$.

4.2.3 Outer Loop

The outer loop search over the set of nonlinear parameters ($\Theta_2$) to maximize the likelihood function:

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

\(^{31}\)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 (Gowrisankaran and Rysman (2012)).
When convergence is reached in both middle and inner loop given $\Theta_2$, I can obtain the predicted probability $\hat{P}_{ijt}(\vec{s}_{it}, N_{it}, \psi_{it}, \delta_{it})$ after integrating out unobserved firm attributes $z_{i2t}$. To construct the objective function, I take every firm $i$’s state variable in the first period of the sample as exogenous given.\textsuperscript{32} 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.\textsuperscript{33}

### 4.3 Identification

Since the discount factor ($\beta$) is not identified in the class of dynamic discrete choice model, so I set $\beta = 0.9$, which is commonly used in the literature. 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 (Arcidiacono and Ellickson (2011)). The key to identify the adaptation cost $\eta_j$ 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 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, $\tilde{\alpha}_{kr}$, $r = 1, 2$ by exploring the variation in the the share of firms that have invested more than once in the same location.\textsuperscript{34} To be more precisely in this

\textsuperscript{32}There is no “initial condition problem” here, because firms’ unobserved heterogeneity is assumed to be an iid process over time.

\textsuperscript{33}I use the Nelder-Mead non-derivative “simplex” search method to get the relatively more robust results for nonlinear parameters. In addition, I restrict the parameters for adaptation costs to be non-negative in the optimization algorithm.

\textsuperscript{34}The two unobserved types of firms iid over time in this paper is a special case of general random coefficients in the discrete choice model, but the identification argument for other parameters form Berry, Levinsohn and Pakes (1995, 2004) still works here.
paper, the coefficient (\(\hat{\alpha}_{k_1}\)) 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. Parameter \(\lambda\) governing the distribution of unobserved productivity shock can be identified directly from the data. Since the unobserved productivity shock is assumed to be iid across firms and over time, then \(\lambda\) is equal to the average share of unproductive firms observed in the data. The productivity in this paper is defined in the same way by the parent firm’s sales as in Yeaple (2009);\(^\text{35}\) then \(\lambda\) is identified as the average share of firms below the mean productivity over time. Regarding the coefficient (\(\hat{\alpha}_{k_2}\)) for the interaction between unobserved firm attributes and location attributes, it can be identified by the magnitude of firms’ response to positive productivity shock. 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 firms over time helps to identify that parameter. For example, if the mean coefficients were simply zero, then less productive firms would with equal probability choose each location, however productive firm would always choose the location with largest market size over time.

5 Results and Counterfactuals

5.1 Estimated Parameters

The parameter vectors (\(\Theta_1\)) and (\(\Theta_2\)) are presented in Table (6).\(^\text{36}\) As we can see, variables gdp, gdp per capita, labor cost, interest rate and distance all have the expected effect. The positive coefficient on market size (gdp) is consistent with the incentives of horizontal FDI to serve a local market. Since Germany is a highly developed country, horizontal FDI made by German firms is likely to attractive to countries with similar economic endowment as in Germany. This effect is confirmed by significantly positive coefficient on GDP per capita. The significant negative effect from distance and labor cost is in line with conventional theory that firms prefer to investing in locations in the proximity and with low labor cost. Positive sign of

\(^{35}\)Because of the limited information about parent firms’ attributes in the data, I could only use sales to define the productivity. The calculation of direct measure-total factor productivity (TFP) requires additional information.

\(^{36}\)As mentioned in the estimation section, I only need to search \(\Theta_2\) in the maximum likelihood function. That’s why these parameters are distinguished into two groups in the table.
Table 6: Estimation results

<table>
<thead>
<tr>
<th></th>
<th>$\Theta_1$</th>
<th>$\Theta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>gdp</td>
<td>0.63</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>(0.5716)</td>
<td>(0.6903)</td>
</tr>
<tr>
<td>gdp per capita</td>
<td>0.13</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>(0.0248)</td>
<td>(0.2558)</td>
</tr>
<tr>
<td>language (German)</td>
<td>0.82</td>
<td>2.68</td>
</tr>
<tr>
<td></td>
<td>(0.0625)</td>
<td>(0.1762)</td>
</tr>
<tr>
<td>CI</td>
<td>-0.18</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>(0.3203)</td>
<td>(0.2745)</td>
</tr>
<tr>
<td>labor cost</td>
<td>-0.36</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>(0.1027)</td>
<td>(0.2545)</td>
</tr>
<tr>
<td>unemployment</td>
<td>-0.13</td>
<td>2.56</td>
</tr>
<tr>
<td></td>
<td>(0.0773)</td>
<td>(0.1814)</td>
</tr>
<tr>
<td>tax rate</td>
<td>-2.96</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>(2.2553)</td>
<td>(0.0515)</td>
</tr>
<tr>
<td>interest rate</td>
<td>-2.80</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(2.4383)</td>
<td>(0.0150)</td>
</tr>
<tr>
<td>distance</td>
<td>-3.45</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(1.0595)</td>
<td>(0.0021)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3273)</td>
<td></td>
</tr>
</tbody>
</table>

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

Language and negative effect of tax rate in the dynamic model confirms that firms like to choose countries speaking the same language and with low tax rate given everything else being equal. In addition, the negative coefficient of the interest rate in Germany indicates a high opportunity costs for German firms investing abroad. However, the coefficient of CI is unexpectedly negative because high investment risk (i.e., low CI) is supposed to be detrimental to FDI. This seems to suggest firms do not really concern about the corruption risk compared with the attraction of other characteristics.\(^{37}\)

The important parameters $\Theta_2$ include adaptation costs in each region as well as random coefficients. All estimated adaptation costs are positive and statistically significant, except in Western Europe.\(^{38}\) Northern Europe $\eta_{NE}$ has the highest adap-

---

\(^{37}\) Please recall that due to intensive computational burden, I don’t directly search the parameters in $\Theta_1$ via MLE; instead I use recovered $\bar{\pi}_{jt}$ in the middle loop optimization to estimate $\Theta_1$. 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.

\(^{38}\) Recall that country-specific adaptation costs within the same region are assumed to be the same for the ease of computation. This may partially explain the insignificance of adaptation costs in Western Europe. There are eight countries grouped in this region. Hence this large region
tation costs for German firms. Adaptation costs in Eastern Asia $\eta_{EA}$ are the second highest and its magnitude is very close to the top one, almost 20 times larger than the lowest costs in Western Europe $\eta_{WE}$. Costs in Southern Europe (SE) are ranked the third highest; while North America $\eta_{NA}$ and Eastern Europe $\eta_{EE}$ are similar to each other in terms of both ranking and magnitude of the costs from German firms’ perspective. 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 these adaptation costs ranging from 1.7% to 34.3% of the mean value of German firms’ expected discounted profit. The mean is defined as the arithmetic average value across all firms over the whole sample period in the data. As the ranking shown in Table (6), the lowest adaptation costs in Western Europe provide the lower bound for the range, and the upper bound is from the highest adaptation cost in Northern Europe.

With respect to the parameter $\alpha$, the significantly positive sign establishes 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, because this diminishing marginal return decreases the option value generated by previous operating experience in some countries. Coefficient $\tilde{\alpha}_1$ is significantly positive, indicating the presence of persistent heterogeneity that firms in large size can earn high profit through the channel of making better use of the given market size. Parameter $\tilde{\alpha}_2$ is also positive as expected, but unfortunately insignificant. This seems to suggest that the timing varying productivity heterogeneity may play a role in affect firms location choice even though the effect are not significant.

In general, all these firms engaging in FDI are are regarded more productive firms compared with those only serve domestic markets. However, the significance of $\tilde{\alpha}_1$ shows that the size difference still generates the non-ignorable heterogeneity among those most productive firms. The heterogeneity among German multinational firms is consistent with recent findings for US firms by by Yeaple (2009) and French firms by Chen and Moore (2010). Finally, parameter $\lambda$ indicates that on average about 13% firms are relatively more productive than others in the data.

may present relatively more variety in the adaptation process than any other regions. Another reasonable explanation would be that German firms actually do not feel any significant adaptation costs in Western Europe.
5.2 Model Fit

Figure 1: Model fit

Before conducting any counterfactual analysis, I present how these estimated parameters fit the data. I first randomly draw iid profit shocks \( (\varepsilon_{ijt}) \) 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 firms engaging FDI in each location every period and then take the average over time as in figure (1). In general, the estimated parameters fit the data quite well for all 24 locations. Thus, I will use the same estimated parameters for counterfactual analysis in the next subsection.

5.3 Counterfactual Analysis

Table 7 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
### Table 7: National regulatory changes 2002-2009

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

*Source: UNCTAD, Investment Policy Monitor database*

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 conduct 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 required 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 two 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 highest magnitude of reduction. Admittedly, it is true given everything else, like location attributes, being equal, but then I can compare it with 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. On one hand, contrary to the first scenario, the relative attractiveness of locations in terms of per-period profit flow does not change for firms that engage in FDI for the first time under the second scenario, because this policy just introduces a constant shift for all alternatives. But it still affect their location choice through the inter-temporal linkage by making future investment in the same location relatively less attractive, thus isolating the impact of
adaptation costs on the allocation of FDI for these firms. On the other hand, it does affect both the per-period profit flow and option value of investing in the same location for firms that already have operating experience under the second scenario. In both counterfactuals, firms with distinguished experience states could 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 calculate the benchmark expected discounted profit (henceforth EDP) for every firm based on each individual observation in the data. For each counterfactual, I then compute the EDP change that is essentially a compensating variation that would induce an equivalent change in the EDP, 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 EDP, the actual and counterfactual shares of firms entering each location are reported.

5.3.1 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., \( \eta_{j}^{\text{new}} = \frac{1}{2} \eta_{j} \) and \( \eta_{j}^{\text{new}} = 0 \), 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 adaptation costs (\( \eta_{j}^{\text{new}} \)) 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.

The actual and counterfactual aggregated share of firms investing in each region and country are reported in figure (2) and (3) respectively. Countries in all regions
Figure 2: Change of the share in regions if scaling down adaptation costs

Figure 3: Change of the share in countries if scaling down adaptation costs

(a) China and North America

(b) China and Northern Europe

Figure 4: Change of firms’ profit if scaling down adaptation costs
except in Western Europe (WE) would become more attractive than the real world if adaptation costs everywhere were scaled down (subsidized) to half or zero. For example, in the zero adaptation costs scenario, the share of firms entering Eastern Asia and Northern Europe regions multiplies about 6 times on average over the sample period. As expected, these two regions would benefit 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 the highest expansion there. Under this counterfactual with the zero adaptation cost everywhere, I observe 28.6% firms would change the location choice on average every period, which implies most of the variation in the counterfactual share comes from the new firm-location matching. What drives the matching process in this counterfactual entirely depends on location attributes. Since China has a huge market and enjoys a high economic growth for decades, it would attracts most firms investing there and become the second most preferred location after the USA for FDI in the absence of adaptation costs. On the contrary, the share of countries in Western Europe (WE) falls by almost half with zero adaptation costs because more attractive countries like China may provide a better match for those firms originally investing in this region. Figure (3b) shows the expansion pattern in the country level for regions with the largest adaptation costs. Even though adaptation costs in Northern Europe are slightly larger than in Easter Asia, the complete subsidization of costs would not make countries within Northern Europe more attractive than China. This figure essentially reinforces the role of matching channels in the counterfactuals that investment would be attracted to countries with large market size as well as other promising economic conditions.

With respect to the change of profits in this counterfactual, Figure (4) 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 it comes from the new matching between locations and firms. Moreover, as market size is 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 would be 7.9% on average over time if adaptation costs in all countries were scaled
down to half; while the increase would soar to 20.6% if all adaptation costs were completely eliminated. Figure (5) shows the decomposition of the changes in firms’ expected discounted profit under the counterfactual with zero adaptation costs. In the absence of adaptation costs, the first period’s profit flow on average across firms decreases 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 19.7% of the increase in expected discounted profit, while an much larger contribution is from the increment in the continuation value, 26.1%.

Figure 5: Decomposition of firms’ profit change with zero adaptation cost

Finally, Figure (6) 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 28.6% 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 (6), most of the switching patterns are driven by the first type that on average 19.4% of firms in each period would be motivated to start to invest again if adaptation costs were completely subsidized. Combined with Figure (5), 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 6: Location switching patterns with zero adaptation cost

\[ \eta_{\text{new}}^j = \eta_j - \eta_{WE}, \text{ for all } j. \]

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

Under this counterfactual, adaptation costs in every location are all cut by the lower bound, i.e., \( \eta_{\text{new}}^j = \eta_j - \eta_{WE}, \) for all \( j. \) In this setting, it is equivalent to normalize the adaptation costs in Western Europe to be 0, that is \( \eta_{WE}^{\text{new}} = 0. \)

Figure 7: Counterfactual: all adaptation costs are cut by the same amount

Figure (7) shows how firms respond to this policy scenario and the predicted aggregate share of firms investing in each region and switching pattern respectively. In this scenario, all regions would start to attracting more firms to their local markets as the average percentage change of firms engaging in FDI are all positive across all regions in Figure (7). However, there is still some variation in the percentage increase among different regions, while the difference of adaptation costs between regions are
the same as in the real world. Although Eastern Asia (EA) still has the second largest adaptation costs, a small magnitude of the costs reduction would enable it to enjoy the highest growth rate of attracting firms among all regions. Whereas, such a small cut would not make any difference in Northern Europe. This again suggests countries in Eastern Asia (EA) are relatively more attractive for investment between those two regions, in the sense that firms can find a better location for FDI to maximize the expected discounted profits there. Under this counterfactual, countries within Northern America (NA) and Western Europe (WE) are still more attractive than other ones based on the absolute share of firms investing in each region. This implies that location attributes in Eastern Asia (EA) are still not good enough to compensate its largest adaptation costs.

The right sub-figure in Figure (7) presents the location switching pattern for FDI if adaptation costs everywhere were cut by the same amount. The general share of firm switching locations is 22.7%, around 6% less than under the first counterfactual. The decomposition shows that the majority change would also be switching from not investing to investing, about 15.4% on average over time. The similar switching pattern seems to suggest that the inter-temporal consideration may be the driving force for location switching. In other words, firms are indeed forward looking to pursue the long run profit of investing in certain locations. The reduction of adaptation costs by the same amount makes the benefit of state dependence shrink. Thus, 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 would be 2.1% under this counterfactual.

6 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 $1.7\% - 34.3\%$ 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.
## 7 Appendix

Table 8: Results of mixed conditional logit model

<table>
<thead>
<tr>
<th>Mean</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
</tr>
</thead>
<tbody>
<tr>
<td>gdp</td>
<td>0.1198**</td>
<td>0.1556**</td>
<td>0.1603***</td>
<td>0.0710</td>
<td>1.0503**</td>
<td>1.0475**</td>
<td>1.0024**</td>
</tr>
<tr>
<td></td>
<td>(0.0504)</td>
<td>(0.0606)</td>
<td>(0.0597)</td>
<td>(0.1737)</td>
<td>(0.5163)</td>
<td>(0.5084)</td>
<td>(0.1904)</td>
</tr>
<tr>
<td></td>
<td>-0.0134</td>
<td>-0.0286**</td>
<td>-0.0305***</td>
<td>-0.0357**</td>
<td>-0.0172</td>
<td>-0.0155</td>
<td>-0.0147</td>
</tr>
<tr>
<td></td>
<td>(0.0998)</td>
<td>(0.0112)</td>
<td>(0.0112)</td>
<td>(0.0143)</td>
<td>(0.0460)</td>
<td>(0.0438)</td>
<td>(0.0384)</td>
</tr>
<tr>
<td>growth rate</td>
<td>0.0747***</td>
<td>0.0767***</td>
<td>0.0791***</td>
<td>0.0772***</td>
<td>0.0671*</td>
<td>0.0685**</td>
<td>0.0704**</td>
</tr>
<tr>
<td></td>
<td>(0.0204)</td>
<td>(0.0209)</td>
<td>(0.0219)</td>
<td>(0.0203)</td>
<td>(0.0343)</td>
<td>(0.0340)</td>
<td>(0.0292)</td>
</tr>
<tr>
<td>labor cost</td>
<td>-0.0130</td>
<td>-0.0048</td>
<td>-0.0034</td>
<td>-0.0016</td>
<td>0.0076</td>
<td>0.0095</td>
<td>0.0154</td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td>(0.0123)</td>
<td>(0.0120)</td>
<td>(0.0160)</td>
<td>(0.0168)</td>
<td>(0.0125)</td>
<td>(0.0161)</td>
</tr>
<tr>
<td>tax rate</td>
<td>-2.4000***</td>
<td>-3.0526***</td>
<td>-3.0149***</td>
<td>-3.2250***</td>
<td>-0.7699</td>
<td>-0.7080</td>
<td>-0.9004</td>
</tr>
<tr>
<td></td>
<td>(0.6530)</td>
<td>(0.7036)</td>
<td>(0.7002)</td>
<td>(0.9753)</td>
<td>(1.5121)</td>
<td>(1.5355)</td>
<td>(1.3508)</td>
</tr>
<tr>
<td>unempt</td>
<td>0.0298**</td>
<td>0.0328**</td>
<td>0.0304**</td>
<td>0.0291*</td>
<td>0.0390**</td>
<td>0.0371***</td>
<td>0.0426***</td>
</tr>
<tr>
<td></td>
<td>(0.0119)</td>
<td>(0.0120)</td>
<td>(0.0118)</td>
<td>(0.0151)</td>
<td>(0.0117)</td>
<td>(0.0119)</td>
<td>(0.0144)</td>
</tr>
<tr>
<td>CI</td>
<td>0.1070</td>
<td>0.1114</td>
<td>0.1748**</td>
<td>0.3287***</td>
<td>0.0474</td>
<td>0.1042</td>
<td>0.2017**</td>
</tr>
<tr>
<td></td>
<td>(0.0701)</td>
<td>(0.0710)</td>
<td>(0.0734)</td>
<td>(0.0841)</td>
<td>(0.0638)</td>
<td>(0.0673)</td>
<td>(0.0783)</td>
</tr>
<tr>
<td>experience-country</td>
<td>2.4282***</td>
<td>2.3869***</td>
<td>4.0643***</td>
<td>2.4099***</td>
<td>3.9562***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2860)</td>
<td>(0.2859)</td>
<td>(0.5213)</td>
<td>(0.2823)</td>
<td>(0.4583)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>experience-region</td>
<td>4.3907***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.2812***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.5484)</td>
</tr>
<tr>
<td>experience-country*CI</td>
<td>-0.2347***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.2206***</td>
</tr>
<tr>
<td></td>
<td>(0.0459)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0399)</td>
</tr>
<tr>
<td>experience-region*CI</td>
<td>-0.3023***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.2932***</td>
</tr>
<tr>
<td></td>
<td>(0.0486)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0483)</td>
</tr>
<tr>
<td>C FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C FE*sector FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>gdp</td>
<td>-7549.85</td>
<td>-7484.30</td>
<td>-7452.69</td>
<td>-7740.97</td>
<td>-7985.68</td>
<td>-7513.97</td>
<td>-7726.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Firm heterogeneity $F_{ij}$ in equation (1) is modeled to follow independent normal distribution, i.e., $F_{ij} \sim N(\mu_j, \sigma_j)$. The associated results are from column S1-S4. For robustness check, I alternatively assume the firm heterogeneity is captured by their ability of generating various profit given the same market size (gdp). Thus I estimate a random coefficient for market size in column S5-S7. In brief, the coefficient for experience in both country and region are significantly positive in all specifications. CI denotes corruption perceived index; C FE denotes country fixed effect. ***, ** and * denote significance at the 1%, 5% and 10%, respectively.
References


