DISCUSSION PAPER SERIES

No. 8536

ON THE GENESIS OF MULTINATIONAL FOREIGN AFFILIATE NETWORKS

Peter Egger, Matthias Fahn, Valeria Merlo and Georg Wamser

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP8536.asp**

ISSN 0265-8003

ON THE GENESIS OF MULTINATIONAL FOREIGN AFFILIATE NETWORKS

Peter Egger, ETH Zürich and CEPR
Matthias Fahn, Universität Konstanz
Valeria Merlo, ETH Zürich
Georg Wamser, ETH Zürich

Discussion Paper No. 8536
August 2011

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8536

August 2011

ABSTRACT

On the Genesis of Multinational Foreign Affiliate Networks*

Multinational enterprises (MNEs) develop their networks of foreign affiliates gradually over time. Instead of exploring all profitable opportunities immediately, they first establish themselves in their home countries and then enter new markets stepwise. We argue that this behavior is driven by uncertainty concerning a firm's success in new markets. After entry, the firm collects information which is used to update its beliefs about its performance at a market. As conditions in different markets are correlated, the information gathered in one of them can also be used to update beliefs elsewhere – with the degree of correlation depending on issues such as the geographical or cultural distance between markets. This correlated learning may render it optimal to enter markets sequentially – an investment in market A is only followed by entry in market B if the firm was sufficiently successful in A. The prediction that firms start their expansion in markets that are closer to their home base and then proceed step by step is supported by our empirical analysis, which features the universe of foreign affiliates held by German multinationals. Based on a rich set of benchmark estimates and sensitivity checks, we identify correlated learning across markets beyond alternative explanations as a key driver of gradualism in the genesis of multinational foreign affiliate networks.

JEL Classification: D83, D92, F23 and L23
Keywords: firm-level data, foreign affiliates, learning, location decisions and multinational firms

Peter Egger
ETH Zürich, KOF
Weinbergstrasse 35
8092 Zürich
SWITZERLAND

Email:
egger@kof.ethz.ch

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=154525

Matthias Fahn
University of Konstanz
PO Box 144
78457 Konstanz
GERMANY

Email:
matthias.fahn@uni-konstanz.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=174080

Valeria Merlo
Zürich, KOF
Weinbergstrasse 35, WEH E7
8092 Zürich,
SWITZERLAND

Email: merlo@kof.ethz.ch

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=174081

Georg Wamser
ETH Zurich
KOF Weinbergstrasse 35
WEH E8
8092 Zurich
SWITZERLAND

Email: wamser@kof.ethz.ch

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=173719

* We thank Ralph Ossa, Mark Satterthwaite, Robert Ulbricht, Mauricio Varela, participants at research seminars in Konstanz and Munich, as well as participants at the 8th Bundesbank MiDi workshop for helpful comments. We gratefully acknowledge support by members of the research center of the German central bank.

Submitted 05 August 2011

1 Introduction

Multinational enterprises (MNEs) tend to pursue a *gradual expansion strategy* of their network of foreign affiliates over time rather than exploring all profitable opportunities *simultaneously*. They typically establish themselves in their home countries and then enter new foreign markets step by step. This paper studies the optimal dynamic behavior of MNEs to explore international growth opportunities. It contributes to the literature on the international organization of firms by investigating sequential versus simultaneous or isolated location decisions.

We propose a model where MNEs face uncertainty concerning their success in new markets and learn about that after entry. Conditions in different markets are not independent, and the information gathered in one country can also be used to learn about conditions in other, in particular similar countries. This so-called correlated learning can explain why many firms expand step by step: market entry is associated with considerable costs, and sequential investments help to economize on these costs by reducing uncertainty. The learning model developed in this paper serves to derive a number of testable hypotheses regarding market entry in general and *simultaneous* versus *sequential* market entry in specific. These hypotheses are assessed in a data-set of the universe of German MNEs and their foreign affiliates. The results provide empirical evidence for correlated learning as a main driver behind international expansion strategies.

Our paper is related to recent work on sequential exporting. For instance, Evenett and Venables (2002) point out that initial exports to one market are typically followed by exports to adjacent markets at the product level. Eaton, Eslava, Kugler, and Tybout (2007) find that Columbian firms start exporting in a single foreign market and gradually enter additional destinations. They also show that further expansions crucially depend on the export market served initially. Using Russian firm-level data, Schmeiser (2009) identifies a similar pattern and demonstrates that export experience determines export dynamics: a typical firm first enters one destination and then slowly expands. More recently, Albornoz, Calvo Pardo, Corcos, and Ornelas (2011) explore how firms learn about their export profitability. They illustrate that firms use their first export market as a "testing ground" to learn about their export profitability and, subsequently, exit, continue to export, or enter further markets. Hence, the first export decision not only provides information about the export market, it also reveals information about the firm itself (in a given market). We argue that learning is particularly crucial for foreign direct investment (FDI; as an alternative to exporting) which, unlike exporting, definitely involves discrete real investments.

Our study also relates to the literature on the mode and depth of firms' international activities. Models of heterogenous firms describe how enterprises make decisions depending on the associated costs and their productivity levels. Assuming that fixed costs are higher for exporting than for domestic sales only, and that they are even higher for foreign plant set-up and running a multinational network than for exporting, the most productive firms engage in FDI, less productive companies export, and the least productive firms stay in the domestic market only (see Helpman, Melitz, and Yeaple, 2004). This theoretically predicted pattern has been supported by a number of empirical studies. Recently, Conconi, Sapir, and Zanardi (2010) have illustrated that learning through exporting matters for the decision of how to serve a market, via exports or FDI. Empirically, many MNEs are multi-plant units which are established gradually. It appears that no research on the *genesis of multinational foreign* *affiliate networks* exists to this date, and it is this paper's purpose to fill this gap.

Given that establishing a multinational network of foreign affiliates is profitable per se beyond other options, further choices are available to the firm. For example, it has to decide on *where* to locate the first foreign entity (location choice). This choice among several alternative first locations may depend on local factor costs, on the accessibility of production factors, or on various measures of proximity to the home market (for empirical investigations on the location choice of MNEs, see Devereux and Griffith, 1998; Becker, Egger, and Merlo, 2009; Chen and Moore, 2010). Managers of the firm then have to answer related questions of the following kind. Should the first investment involve high or low capacity levels? Is the first investment the basis for other investments in the region? Given the location choice of previous investments, where should subsequent affiliates be located?

In this context, the present study analyzes foreign location decisions of MNEs, why sequential entry patterns can be optimal, and how decisions depend on earlier location choices. Our theoretical approach is related to the theoretical learning (or bandit) literature. Early contributions to this literature include Bellman (1956) and Berry and Fristedt (1979), while a learning process similar to ours has been applied recently by Bergemann and Hege (1998, 2005) and Keller, Rady, and Cripps (2005). Specific to our model is the possibility that entry decisions in different countries depend on each other, since market conditions exhibit similarities. How consumer preferences or attitudes of employees differ across countries depends on issues such as geographical or cultural distance. If the correlation between market features is sufficiently high, a firm can make use of the knowledge it gains in one market to learn about conditions elsewhere. Then, a firm may want to enter a second country if it was sufficiently successful in the first one. This leads to one of our main results: even if expected profits in a market are positive, it can be optimal to delay or later on even abandon subsequent entry. The reason is that market entry is costly, and sequential investments can increase expected profits by using information gathered elsewhere. On the other hand, the reduced uncertainty through delayed entry comes at the cost of foregone profits. This result is related to the vast literature on investment under uncertainty (starting with McDonald and Siegel, 1986; see Dixit and Pindyck, 1994, or Carruth, Dickerson, and Henley, 2000, for overviews). If the value of an irreversible investment project follows a stochastic process, the option to wait for a better realization is valuable even if immediate entry would be profitable. Our result follows a similar logic. Uncertainty combined with correlated learning creates an option value of waiting, and a sufficient amount of uncertainty must exist to make sequential entry potentially optimal. The main difference is that learning is not induced exogenously but by a firm's activities elsewhere. Thus, the firm can influence the degree of learning by adjusting its investment levels in other markets.

Based on the proposed theoretical model, we derive several testable hypotheses. First, entry should be more likely in foreign markets where expected profits are higher. Expected profits do not only increase with market size and productivity but also with proximity to the domestic market. The reason for the latter is the following. Firms for which FDI is possibly attractive are successful at home. Such firms will first enter closer foreign markets, since their positive experience at home is ceteris paribus more valuable there. Second, sequential entry rather than simultaneous entry abroad can be optimal with sufficient uncertainty about returns on FDI and high-enough success at markets entered first. Then, with sequential entry, more proximate countries should be entered first on average. The reason is that uncertainty creates a value of waiting, rendering it worthwhile to stagger FDI decisions across markets in an order which declines in expected profits. Third, subsequent foreign entry is more likely in markets which are proximate to previous investments for the same correlated learning reasons as before.

We assess these hypotheses empirically using a unique micro-level panel data-set provided by Deutsche Bundesbank (the German Central Bank) that allows us to track German MNEs' sequential location decisions over time. We are able to identify the first, the second, etc., up to the eleventh location decision of firms across large-enough samples. Using a conditional logit model for the empirical analysis, we find that first foreign entry is more likely for countries that are closer to the MNE's home base and where higher profits may be expected in general terms. This finding is supported by variables measuring the proximity of markets at large, e.g., whether the same language is spoken or if the target country used to be a colony of the home country. Second, proximate countries tend to be entered first as a multinational network evolves. And third, subsequent entry in later expansion phases is generally more likely in markets that are closer to the ones entered previously.

To analyze whether it is actually correlated learning that drives the observed expansion patterns or not, we conduct a number of tests. An important result of our theoretical model is that the average capacity of investments should ceteris paribus be higher in case of a sequential entry pattern compared to the average capacity of investments when entities are established simultaneously. The reason is that in the former case a higher capacity in a country not only raises expected revenues there, but also the amount of learning about other markets.

Furthermore, the reason for not observing sequential investments could either be that *isolated entry* or *simultaneous entry* was intended from the beginning, or that a firm initially planned *sequential entry* but was not sufficiently successful in the first market. We can use the result that the average capacity in sequentially entered markets is higher than in simultaneously entered ones. This allows us to hypothesize that, if a firm only enters one market in one phase and does not establish any subsequent affiliates, the more successful ones should have a lower capacity than the others. The latter is consistent with the notion that affiliates with an above-average capacity were intended to be followed by sequential investments elsewhere. Then, one reason for a lack of subsequent investments to high-capacity first investments should be that first investments were not sufficiently successful.

Finally, correlated learning makes the firm ceteris paribus more optimistic about the prospects in a market. Thus, it will lower its requirements for later entry with respect to market size or entry costs over time.

We find support for all of these hypotheses in our empirical analysis, leading us to the conclusion that the proposed correlated learning mechanism is indeed an important factor determining international expansion strategies of multinational foreign affiliate networks. To investigate whether other alternative mechanisms can explain the results of our model as reported above, we analyze alternative mechanisms such as stochastic shocks across markets, diseconomies of scale (i.e., constrained resources available to firms), or learning by doing. All of these mechanisms may be used to derive dynamic expansion strategies of MNEs. But, as we illustrate, none of these models fits the data as well as correlated learning does.

The remainder of the paper is organized as follows. We present a theoretical model and main results in Sections 2 and 3. Section 4 derives testable hypotheses, introduces the data and empirical model, and summarizes the benchmark estimates. Section 5 provides extensions and robustness tests, while Section 6 develops alternative models that might also explain the observed firm behavior. Section 7 offers concluding remarks.

2 Model Setup

The following model portrays the international expansion pattern of a firm. This firm (or "multinational", or "MNE") is active in two periods, $t = 1, 2$, and considers establishing affiliates in two countries, $j = \{A, B\}$. Entry in country j is possible in either period. Upon entry, the firm chooses a capacity level X_j which can not be adjusted subsequently.[1] In the period of entry, the investment level X_j is associated with costs $K_j(X_j) = F_j + k_j \frac{(X_j)^2}{2}$, where $F_j \geq 0$ are fixed entry costs and $k_j > 0$ captures marginal investment costs.

Each investment may be profitable or not. More precisely, the firm possesses an exogenously given type θ_j in country j, with $\theta_j \in \{0, \theta^h\}$ and $0 < \theta^h \leq 1$. The type θ_j covers firm- as well as market-specific characteristics and is related to the idea of a matching quality between the firm and the market in the spirit of Jovanovic (1979).[2] If $\theta_j = \theta^h$, the affiliate generates a constant return $R_j > 0$ with probability $X_j\theta^h$ in each period.[3] Future profits are discounted with the factor $\delta \in [0, 1]$. If $\theta_j = 0$, the project does not yield any profits. Formally, per-period returns are denoted by $Y_{jt} \in \{0, R_j\}$, with $\text{Prob}(Y_{jt} = R_j \mid X_j, \theta_j) = X_j\theta_j \in [0, 1]$. The latter requires sufficiently high marginal investment costs, $k_j \geq \theta^h R_j(1 + \delta)$, $j = A, B$, which we assume subsequently. Finally, the firm is not financially constrained.

Note that, for the sake of simplicity, we restrict the firm's success to two states – an affiliate is either profitable or not. Allowing for several or even continuous degrees of success would not have any qualitative impact on our results but substantially complicate the analysis.

2.1 A Firm's Belief About Its Type

One crucial element of the proposed model is that, before market entry, an MNE does not know whether its type in country j is high or low. Instead, it assigns the (subjective) probability ρ_j to being the high type. In the following, we call this probability the firm's *belief*. The belief is given before period $t = 1$ and may have been formed by previous activities in this market such as market research.

For a firm in our model, there are three relevant markets in each of which the firm is either of the high or the low type: home, A, and B. All firms are initially active at home and learn about their type there over time. We confine our interest to firms of the high type at home. This is consistent with results in Helpman, Melitz, and Yeaple (2004) suggesting that only the most productive firms

[1] This assumption has no qualitative impact on our results; see Appendix III for a characterization of capacity investments when X_j can be adjusted later on.

[2] For instance, θ_j captures the success of a marketing campaign and other specific characteristics of demand for a firm's products in country j. It could also reflect the ability to make use of natural resources and other local factors in country j and, in general, the efficiency of the firm's production process there.

[3] We interpret R_j quite broadly. It may include revenues attributable to the investment in j but also general efficiency gains to the firm through a foreign investment. Furthermore, we choose the probability function to be linear in X_j for convenience. Generally, any probability function which is monotonic in X_j and less convex than the cost function $K_j(X_j)$ would serve our purpose.

become MNEs. While decisions about home do not feature in our analysis, the type there is relevant because conditions in different markets – and thus the realizations of types – are not independent. For example, geographical or cultural neighborhood across markets is a source of such correlation. A firm's type is specific to a market so that being of the high type at home does not guarantee being of the high type also in a foreign market.[4]

In our model, the type at home is drawn first. Recall that we focus on firms of θ^h at home. Then, the type in country A is realized. The type is determined by two different components. The first component relates to the type at home, the second one is idiosyncratic. To be precise, with weight $r_A \in [0,1]$ the type in A is high with the same probability – namely unity – as at home. r_A captures the proximity between home and country A and is larger if markets are geographically and culturally close to each other. With weight $(1-r_A)$, the type in A is high with an idiosyncratic probability of ρ_A^0. The latter is formed by generally available information, market research, or other previous activities. Thus, the firm's subjective ex-ante belief of being a high type in country A equals

$$\rho_A = r_A + (1 - r_A)\rho_A^0.$$

Finally, the type in country B is realized. It is identical to home with weight r_B. With weight $(1-r_B)$, the type is determined by B's idiosyncratic component and its proximity to country A. Formally, the belief in B is characterized by

$$\rho_B = r_B + (1 - r_B)[r_{AB}\rho_A^0 + (1 - r_{AB})\rho_B^0].$$

The parameter r_{AB} captures potential correlations between A and B that are not already covered by the proximity to home, r_A and r_B, respectively. Hence, we introduce different dimensions of proximity. For example, assume that home is Germany and that the MNE considers investments in Austria, Switzerland, and Denmark. The (geographical and cultural) distance of each of these countries to Germany is quite low. However, while Austria and Switzerland share a common language and other cultural aspects with each other so that they are quite proximate in general terms (high r_{AB}), the geographical and cultural distance between Austria and Denmark is much bigger (lower r_{AB}) than the one between Austria and Switzerland.

The beliefs are increasing in the respective proximity parameters r_j, so that the firm is more optimistic about a country closer to home. The parameters ρ_j^0, r_j and r_{AB} are known ex ante and determine the subjective beliefs ρ_j. After the first period the respective output values are observed and these observations are used to update beliefs using Bayes' rule. We will explore the updating process in more detail below when analyzing respective entry patterns.

[4] Prior to foreign market entry, a firm faces substantial uncertainty concerning its profitability there. It may be argued that firms considering FDI in some market have already gathered information about local market conditions. This does not contradict the maintained assumptions, as long as there is still some uncertainty left.

3 Optimal Behavior

3.1 Entry Patterns

Conditional on being active abroad, the MNE will choose one of the three following options.

- **Isolated Entry**: Entry into one country in period 1, no further entry in period 2.
- **Simultaneous Entry**: Entry into both countries in period 1.
- **Sequential Entry**: Entry into one country in period 1 and into the other one in period 2, conditional on a success of the first foreign investment.

All other possibilities are dominated by one of the options mentioned above. Isolated or simultaneous entry in the second period would come at the cost of gone profits in the first period. Entering a second market sequentially after a failure in the first one would be dominated ex ante by simultaneous entry. The reason for the latter lies in the correlated updating of beliefs, as will become clear below.

The MNE will choose isolated entry into one country if beliefs in the other country are too low to ever justify an investment there. Otherwise, the firm will consider simultaneous or sequential entry, facing the following tradeoff. Under sequential entry, the firm loses potential profits in the first period. On the other hand, the risk of wasting investment costs is reduced, i.e., there is a value of waiting. The reason is that the firm learns something about the conditions in the second market because of the correlation of the firm's types in the two markets. We will analyze this tradeoff in more detail below. The proofs of the subsequently stated Propositions 1-3 can be found in Appendix I.

3.2 Isolated Entry

In this section, the MNE only considers entry into one country, which allows us to omit the country subscript. If entry is optimal, it will occur in period 1. Then, the firm chooses a capacity X to maximize the expected discounted profit stream

$$\Pi^{iso} = X\left(\rho_1 R\theta^h + \delta \mathrm{E}[\rho_2] R\theta^h\right) - K(X), \tag{1}$$

where expectation is taken concerning ρ_2, the belief in period 2. To be able to characterize the optimal level of X, we have to consider the updating process. ρ_2 is derived using Bayes' rule, given the initial belief and capacity X. To simplify issues, we omit time subscripts and denote the initial belief by ρ. After a success ($Y_1 = R$), the period-2 belief equals ρ^+, while after a failure ($Y_1 = 0$), that belief equals ρ^-. As the bad type always fails, a success immediately reveals a good type, and

$$\rho^+ = 1.$$

After a failure, we have[5]

$$\rho(X)^- = \frac{(1 - X\theta^h)\rho}{1 - \rho X\theta^h}.$$

[5] Recall that $\rho(X)^- = \mathrm{Prob}[\theta = \theta^h \mid Y_1 = 0] = \frac{\mathrm{Prob}[\theta=\theta^h \cap Y_1=0]}{\mathrm{Prob}[Y_1=0]}$
$= \frac{\mathrm{Prob}[Y_1=0|\theta=\theta^h]\mathrm{Prob}[\theta=\theta^h]}{\mathrm{Prob}[Y_1=0|\theta=\theta^h]\mathrm{Prob}[\theta=\theta^h]+\mathrm{Prob}[Y_1=0|\theta=0]\mathrm{Prob}[\theta=0]} = \frac{(1-X\theta^h)\rho}{(1-X\theta^h)\rho+1(1-\rho)}.$

As $\frac{\partial \rho(X)^-}{\partial X} < 0$, a larger investment is generally associated with more learning. More updating occurs for intermediate values of the initial belief, while there is less updating if the belief is close to zero or unity.[6]

As running the affiliate requires no further costs once the capacity is set, it is never optimal to exit a market, no matter how low the belief may be.[7] The assumption of no operating costs has no substantial impact on our results. In a more general model with operating costs and a longer time horizon, the possibility of exit in the future would affect the decisions associated with entry. However, this impact is lower the further ahead a potential exit lies in the future. In our analysis, firms are sufficiently productive to engage in FDI and face substantial entry costs. This implies that the belief necessary to make entry optimal is high enough to make an immediate exit after a few failures very unlikely. The optimal activities of the MNE in case of isolated entry may be described as follows.

Proposition 1 (Isolated Entry): Given market entry, the optimal capacity under isolated entry equals

$$X^{iso} = \frac{\rho\theta^h R(1+\delta)}{k}, \tag{2}$$

and entry is only optimal if

$$\Pi^{iso} = \frac{(\rho\theta^h R)^2(1+\delta)^2}{2k} - F \geq 0. \tag{3}$$

Throughout, we assume that the condition in (3) is satisfied for each country if $\rho_j = 1$, i.e., market entry is profitable at profit-maximizing capacity levels for sufficiently high beliefs. Comparative statics can easily be derived. The capacity is increasing in R (which could reflect fundamentals such as market size). It is decreasing in the distance to home (i.e., a larger value of r) and in investment costs k. Higher R and r as well as lower k also render entry more likely, whereas larger fixed costs (which in a market perspective could reflect fundamentals such as corruption or investment freedom) make entry less likely.

3.3 Simultaneous Entry

When choosing simultaneous entry, the firm enters both countries A and B at the beginning of period 1. Now, beliefs in country A are also affected by outcomes in B (and vice versa). But, after capacities are set, events in A do not have an impact on decisions in B and vice versa. Thus, correlated learning does not provide an additional benefit under simultaneous entry, and we postpone the analysis of the correlated updating process to the case of sequential entry. Total expected profits of the firm just equal the sum of profits in each country under isolated entry:

$$\Pi^{sim} = \left(X_A^{sim}\rho_A\theta^h R_A + X_B^{sim}\rho_B\theta^h R_B\right)(1+\delta) - K_A(X_A^{sim}) - K_B(X_B^{sim}).$$

[6] This is the case since $\rho - \rho(X)^- = \rho X\theta^h \frac{1-\rho}{1-\rho X\theta^h}$.

[7] Thus, we do not consider the "standard" value of learning, namely the option to stop the project. If running the affiliate was costly, this option value would make the firm willing to accept some expected short-term losses in the first period.

Therefore, the chosen capacity levels are identical to above and we get

$$X_j^{sim} = \frac{\rho_j \theta^h R_j (1+\delta)}{k_j},$$

yielding total profits

$$\Pi^{sim} = \frac{(\rho_A \theta^h R_A)^2 (1+\delta)^2}{2k_A} + \frac{(\rho_B \theta^h R_B)^2 (1+\delta)^2}{2k_B} - F_A - F_B.$$

The non-negativity condition is identical to the one under isolated entry and has to be satisfied here as well. Comparative statics for capacity and the likelihood of entry are also the same as under isolated entry.

3.4 Sequential Entry

When choosing sequential entry, the firm uses information gathered in one country, say A, to update its beliefs about B. At the beginning of period 1, it enters A. Observing a success, it subsequently invests in B in period 2. Otherwise, it just remains in A without any further investments. Note that entry in B after a failure in A can not be optimal since this would be dominated by simultaneous entry.

The (relative) profitability of sequential entry depends on several aspects. As already mentioned, the firm faces a tradeoff when comparing sequential and simultaneous entry. Under the former regime, it can reduce its risk and only has to bear investment costs for relatively high beliefs. On the other hand, it loses potential profits from the second country in period 1. Crucial for the aspect of risk reduction is the actual amount of correlated learning, which determines the option value of waiting. This depends on the distance between A and B, captured by the parameter r_{AB}. Furthermore, observing a success in A has to be a sufficiently strong signal. In case that r_A is very close to unity, a success in A does not contain much new information, as the firm is already quite optimistic to face a high type there ex ante. This limits updating in B, rendering the gains of sequential entry negligible. Therefore, a considerable amount of uncertainty in A has to prevail for sequential entry to be optimal.

3.4.1 Beliefs and Correlated Learning

Considering correlated learning, the updating process is slightly different from above, as the outcome in one country also affects beliefs in the other one.

Recall that ex-ante beliefs (or *priors*) about markets A and B equal $\rho_A = r_A + (1-r_A)\rho_A^0$ and $\rho_B = r_B + (1-r_B)[r_{AB}\rho_A^0 + (1-r_{AB})\rho_B^0]$, respectively, where r_j is a proxy for the (cultural or geographical) distance of country $j \in \{A, B\}$ to the MNE's home market, while r_{AB} captures the proximity between the two foreign target countries. With sequential entry, we can not analyze both countries' beliefs in isolation anymore and have to consider four possible states for the set of types (θ_A, θ_B), namely (θ^h, θ^h), $(\theta^h, 0)$, $(0, \theta^h)$, and $(0, 0)$. Updating occurs conditional on observing the outcome (Y_A, Y_B), which takes one of the realizations (R_A, R_B), $(R_A, 0)$, $(0, R_B)$, or $(0, 0)$. As before, we use superscripts to denote updated beliefs. For instance, ρ_j^{+-} denotes the period-2 belief in country j after a success in A and a failure in B were observed. Note that not entering B in period 1 automatically implies that a failure was observed there. In Appendix II, we derive general characterizations of the updated beliefs.[8] Note that, after a success, the belief in the respective country still jumps to unity. What we need for the analysis of sequential entry – i.e., entry in the second country after a success in the first one – is the updated belief in B after a success was observed in A, and vice versa.

Under sequential entry starting in A and with $Y_A = R_A$ (which implies that the type in A must be high), the belief in B becomes

$$\rho_B^{+-} = \left(r_B + (1 - r_B)[r_{AB} \frac{\rho_A^0}{\rho_A} + (1 - r_{AB})\rho_B^0] \right).$$

Obviously, $\rho_B^{+-} > \rho_B$. Starting out by investing in B and observing a success there yields

$$\rho_A^{-+} = \left(r_B + (1 - r_B)[r_{AB} \frac{\rho_A^0}{\rho_A} + (1 - r_{AB})\rho_B^0] \right) \frac{\rho_A}{\rho_B} = \rho_B^{+-} \frac{\rho_A}{\rho_B}.$$

3.4.2 Profits and Capacities

In this section we derive profits and capacities under sequential entry. For convenience, we continue to assume that the MNE enters country A first, and B subsequently.

Sequential entry yields total expected profits of

$$\Pi^{seq} = X_A^{seq} \rho_A \theta^h R_A (1+\delta) - k_A \frac{(X_A^{seq})^2}{2} - F_A + \delta X_A^{seq} \rho_A \theta^h \left(X_B^{seq} \rho_B^{+-} \theta^h R_B - k_B \frac{(X_B^{seq})^2}{2} - F_B \right). \quad (4)$$

$X_A^{seq} \rho_A \theta^h R_A (1 + \delta) - k_A \frac{(X_A^{seq})^2}{2} - F_A$ collects profits generated in A. The term $X_A^{seq} \rho_A \theta^h \left(X_B^{seq} \rho_B^{+-} \theta^h R_B - k_B \frac{(X_B^{seq})^2}{2} - F_B \right) = X_A^{seq} \rho_A \theta^h \Pi_B^{seq}$ describes expected profits from entering country B valued in period 2. It is the product of the probability that this actually happens, i.e., the probability of success in A, and the expected profits in B given entry there. Then, the capacity in B equals

$$X_B^{seq} = \frac{\rho_B^{+-} \theta^h R_B}{k_B},$$

yielding $\Pi_B^{seq} = \frac{\left(\rho_B^{+-} \theta^h R_B\right)^2}{2k_B} - F_B$.
When determining X_A^{seq}, the potential profits in B are taken into account, and we obtain

$$X_A^{seq} = \frac{\rho_A \theta^h R_A (1 + \delta) + \delta \rho_A \theta^h \Pi_B^{seq}}{k_A}. \quad (5)$$

Proposition 2 (Sequential Entry): The capacity chosen in the first country under sequential entry is larger than the capacity in this country under simultaneous entry.

The capacity in A is higher under sequential than under simultaneous entry because it not only

[8] We also show that under correlated learning beliefs follow a martingale – implying that they do not change in expectation – just as before when we considered the updating process for one country in isolation.

raises expected revenues there but also the probability for entry in B for a given belief ρ_B^{+-}. As the expected profits in B are positive by construction (otherwise, isolated entry would be better), a larger capacity in A increases the likelihood of a realization of these profits. This implies that expected total profits (net of investment costs) in A are lower than under simultaneous entry. Whether the capacity in B is lower or higher depends on the size of the discount factor.

Different from the standard literature on investment under uncertainty (see McDonald and Siegel, 1986; Dixit and Pindyck, 1994; Carruth, Dickerson, and Henley, 2000), the degree of learning is not exogenously given but implied by the capacity choice in country A. Thus, the firm balances costs of learning (higher capacity in A than individually optimal) with its benefits (higher probability of realizing profits in B).

Finally, sequential entry upon entry in A gives expected profits

$$\Pi^{seq} = \frac{1}{2k_A}\left[\rho_A\theta^h R_A(1+\delta) + \delta\rho_A\theta^h \left(\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B} - F_B\right)\right]^2 - F_A. \tag{6}$$

Similar to the cases of simultaneous and isolated entry, the probability of sequential entry to be profitable (which does not mean that it is actually optimal) increases with expected profits, i.e., in ρ_j, r_j,[9] θ^h, and R_j, and it decreases with costs, i.e., in k_j and F_j. The next proposition contains our first main result.

Proposition 3: Sequential entry or simultaneous entry can be optimal, depending on parameter values. Sequential entry is even possible if individual expected profits in both countries are positive at the beginning.

This proposition links our results to the one proposed by optimal investment decisions under uncertainty (McDonald and Siegel, 1986; Dixit and Pindyck, 1994). If expected profits in both countries are strictly positive ex ante, delaying entry for one of them might still be optimal. Despite the positive net present value of an investment, the option value of waiting may be higher.

However, sequential entry will only be optimal if sufficient correlated learning occurs, which requires two elements. First, the distance between countries A and B must not be too high (r_{AB} must be high enough). Second, some uncertainty has to prevail, as otherwise no substantial updating can occur. If a firm is already very optimistic about its type in one country (ρ_j is close to unity), beliefs will only be updated marginally.

Let us use these results and attend to the tradeoff the firm faces when considering sequential entry. It can reduce the total risk of investment (investment costs in markets with a low type) by using information gained in country A for activities in B.[10] The information gathered in A is only valuable for B, if the difference between ρ_B^{+-} and ρ_B is sufficiently large. To see this, consider the extreme case where $\rho_B = \rho_B^{+-}$ (which will be the case if either $\rho_A = 1$ or $\rho_B = 1$). Then, simultaneous or isolated entry always dominates sequential entry (for a formal analysis see the proof related to Hypothesis 2 in

[9] However, note that obtaining comparative statics with respect to r_j is less straightforward here, since a higher level of r_A decreases ρ_B^{+-}. We show below – in the proof associated with Hypothesis 1 – that this claim is true under sequential entry.

[10] For example, a high level of F_B will make the risk reduction through waiting and learning more valuable.

Appendix I). This remains the case as long as $\Delta\rho_B^{seq} \equiv \rho_B^{+-} - \rho_B$ is relatively small. Since this is an important aspect, let us take a closer look at $\Delta\rho_B^{seq} = (1-r_B) r_{AB} \frac{\rho_A^0}{\rho_A}(1-\rho_A)$. As $\rho_A = r_A + (1-r_A)\rho_A^0$, we arrive at the following comparative static results:

$$\frac{\partial \Delta\rho_B^{seq}}{\partial r_{AB}} = (1 - r_B)\left(\frac{\rho_A^0}{\rho_A} - \rho_A^0\right) > 0, \tag{7}$$

$$\frac{\partial \Delta\rho_B^{seq}}{\partial r_B} = -r_{AB}\left(\frac{\rho_A^0}{\rho_A} - \rho_A^0\right) < 0, \tag{8}$$

$$\frac{\partial \Delta\rho_B^{seq}}{\partial r_A} = -(1 - r_B) r_{AB} \frac{\rho_A^0(1 - \rho_A^0)}{\rho_A^2} < 0. \tag{9}$$

Expression (7) implies that if conditions in countries A and B are more similar to each other, the correlation in learning is higher, and a positive outcome in A is a stronger signal concerning the profitability in B. Equation (8) states that if B is already relatively close to home, then the additional amount of learning from A about B is smaller. Finally, expression (9) states that if A is closer to home, entry there makes the firm learn less about the conditions in B. In the extreme case, if $r_A = 1$, the type in A can not be distinguished from the type at home. Then, $\Delta\rho_B^{seq} = 0$, and it is not possible to learn something from A about the conditions in B. Again, this part relates to the question of optimal investment under uncertainty. A higher degree of uncertainty increases the option value of waiting and thus raises the threshold of required profits to make entry actually optimal. Here, a sufficient degree of learning as characterized by $\Delta\rho_B^{seq}$ is required to render sequential entry an optimal choice.[11]

To sum up, for sequential entry to be optimal, a success must reveal sufficient information about the first country that is entered as well as about the second one.

4 Empirical Analysis

4.1 Testable Hypotheses

In this section, we use the theoretical model to derive predictions and formulate them in a way that allows us to test them empirically. We will refer to these *predictions* as testable *hypotheses* (all proofs associated with Hypotheses 1-3 can be found in Appendix I).

Hypothesis 1: *Foreign market entry should be more likely for larger levels of R_j and θ^h, for lower costs k_j and F_j, and for a larger value of general proximity r_j.*

R_j or θ^h capture a firm's profitability in a market provided that it is generally successful there (its type is high). Profitability may be affected by market size, which can be measured by a country's GDP, and other aspects that have a direct impact on profits, such as a country's profit tax rate. Concerning the costs of market entry, we consider measures such as corruption, investment freedom, or the general costs of starting a business in a country (see Chen and Moore, 2010). Finally, a high proximity to

[11] Note that we can not establish a simple monotone rule claiming that a higher level $\Delta\rho_B^{seq}$ increases the profitability of sequential relative to simultaneous entry. The reason is that it is not possible to analyze a change of $\Delta\rho_B^{seq}$ in isolation, as all its components have an impact on other variables (other than the degree of learning or uncertainty).

the home country should render entry into a market more likely. Geographical distance is obviously a good proxy for the parameter r_j. But also cultural factors such as a common language with the home market are expected to positively affect a firm's propensity to enter a foreign market.

Hypothesis 2: *Sequential entry can be the optimal entry mode. If it is chosen, more proximate countries (higher r_j) should generally be entered first.*

Hypothesis 2 states that the first foreign investment of a firm, which may relate to sequential or simultaneous entry, should on average be closer to home.[12] This also implies that the marginal effect of the proximity parameters should be larger in absolute value for earlier compared to later entry.

With sequential entry, the closer country should generally be entered first as long as two foreign markets do not differ too much in size (and, hence, profitability). If r_j is close to unity, almost no correlated learning occurs; see equations (7) and (9) and the related discussion. Then, simultaneous entry is optimal.

Hypothesis 3: *Provided that market A is entered in period 1 but B is not, a higher value of proximity between A and B, r_{AB}, should increase the probability that the MNE enters B in period 2.*

Hypothesis 3 predicts that a greater (geographical) distance between countries of different expansion phases will reduce the probability to enter a country at a later stage. In this sense, later expansion phases depend on all previous investments.

The above hypotheses suggest that a firm should rather enter more promising markets in terms of market size and costs than others. Furthermore, an expansion of a multinational network of foreign affiliates should, on average, follow a certain pattern – starting in closer countries, then gradually increasing in distance from home but remaining close to markets entered previously. In the following, we show that such entry patterns are indeed observed in our data. For this empirical analysis, we use a unique micro-level data-set provided by Deutsche Bundesbank (the German Central Bank) that allows us to track the universe of German MNEs' sequential location decisions over time. We will see that the patterns observed in the data are largely in line with those hypotheses.[13]

4.2 Empirical Model Specification

Let us index German MNEs by $i = 1, ..., N$ and focus on the location choice of their affiliates among $j = 0, 1, ..., J$ foreign host countries. In any phase $p = 1, ..., P$ (corresponding to periods in our theoretical model), MNE i can choose among the J host markets with regard to location of its foreign entities.[14] Since we are interested in the genesis of MNEs' foreign affiliate networks, we associate

[12] Note that, after the first investment, all subsequent investments must be part of a sequential entry strategy.

[13] At this point, the patterns described by Hypotheses 1-3 and found in the data could still be generated by other mechanisms than the proposed correlated learning channel. However, we analyze this issue theoretically below and provide evidence supporting the proposed learning process for the genesis of multinational foreign affiliate networks.

[14] Notice that investment phases are unequally spaced in real time across firms. Hence, *phases* should not be confused with *years*. For instance, the first foreign investment of firm i may take place in any year covered by our sample period. Hence, a phase is associated with a vintage of foreign investments per firm.

expansion phases of the network with p. While MNEs typically set up one foreign affliate per phase p, in some cases they locate in several markets simultaneously in p. Each of these decisions will be treated individually below.[15] There is a number of options for modeling such a multinomial choice problem by means of nonlinear multinomial probability models. Examples thereof are the classes of multinomial probit-type models and multinomial logit-type models. A great advantage of the latter is that they follow from utility maximization of households or, as in our case, profit maximization of firms (see Wooldridge, 2002, p. 500f.). The same would be true for multinomial probit-type models, but with a huge number of $N = 15,171$ firms choosing among as many as $J = 104$ host countries as in our case,[16] it is natural to resort to multinomial logit-type models due to their tractability and numerical stability.[17] In the class of logit-type models, the conditional logit is a natural candidate for modeling the problem at stake, since it allows for regressors which are indexed by alternative j (and possibly also by firm i).[18]

We postulate that firm i would receive latent net profits Π^*_{ijp} from locating an affiliate at market j in phase p consistent with our theoretical model according to the process

$$\Pi^*_{ijp} = Z_{ijp}\beta_p + \alpha_{ijp}, \quad i = 1, ..., N,\ j = 0, 1, ..., J,\ p = 1, ..., P \tag{10}$$

where the $1 \times L_p$ vector Z_{ijp} contains determinants of the profits which depend on the alternative and, eventually, on firm i in any phase p. The $L_p \times 1$ vector of weights β_p on Z_{ijp} are unknown and will be estimated by maximum likelihood estimation. α_{ijp} represents unobservable variables affecting the choice. The actual choice $C_{ip} \in \{0, 1, ..., J\}$ is based on the maximum attainable profit, $\arg\max(\Pi^*_{i0p}, ..., \Pi^*_{iJp})$. Following McFadden (1974) in assuming that the α_{ijp} are independently distributed across alternatives with a type I extreme value distribution and using the notation $Z_{ijp} = (Z_{i0p}, ..., Z_{iJp})$,

$$P_{ijp} \equiv Pr(C_{ip} = j|Z_{ijp}) = \frac{\exp(Z_{ijp}\beta)}{\sum_{j=0}^{J}\exp(Z_{ijp}\beta)}, \quad \text{for all } i, j, p \tag{11}$$

is the probability of the actual choice C_{ip} being j. The marginal effect of the kth variable Z_{ijp} is $\partial P_{ijp}/\partial Z_{ijpk} = P_{ijp}(1 - P_{ijp})\beta_{pk}$ for all i, j, p, k and $\partial P_{ijp}/\partial Z_{i\ell pk} = -P_{ijp}P_{i\ell p}\beta_{pk}$ for all $i, \ell \neq j, p, k$. A well-known assumption taken by this approach is the one of independence from irrelevant alternatives (i.e., that the choices taken with regard to alternatives j versus ℓ are not affected when adding further alternatives).[19]

Furthermore, note that the restrictions to two firms, two host countries, and two periods in our theoretical model has no qualitative impact on the derived hypotheses.

[15] Accordingly, index i in fact denotes the choice of an MNE about a specific affiliate. However, for the ease of presentation, it is sufficient to refer to i as a firm.

[16] In principle, MNEs may enter as many as 162 countries, but in 58 of them not a single investment occurs so that those choices are dropped in the analysis.

[17] Multivariate probit-type models require integrating numerically a multivariate normal whose dimensions are determined by the number of choices taken. In spite of the efficient simulation algorithms available nowadays, for a choice problem as large as ours and a data-set which is not accessible locally so that computers can not be employed over extended time spans, it is virtually impossible to run multinomial probit-type or nested logit-type models.

[18] What is referred to as the multinomial logit model in a narrow sense assumes that the regressors only vary across firms i but not alternatives j in any phase p, while the parameters on those regressors vary across alternatives. It is well known that this model can be represented by the conditional logit model, where regressors rather than parameters are specific to the alternatives. Again, for as many alternatives as in our case, it appears unnatural to estimate J parameter vectors.

[19] Alternative modeling choices such as multivariate probit or nested logit models do not assume an independence

4.3 Data

We use data on the universe of German MNEs' foreign entities according to the classification taken by Deutsche Bundesbank[20] and as collected in and made available through the Bank's MiDi (Micro-database Direct Investment) database (see Lipponer, 2009, for details). Individual MNEs and their affiliates can be tracked annually in MiDi since 1996. The Census-type coverage of MNEs and their foreign affiliates renders this database particularly suited for an analysis of the genesis of multinational networks of foreign affiliates.

The vector of determinants of location decisions of firm i in phase p, Z_{ijp} in equation (10), contains the following regressors. The statutory corporate tax rate of the host country, $Tax_{jp} \in [0,1]$, reduces a firm's profitability ceteris paribus. The log of real GDP at constant U.S. dollars of the year 2000, $\log GDP_{jp}$, is a measure of j's market size. A number of variables are supposed to reflect fixed investment costs F in terms of our theoretical model, namely an investment freedom index, $InvestFree_{jp} \in [0,100]$, and a corruption perception index $CPI_{jp} \in [0,10]$,[21] as inverse measures of investment costs, as well as $InvestCost_{jp} \in [0,\infty)$ which reflects costs of starting a business relative to income per capita. The stock of German investments prior to firm i's investment in j and phase p, $StockInv_{jp} \in [0,\infty)$, is included as a general measure of market j's attractiveness for German investors beyond the aforementioned measures thereof. Furthermore, a number of variables determine the correlation between markets entered in phases p and ℓ, $1 \leq \ell < p$, in terms of economic, cultural, and geographical proximity. Note that, when considering the first location decision ($\ell = 1$), ℓ always refers to Germany. The following measures of distance or proximity across markets are employed: host market j's geographical distance to Germany, $\log Distance_{jp}$ *to parent* $\in (-\infty,\infty)$, a common border indicator between Germany and host market j, $Border_{jp}$ *to parent* $\in \{0,1\}$, a common language indicator between Germany and host market j, $Language_{jp}$ *same as parent* $\in \{0,1\}$, a former colony indicator between Germany and host market j, $Colony_{jp}$ *of parent* $\in \{0,1\}$, and a preferential trade agreement indicator between Germany and host country j, GTA_{jp} *with parent* $\in \{0,1\}$.

When analyzing subsequent investment decisions (see below) for $p \geq 2$, we will also control for the indicator variable $Same_{jp} \in \{0,1\}$, which is unity if host country j and the country of the previous investment i are the same. Since $Same_{j1} = 0$ for all host countries j in the sample, $Same_{jp}$ is included only in the specifications for the second and subsequent investment phases. The sources for the data on the control variables are the World Bank's World Development Indicators 2009 ($\log GDP_{jp}$, $InvestCost_{jp}$), International Bureau of Fiscal Documentation, Ernst&Young, and Price Waterhouse Coopers (Tax_{jp}), Transparency International (CPI_{jp}), Deutsche Bundesbank's MiDi ($StockInv_{jp}$), the Centre d'Études Prospectives et d'Informations Internationales ($\log Distance_{jp}$ *to* $\ell - th$, $Border_{jp}$ *to* $\ell - th$, $Language_{jp}$ *to* $\ell - th$, $Colony_{jp}$ *to* $\ell - th$, $Same_{jp}$ *to* $\ell - th$), and the World Trade Organization as well as individual preferential trade agreement secretariates' webpages (GTA_{jp} *to* $\ell - th$).

---

of the relative odds between choices j and ℓ from irrelevant alternatives. However, as said before, these models are computationally demanding and, with a choice and firm data-set as large as ours and the conditions imposed on empirical analysis through computing at the site of the data source, even infeasible to estimate.

[20] All German firms and households which hold 10% or more of the shares or voting rights in a foreign enterprise with a balance-sheet total of more than 3 million Euros are required by German law to report balance-sheet information to Deutsche Bundesbank. Indirect participating interests have to be reported whenever foreign affiliates hold 10% (50% as of 2007) or more of the shares or voting rights in other foreign enterprises. These reporting requirements are set by the Foreign Trade and Payments Regulation.

[21] Higher values of that index measure lower levels of perceived corruption.

Since the purpose of our analysis is to shed light on the determinants of an establishment of foreign affiliates per phase of investment, we restrict our interest to those firms for which we know that they did not operate any foreign affiliates in the first available year of the data, 1996. Hence, phase $p = 1$ with the first foreign investment of firm i may correspond to 1997 or any later year. Our data-set covers all first or subsequent investments of firms that became MNEs in 1997 or thereafter until 2008. Moreover, there are as many as $P - 1 = 11$ subsequent expansion phases possible in the data between 1998 and 2008. All of a firm's new affiliates which are founded across different years are associated with specific phases p and dubbed *sequential* investments, while a number of new affiliates founded within the same year are associated with the same expansion phase p and dubbed *simultaneous* investments. The design is such that $p = 1$ refers to the first set-up of one or more affiliates of firm i abroad, no matter in which year between 1997 and 2008 it occurred, and similarly for subsequent phases $p \geq 2$.

First Foreign Investments (p=1):

For first investments, $Z_{ij1} = Z_{j1}$ in (10) includes only determinants which pertain to the host country the first affiliate may be or is located in. First foreign investments can in principle occur in more than a single host market as investments in any phase p. We will relate subsequent investments to the largest investment in phase $p - 1$ in terms of *fixed assets* for any phase $p \geq 2$. Using *total assets* as an alternative criterion does not lead to alternative conclusions. See also Section 4.5 for further sensitivity checks on this issue.

Second and Subsequent Foreign Investments (p>1):

According to our theoretical model, firm-specific decisions about first investments matter for subsequent foreign investments. Therefore, the determinants for subsequent expansions of the MNE foreign affiliate network will be collected in the matrix Z_{ijp} for $p \geq 2$ in equation (10), which is indexed by i as well as j apart from p. In phases $p \geq 2$, Z_{ijp} includes regressors which are specific to host market j for the p-th investment, but it also includes ones that are firm-and-host-market specific in the sense that they relate to previous investments for firm i in phases $\ell < p$. By design, the number of regressors is $L_p > L_\ell > L_1$ for all phases $\ell < p$ with $p \geq 2$.

Covariates which relate foreign investments in phase p to previous ones are the following: the log distance of an affiliate set up in market j and phase p to the investments in earlier phases, log $Distance_{jp}$ *to* $\ell - th$; a common border indicator between an affiliate set up in market j and phase p with the investments in earlier phases, $Border_{jp}$ *to* $\ell - th$; a common language indicator between country j entered in phase p and countries entered in previous phases, $Language_{jp}$ *to* $\ell - th$; and similarly with colonial relationships ($Colony_{jp}$ *to* $\ell - th$), same country relationships ($Same_{jp}$ *to* $\ell - th$), and membership in a common goods trade agreement (GTA_{jp} *to* $\ell - th$). Table 1 presents descriptive statistics of all variables.

– Insert Table 1 and Figures 1-5 about here –

Figures 1-5 in the Appendix illustrate the frequency of first to fifth foreign investments by German MNEs. They indicate that market size is clearly important for the set-up of new affiliates in any phase. However, the distance to Germany is obviously much more important for first than for subsequent investments. While the figures provide some interesting descriptive information concerning the allocation of Germany's foreign affiliates, they can not substitute a multivariate analysis. Especially, the figures do not provide any insights into the influence of a sequence of previous investment decisions on subsequent ones.

4.4 Estimation Results

Table 2 summarizes results for sequential location decisions of MNEs. In every phase p, firms choose among approximately 100 host countries.

– Insert Table 2 about here –

We observe 15,165 *first* location decisions of MNEs in our sample analyzed in Column 1 of Table 2. In accordance with Hypothesis 1, a bigger market size ($\log GDP$) raises the probability of an investment. A higher tax burden measured by the statutory tax rate of a country (Tax) implies a lower probability to choose a location. This is consistent with the impact of R_j on the location choice in the theoretical model. Moreover, as stated by Hypothesis 1, lower costs of entry as captured by more investment freedom ($InvestFree$), lower fixed costs ($InvestCost$), and less corruption perception (CPI) are associated with a higher probability to locate in a country. This is consistent with the impact of F_j and k_j on the location choice in the theoretical model. Finally, the included measures of proximity suggest that the probability of choosing a location increases with r_j as stated in Hypothesis 1. For instance, a larger distance between Germany and a potential host country ($\log$ *Distance to parent*) reduces the probability of a first investment there. Similarly, if a potential host country shares a border with Germany (*Border to parent*), the location probability of a first investment increases. The variables *Language same as parent* and $Colony\ of\ parent$ measure proximity in terms of cultural similarity and historic ties, respectively. In both cases, the impact on the location probability of a first investment is positive. Finally, if Germany has signed a goods trade agreement with a host country (*GTA with parent*), a *first* foreign investment decision becomes more likely.

Columns 2 to 5 of Table 2 summarize the results for the *second* up to the *fifth* location decision (phase). The findings with respect to the (unilateral) host-country variables are qualitatively very similar and all coefficient point estimates have the expected signs. Note that a positive *fifth* location decision is observed for only 958 affiliates but the number of (columns in Z_{ij5} and) parameters to be estimated is largest among all models in Table 2. Hence, the coefficients in the last column of Table 2 are estimated with less precision than the ones pertaining to the *first* to the *fourth* investments. All of that is also broadly consistent with Hypothesis 1.

The results for the *second*, *third*, and *fourth* location decisions reveal an interesting pattern, confirming our theoretical considerations as stated in Hypotheses 2 and 3. While the distance effect between foreign investments in phase p to ones in phase $p-1$ is always negative (Hypothesis 3), it becomes less important in terms of magnitude over the expansion path of a multinational network of foreign affiliates (Hypothesis 2). This pattern clearly confirms some form of regional development of MNE foreign affiliate networks, similar to the development of export networks identified in the literature on sequential exporting (see Evenett and Venables, 2002; Albornoz, Calvo Pardo, Corcos, and Ornelas, 2011). This feature does not accrue to the sample composition but is also reflected in the marginal effects (see Table 3).

– Insert Table 3 about here –

Table 3 presents marginal effects of proximity variables and other determinants of location decisions about foreign affiliates. Lines 1 and 2 imply that the marginal impact of parameters referring to potential profitability in a market declines over time, wheres lines 3-5 suggest that the marginal impact of lower fixed costs gets broadly less important for later stages. We use this result as a robustness test below, addressed in Hypotheses R4 and R5 in Section 5.

Let us particularly emphasize two results in Table 2. First, whether or not a host country was a former colony seems relatively important in expansion phases $p \geq 2$. One reason for this result may be that the variable Colony captures many different aspects of *proximity.* Second, while we find that having a goods trade agreement (GTA $= 1$) with the parent makes it less likely to locate in a country, trade agreements between countries of subsequent location decisions increase the probability of establishing affiliates there. This stays in contrast to the literature on tariff-jumping FDI which stipulates that trade agreements may lead to a consolidation of foreign affiliates in response to preferential tariff liberalization (see Raff, 2004).

4.5 Sensitivity Analysis

In contrast to the models estimated in Table 2, the ones in Table 4 include the total stock of German investments in market j and phase p (StockInv$_{jp}$) prior to a firm's location decision there, while otherwise including the same regressors as in Table 2. This modification aims at checking whether or not the estimated coefficients are mainly driven by agglomeration effects – such as a general tendency of German firms to locate in just a few countries. StockInv should be a good measure of a market's general attractiveness for German investors beyond the dimensions captured by the covariates included in the regressions of Tables 1 and 2. The results in Table 4 suggest that the earlier findings are robust against the inclusion of StockInv. In fact, most of the coefficients are hardly affected by the additional control variable and, hence, are not biased due to omitted determinants of location choice.

– Insert Table 4 about here –

Recall that MNEs may establish more than one affiliate in an expansion phase p. If two or more investments are conducted in different countries in phase $\ell < p$, the reference of investments in p to ones in phase ℓ through Distance, Border, Language, Colony, etc., is no longer clear. We solved this problem in Table 2 by using the country of the largest previous investment in terms of *fixed assets* as the reference country in phase p. In Table 5 we use the largest previous investment in terms of *total assets* as alternative criterion to determine the reference country in phase p. The results displayed in Table 5 show that using an alternative criterion does not lead to alternative conclusions.

– Insert Table 5 about here –

Table 6 presents results for a subsample of firms and affiliates where all investments of any previous phase $p - 1$ occurred in only one country (the firms might have established several affiliates in this country, though). Then, the *bilateral* variables Distance, Border, Language, Colony, etc., refer to a unique reference country throughout. The findings in Table 6 confirm our previous results in broad terms. However, we should note that the strategy applied in Table 6 leads to a significant loss of degrees of freedom along the expansion path of MNE foreign affiliate networks. The reason is that many MNEs set up foreign entities simultaneously in several countries at some point of the genesis of their network of foreign affiliates. Therefore, from the *third* location decision onwards, the coefficients can not be estimated precisely any more, due to the reduction in sample size as compared to the findings in Tables 2, 4, and 5.

– Insert Table 6 about here –

5 Further Hypotheses Tests and Robustness

Although our results concerning the genesis of a multinational network of foreign affiliates appear to be robust regarding some general features, it is not per se obvious that learning under uncertainty is the main factor driving the observed patterns of investments. In what follows, we will derive further – more specific – hypotheses, referred to as R1-R5, arising from our theoretical model (all proofs associated with Hypotheses R1-R5 can be found in Appendix I).

Let us first address the point that sequential entry is only observable *ex post*. More precisely, when there is foreign market entry in one period but no subsequent expansion of the network of foreign affiliates, we do not know whether this was intended from the beginning or not. An MNE could have planned to enter markets sequentially, but it might have turned out that it was not sufficiently successful in the first-entered markets to undertake subsequent investments. By this reasoning, firms that actually take the second step and make a sequential investment should have been relatively more successful in their first market(s). On the other hand, the pool of firms that only remained in their initial markets does not only include those that chose isolated or simultaneous foreign investments and were either successful or not; it also includes those that had planned sequential entry but were not successful in the first market(s). Furthermore, an investment in a country entered as second under sequential entry will on average be more successful than an otherwise identical investment in a country entered under simultaneous or isolated entry. The reason is that correlated learning together with previous success raises its belief, implying higher expected profits. This gives

Hypothesis R1: *Firms where sequential entry is observed are on average more successful than firms where isolated or simultaneous entry is observed in otherwise identical markets.*

The following hypothesis uses the result that, if a country is chosen as the first market of a planned sequential entry path, an MNE will have a larger capacity there than otherwise. Although the pool of observed isolated or simultaneous entries also contains planned but not realized sequential patterns, the capacity there should on average be smaller. Thus, we state

Hypothesis R2: *Firms where sequential entry is observed have on average a larger capacity than firms where only isolated or simultaneous entry is observed in otherwise identical markets.*

Hypotheses R1 and R2 relate to the size and profitability of MNEs. In particular, MNEs may differ in these dimensions depending on whether they enter markets simultaneously or sequentially. Table 7 presents regression results, where we use the indicator variable *Sequential entry* to distinguish between sequential and simultaneous entries. To be precise, the variable *Sequential entry* is unity if we identify an observation as a sequential entry and zero otherwise.

We analyze three different dependent variables in Table 7: the *fixed assets*, the *total assets*, and the *sales-to-total-assets ratio* of the *average* foreign investment, respectively. The different columns refer to the maximum number of entities a firm consists of over the whole time-span considered. For example, the column denoted by (3) includes firms that have established only 2 or only 3 entities in the period 1997–2008. In this example, by focusing on firms that are always 2-plant or 3-plant MNEs, we can distinguish between simultaneous and sequential market entry.

All results support Hypotheses R1 and R2: if sequential entry is observed ex post, the previously established affiliates are on average more successful and larger than those where only simultaneous or isolated entry is observed.

– Insert Table 7 about here –

Now, let us only consider investments where no sequential entry is observed (yet). As pointed out above, such investments may include ones where sequential entry was intended but not (yet) exercised.[22] On average, the corresponding firms should be less successful than other MNEs. As their capacity is higher as well, we formulate

Hypothesis R3: *For firms where simultaneous or isolated entry is observed, the more successful ones should on average exhibit a lower capacity.*

Table 8 presents a test of Hypothesis R3, focusing on one-plant MNEs. The dependent variable is fixed assets of a foreign affiliate. Consistent with Hypothesis R3, we observe that for firms where only isolated entry is observed ex post, more profitable ones (measured by the sales-to-fixed-assets ratio or the sales-to-total assets ratio) have lower amounts of fixed assets. We conduct the same test in column 2, but additionally include all multiple-first-foreign-affiliate units with simultaneous entry in the estimation sample. The findings are very similar.

– Insert Table 8 about here –

Furthermore, a crucial component of learning is that, if a market is entered at a later phase, the belief about that market is higher compared with an earlier entry. This can have interesting implications on the (marginal) propensity to enter a market. Take a country that can either be entered using simultaneous entry or, as a second investment, under sequential entry. To make entry optimal, the associated R_j must exceed (and equivalently fixed costs F_j must be lower than) a certain threshold, for a given belief. Since the belief is higher if the market is entered under sequential entry (and if a success

[22]This also contains investments where sequential entry is still planned. However, by selecting on firms that did not exercise sequential entry ex post within a given time span, there is an intended bias towards firms which will not exercise sequential entry in the future. The latter should be sufficient for the proposed inference.

in the first country was observed), the relevant threshold making entry optimal for R_j should be lower and the one for F_j higher than when this country is entered under simultaneous entry. Although this result is less straightforward when the country is entered first under sequential entry, if countries are not too different, we can confirm the above finding. Thus, we propose

Hypothesis R4: *If a country is entered at later expansion phases, the minimum market size necessary to enter should be smaller. Moreover, the maximum fixed costs making entry just profitable should be higher than for ones entered earlier.*

Lines 1-5 of Table 3 give marginal effects of variables characterizing market size (lines 1 and 2) and fixed costs of market entry. Until the third investment, the marginal effects have the predicted patterns. Marginal effects of *Tax* and *log GDP* decrease in absolute terms, consistent with correlated learning being important for entry decisions. Furthermore, while the marginal effect of *InvestCost* does not seem to differ much for different entry stages, the development of the marginal effects of *InvestFree* and *CPI* are largely as predicted.

Finally, let us establish another hypothesis that makes use of correlated learning. For higher fixed entry costs, learning is more valuable, i.e., the option value of waiting is larger.

Hypothesis R5: *If fixed entry costs in one country are ceteris paribus higher, it is more likely that this country is chosen as second under sequential entry. Furthermore, the relative profitability of sequential compared to simultaneous entry increases if the fixed entry costs in the second target country are higher. Thus, fixed costs should on average matter less for countries entered at later stages.*

Hypothesis R5 is supported by Table 1, where lines 3-5 of the host-country variables give average values for parameters capturing fixed entry costs. There, especially the estimated parameters of *InvestFree* and *InvestCost* are as predicted, where the former decreases and the latter increases along expansion phases.

Furthermore, Hypothesis R5 can help to explain the seemingly counterintuitive impact of trade agreements (GTA) with the parent for countries entered in the second and later phases. Whereas the first investment is positively affected by such an agreement, the impact is negative for later ones. If GTAs are associated with fixed cost, this contradicts the (otherwise empirically supported) hypothesis that higher fixed costs should generally be associated with a lower probability of entry. However, if firms enter countries with high fixed cost, this will rather happen at later stages than at earlier stages of the genesis of multinational foreign affiliate networks.

6 Alternative Explanations for the Genesis of Multinational Networks

Although the observed sequential entry and expansion patterns of MNE foreign affiliate networks can not be explained by static models of market entry, there is a number of alternative dynamic models which could lead to predictions that are qualitatively similar to the ones derived from our model. Here, we briefly analyze three prime candidates of alternative models, namely stochastic shocks, diseconomies of scale, and learning by doing. The main difference between those models and ours is that, in the alternative modeling environments, the MNE faces uncertainty concerning its type in the respective market in each period. While we concede that any one of the three alternative models may matter and be consistent with some features of the gradual expansion of MNE foreign affiliate networks, we will show that any one of them fails to explain important facets of the data. The reason is that the proposed correlated learning model renders decisions in later periods contingent on the outcome in earlier periods, while the three alternative explanations do not. In the absence of uncertainty concerning success in a market, second-period actions are generally independent of success in the first period.

Alternative 1: Stochastic Shocks

One reason for why a firm might not want to enter all markets simultaneously is that exogenous factors affect its profitability there. Then, it will not invest unless market conditions turn out to be sufficiently good.

The setup for such a model is identical to the one derived above, with two exceptions. The firm's type in each country is not identical over time, but a new realization is drawn at the beginning of each period. The probability that the type in country j is high in a given period equals q_j, $j = A, B$. We impose no further structure on q_j. It could depend on the distance to home (r_j) or to the other potential host country (r_{AB}). Furthermore, the firm can observe the realizations of θ_j, $j = A, B$, in each period, so that the only uncertainty it faces concerns next period's value of θ_j.

The MNE's entry decision with respect to country A is now independent of its entry decision for B (and vice versa). The reason is that past decisions do not have an impact on the likelihood of having a high type in the future. Thus, we can focus on optimal actions for just one market.

In the first period, the firm will not enter country j if $\theta_j = 0$, since this would yield negative profits in period 1 (without a positive impact on future profits). If $\theta_j = \theta^h$ in period 1, expected profits for a given capacity X_j are

$$X_j \theta^h R_j (1 + \delta q_j).$$

Conditional on entry, the firm will choose a capacity level $X_j = \frac{\theta^h R_j (1+\delta q_j)}{k_j}$ and finally enter the market if $\Pi_j = \frac{(\theta^h R_j)^2 (1+\delta q_j)^2}{2k_j} - F_j \geq 0$. If $\theta_j = 0$ in the first period, the MNE will enter the market in period 2 if the type is then high and if expected profits are positive, i.e., whenever $\Pi_j = \frac{(\theta^h R_j)^2}{2k_j} - F_j \geq 0$. If these conditions are satisfied, entry in periods 1 and 2 occurs with probability q_j and $(1 - q_j)q_j$, respectively. The total likelihood of entry thus equals $2q_j - q_j^2$.

A bigger market size and lower entry costs are also associated with a higher likelihood of entry. We might even construct a sequential entry pattern as observed in the data, with closer countries entered first, followed by a gradual expansion to markets farther away. This would require the assumption $\frac{\partial q_j}{\partial r_j} > 0$. It would be less straightforward – yet not impossible – to construct assumptions such that the role of r_{AB} would be similar to that in our benchmark model. Hypotheses R1-R5, however, will definitely not hold. Take Hypothesis R1, where we claim that firms that enter sequentially are on average more successful. Assume A is entered in the first period but B not. Then, the decision whether to enter B in period 2 is independent of what happened in A. Thus, expected profits in A are always the same, no matter whether B is entered in period 1, 2, or not at all. Similar arguments can be used to reject Hypothesis R2 (sequential entry is associated with higher capacity levels), and Hypothesis R3 (when countries are entered simultaneously, the ones with a lower capacity should be more successful). Furthermore, Hypothesis R4 is not supported. With stochastic shocks, the maximum tenable level of fixed costs or the minimum necessary level of market size would ceteris paribus be the same for first and subsequent investments. Hence, we can reject the first alternative model as an explanation for the observed empirical patterns in comparison to the proposed correlated learning model.

Alternative 2: Diseconomies of Scale

Here, we take into account that an MNE's resources in one period might be constrained. For simplicity, let us focus on financial resources and assume that investment costs in one period may not exceed the value D. Also, let us assume that it is known that the MNE's type is high in both markets. Everything else is identical to the original model setup. Thus, without financial constraint, the MNE would enter both countries at the beginning of period 1. Then, the firm would choose capacities $X_j = \frac{(\theta^h R_j)(1+\delta)}{k_j}$ and obtain expected profits $\Pi_j = \frac{(\theta^h R_j)^2(1+\delta)^2}{2k_j} - F_j$. For $D \geq k_A \frac{(X_A)^2}{2} + k_B \frac{(X_B)^2}{2} + F_A + F_B$, the budget constraint does not bind and simultaneous entry occurs. To simplify issues, let us assume that $k_j \frac{(X_j)^2}{2} + F_j \leq D \leq k_j \frac{(X_j)^2}{2} + F_A + F_B$. Accordingly, it is feasible to enter one country with the first-best capacity, but not possible to enter the second one at all. Hence, higher revenues or lower costs render first entry more likely again, and Hypothesis 1 would hold. If we further assume that R_j decreases with the distance to home, the geographically closer country would more likely be entered first, which is in line with Hypothesis 2. It would be more difficult to justify why the distance between two host markets A and B should matter for the sequential entry pattern, and Hypothesis R1 could only be obtained for this model under the assumption that the budget constraint in the second period is relaxed after the realization of a success in period 1. But Hypotheses R2 and R3 would definitely not flow from the diseconomies of scale model, since chosen capacities are independent of other entries. Note that Hypothesis R2 (first country of sequential entry has larger capacity) holds for otherwise identical markets and we can not use the argument that affiliates at initially entered markets are more successful. Finally, while Hypothesis R5 is in line with a model of diseconomies of scale (countries with higher fixed costs are on average entered later), this is not true for Hypothesis R4, since the relative thresholds above which entry is profitable do not change along entry phases.

Although the diseconomies of scale model does a relatively better job in explaining the pattern of MNE foreign affiliate network formation observed in the data than the stochastic shocks model, neither of them provides an explanation for the different capacity levels and their correlation with observed success. On those grounds, the second alternative model can be rejected as an explanation for the observed empirical patterns in comparison to the proposed correlated learning model.

Alternative 3: Learning by Doing

Finally, let us assume that second-period returns in both countries depend on first-period production, i.e., $R_j(X_A, X_B)$. We use the explicit linear expression for the second-period returns, which equal $R_A(X_A, X_B) = rX_A + \alpha r X_B + T_A$ and $R_B(X_A, X_B) = \alpha r X_A + rX_B + T_B$, with $\alpha, r \geq 0$, and $\alpha \leq 1$.

First-period returns equal T_j. Now, sequential entry might also be used to save investment costs in the first period and use learning benefits from A about B. It is easy to derive formal results, which we omit here; just note that Hypothesis 1 is consistent with learning by doing as with the proposed correlated learning model. If T_j is larger for the closer country and α increases with the two host countries' proximity to each other, Hypotheses 2 and 3 are consistent with learning by doing.

Yet, since entry into the second country in period 2 occurs for sure, Hypotheses R1-R5 are not generally consistent with learning by doing. Hypotheses R2 (sequential entry is associated with a larger capacity), R4 (entry thresholds differ across phases) and R5 (fixed costs are higher and matter less in later entry phases) may or may not be true, depending on parameter values. Hypotheses R1 (observed sequential entry associated with higher success) and R3 (larger capacity correlated with lower success for simultaneous or isolated entry) are not consistent with learning by doing, because outcomes in the first period do not have an impact on later decisions. Hence, the third alternative model can also be rejected as an explanation for the observed empirical patterns in comparison to the proposed correlated learning model.

7 Conclusions

This paper provides an explanation for the fact that multinational enterprises develop their networks of foreign affiliates gradually over time. Instead of exploring all profitable opportunities immediately, they first establish themselves in their home countries and then enter new markets stepwise. We explain this gradualism by proposing a model where MNEs face uncertainty concerning their success in new markets and learn about that after entry. Conditions in different markets are not independent, and the information gathered in one country can also be used to learn about conditions in other, in particular similar, countries.

This so-called correlated learning mechanism serves us to derive a number of testable hypotheses regarding market entry in general and *simultaneous* versus *sequential* market entry in specific. These hypotheses are assessed in a data-set of the universe of German MNEs and their foreign affiliates provided by Deutsche Bundesbank. We find that first foreign entry is more likely for countries that are closer to the MNE's home base and where higher profits may be expected in general terms. This finding is supported by variables measuring the proximity of markets at large, e.g., whether the same language is spoken or if the target country used to be a colony of the home country. Moreover, proximate countries tend to be entered first as a multinational network of foreign affiliates evolves. Third, subsequent entry in later expansion phases is generally more likely in markets that are closer to the ones entered previously.

Although other reasons such as stochastic shocks, diseconomies of scale, or learning by doing may certainly co-determine a multinational firm's expansion, a number of additional tests suggest that correlated learning is consistent with more facets of the data on the *genesis of multinational foreign affiliate networks* than the considered alternative models.

The way how MNEs expand their networks of foreign affiliates over time and, in particular, correlated learning as identified in this paper may have important policy implications. Understanding whether, how, and where firms grow is crucial for policy makers – not only with respect to domestic policies (such as tax policy) but also with regard to international policies (such as bilateral or multilateral preferential agreements on trade or investment). This is especially important since market characteristics – and, hence, in part economic policy – do not only affect location decisions of MNEs in the long run but also their timing and relative size.

In our future research, we aim at analyzing learning processes in more detail. For example, the role of learning might be different contingent on whether a firm acquires an existing affiliate or establishes a new plant. Correlated learning can also have an impact on market exit, which we have abstracted from, here. In subsequent work, we plan to pay attention also to the latter.

References

Albornoz, Facundo, Hector F. Calvo Pardo, Gregory Corcos, and Emanuel Ornelas, "Sequential Exporting," CEPR Discussion Paper No. DP8103, 2010.

Becker, Sascha O., Peter Egger, and Valeria Merlo, "How Low Business Tax Rates Attract Multinational Headquarters: Municipality-Level Evidence from Germany," CESifo Working Paper No. 2517, 2009.

Bellman, Richard, "A Problem in the Sequential Design of Experiments," *Sankhyā: The Indian Journal of Statistics (1933-1960)*, 16 (1956), 221–229.

Bergemann, Dirk, and Ulrich Hege, "Venture Capital Financing, Moral Hazard and Learning", *Journal of Banking and Finance*, 22 (1998), 703–735.

——, "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 36 (2005), 719–752.

Berry, Donald A., and Bert Fristedt, "Bernoulli One-Armed Bandits – Arbitrary Discount Sequences," *The Annals of Statistics*, 7 (1979), 1086–1105.

Carruth, Alan A., Andrew P. Dickerson, and Andrew Henley, "What do we know about Investment under Uncertainty?," *Journal of Economic Surveys*, 14 (2000), 119–153.

Chen, Maggie X., and Michael O. Moore, "Location Decision of Heterogeneous Multinational Firms," *Journal of International Economics*, 80 (2010), 188–199.

Conconi, Paola, André Sapir, and Maurizio Zanardi, "The Internationalization Process of Firms: From Exports to FDI?," Working Paper Research 198, National Bank of Belgium, 2010.

Devereux, Michael, and Rachel Griffith, "Taxes and the Location of Production – Evidence from a Panel of US Multinationals," *Journal of Public Economics*, 68 (1998), 335–367.

Dixit, Avinash K., and Robert S. Pindyck, *Investment under Uncertainty*, (Princeton, NJ: Princeton University Press, 1994).

Eaton, Jonathan, Marcela Eslava, Maurice Kugler, and James Tybout, "Export Dynamics in Colombia: Firm-Level Evidence," NBER Working Paper No. W13531, 2007.

Evenett, Simon J., and Anthony J. Venables, "Export Growth by Developing Economies: Market Entry and Bilateral Trade," unpublished manuscript, University of St. Gallen, 2002.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple, "Export versus FDI with Heterogeneous Firms," *American Economic Review*, 94 (2004), 300–316.

Jovanovic, Boyan, "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87 (1979), 972–990.

Keller, Godfrey, Sven Rady, and Martin Cripps, "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73 (2005), 39–68.

Lipponer, Alexander, "Microdatabase Direct Investment – MiDi. A Brief Guide," Deutsche Bundesbank Technical Documentation, Frankfurt, 2009.

McDonald, Robert, and Daniel R. Siegel, "The Value of Waiting to Invest," *Quarterly Journal of Economics*, 101 (1986), 707–727.

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, P. Zarembka ed. (New York: Academic Press, 1974).

Raff, Horst, "Preferential Trade Agreements and Tax Competition for Foreign Direct Investment," *Journal of Public Economics*, 88 (2004), 2745–2763.

Schmeiser, Katherine, "Learning to Export: Export Growth and the Destination Decision of Firms," unpublished manuscript, Mt. Holyoke College, 2009.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, (Cambridge, MA: MIT Press, 2002).

Appendix I – Proofs

Proof of Propositions 1-3

Proposition 1 (Isolated Entry): Given market entry, the optimal capacity under isolated entry equals $X^{iso} = \frac{\rho\theta^h R(1+\delta)}{k}$, and entry is only optimal if $\Pi^{iso} = \frac{(\rho\theta^h R)^2(1+\delta)^2}{2k} - F \geq 0$.

Proof: Note that beliefs follow a martingale, i.e., they do not change in expectation:
$\mathrm{E}[\rho_{t+1} \mid \rho_t, X_t] = X_t\rho_t\theta^h\rho_t^+ + (1 - X_t\rho_t\theta^h)\rho_t^-(X_t) = \rho_t$. Thus, (1) can be rewritten as
$\Pi^{iso} = \max\limits_{X\geq 0}\left[\{1\}_{X=0}0 + \{1\}_{X>0}\left(X\rho R\theta^h(1+\delta) - [F + k\frac{(X)^2}{2}]\right)\right]$.
Since there is no market exit after a failure and beliefs assume the martingale feature, expected profits in period 1 and 2 are identical (from the perspective of period 1). The first-order condition yields (2). The second-order condition is satisfied by the assumption of convexity of the investment cost function. As entry will only occur for non-negative profits, fixed investment costs have to be covered as well in expectation, i.e., $X(\rho\theta^h R - c)(1+\delta) - [F + k\frac{(X)^2}{2}] \geq 0$, yielding (3).

Proposition 2 (Sequential Entry): The capacity chosen in the first country under sequential entry is larger than the capacity in this country under simultaneous entry.

Proof: This immediately follows from comparing $X_A^{seq} = \frac{\rho_A\theta^h R_A(1+\delta)+\delta\rho_A\theta^h\left(\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B} - F_B\right)}{k_A}$ with $X_A^{sim} = \frac{\rho_A\theta^h R_A(1+\delta)}{k_A}$. The term $\left(\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B} - F_B\right)$ has to be positive as otherwise entry into B would not occur.

Proposition 3: Sequential entry or simultaneous entry can be optimal, depending on parameter values. Sequential entry is even possible if individual expected profits in both countries are positive at the beginning.

Proof: Assume $\Pi_B^{sim} = \frac{(\rho_B\theta^h R_B)^2(1+\delta)^2}{2k_B} - F_B^* = 0$, implying that isolated entry (only in A) and simultaneous entry yield identical profits.
$\Pi^{seq} > \Pi^{sim}$, if
$\frac{1}{2k_A}\left[\rho_A\theta^h R_A(1+\delta) + \delta\rho_A\theta^h\left(\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B} - F_B^*\right)\right]^2 - F_A > \frac{(\rho_A\theta^h R_A)^2(1+\delta)^2}{2k_A} - F_A$, or
$\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B} - F_B^* > 0$.
Thus, we need $\rho_B^{+-} > \rho_B(1+\delta)$. As $\rho_B^{+-} > \rho_B$, there is always a δ such that this is satisfied.

For the part that sequential entry can be optimal even if ex-ante profits in country B are strictly positive, assume that $F_B = F_B^* - \left(\rho_A\theta^h\right)^2\delta\varepsilon\frac{(2R_A+\delta\varepsilon)(1+\delta)^2}{2k_A}$, $\varepsilon > 0$, and entry into B already in the first period would yield a profit $\left(\rho_A\theta^h\right)^2\delta\varepsilon\frac{(2R_A+\delta\varepsilon)(1+\delta)^2}{2k_A}$.

For $\Pi^{seq} > \Pi^{sim}$, we need

$\frac{1}{2k_A}\left[\rho_A\theta^h R_A(1+\delta)+\delta\rho_A\theta^h\left(\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B}-F_B^*+\left(\rho_A\theta^h\right)^2\delta\varepsilon\frac{(2R_A+\delta\varepsilon)(1+\delta)^2}{2k_A}\right)\right]^2-F_A$
$>\frac{(\rho_A\theta^h R_A)^2(1+\delta)^2}{2k_A}-F_A+\frac{\left(2\delta\rho_A\theta^h\rho_A\theta^h R_A\varepsilon+\left(\delta\rho_A\theta^h\varepsilon\right)^2\right)(1+\delta)^2}{2k_A}$, or
$\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B}-F_B^*>\varepsilon\left((1+\delta)-\delta\left(\rho_A\theta^h\right)^2\frac{(2R_A+\delta\varepsilon)(1+\delta)^2}{2k_A}\right).$
This is possible for ε sufficiently small.

Finally, we have to make sure that entry into B after a failure in A is not optimal, which requires
$\frac{\left(\rho_B^{--}\theta^h R_B\right)^2}{2k_B}-F_B^*+\left(\rho_A\theta^h\right)^2\delta\varepsilon\frac{(2R_A+\delta\varepsilon)(1+\delta)^2}{2k_A}<0.$
We know that $\frac{\left(\rho_B\theta^h R_B\right)^2(1+\delta)^2}{2k_B}-F_B^*=0$ and that $\rho_B^{--}<\rho_B$ for $X_A^{seq}>0$. Thus, the above condition is satisfied for ε sufficiently small.

For the potential optimality of simultaneous entry, see Lemma A2 below, which states that there exists a value r_A^* such that for $r_A\geq r_A^*$, sequential entry is never chosen. Then, there are always values for F_A and F_B making simultaneous (and not isolated) entry optimal.

Proofs of Propositions underlying Hypotheses 1-3

Hypothesis 1: *Foreign market entry should be more likely for larger levels of R_j and θ^h, for lower costs k_j and F_j, and for a larger value of general proximity r_j.*

Proof: We aim at showing that the marginal impact on respective profits of R_i, r_i, and θ^h is positive, while it should be negative for k_i and F_i. The claim is obvious for isolated and simultaneous entry, where individual profits equal $\frac{(\rho_i\theta^h R_i)^2(1+\delta)^2}{2k_i}-F_i$ and comparative statics yield the predicted signs. Total profits under sequential entry are
$\Pi^{seq}=X_A^{Seq}\rho_A\theta^h R_A(1+\delta)-k_A\frac{(X_A^{Seq})^2}{2}-F_A+\delta X_A^{Seq}\rho_A\theta^h\left(X_B^{Seq}\rho_B^{+-}\theta^h R_B-k_B\frac{(X_B^{Seq})^2}{2}-F_B\right)$
$=\frac{1}{2k_A}\left[\rho_A\theta^h R_A(1+\delta)+\delta\rho_A\theta^h\left(\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B}-F_B\right)\right]^2-F_A.$
The Hypothesis is easily satisfied for entry into B, where profits, given a success in A was observed, equal $\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B}-F_B$, and $\frac{\partial\rho_B^{+-}}{\partial r_B}>0$. Concerning entry in period 1, comparative statics with respect to R_A, θ^h, k_A and F_A are unambiguous. This is different for r_A, as $\frac{\partial\rho_B^{+-}}{\partial r_A}<0$, and we can not exclude $\frac{\partial\Pi^{seq}}{\partial r_A}<0$. Still, to determine the likelihood of entry, we focus on the margin, i.e., where $\Pi^{seq}=0$. But if $\frac{\partial\Pi^{Seq}}{\partial r_A}\mid_{\Pi^{seq}=0}<0$, the MNE would choose isolated or simultaneous instead of sequential entry. As derived above, this becomes more likely for a larger value of r_A.

Hypothesis 2: *Sequential entry can be the optimal entry mode. If it is chosen, the country with a higher level of proximity r_j should generally be entered first.*

Proof: For the part that sequential entry might be optimal, see Proposition 2. When choosing sequential entry, we first show that for two countries which are identical and only differ in their distance to home, the MNE will enter the closer country first. Afterwards, we compare the profits under sequential entry when A is entered first with those when B is entered first. We look at the impact of r_A on the difference between these two measures and show that – as long as the countries are not too different – this impact works in favor of first entering A. Here, we are mainly interested on the impact of distance, i.e., if we expect to observe entry into closer countries first. Although a larger r_A decreases the updating in B, it generally makes it more likely that A is entered first. Let us first derive the result for the most stylized case where both countries are identical except for their distance to home. Then, it can be shown that entry first occurs into the country with the higher value of r_j.

Lemma A1: Assume $R_A = R_B \equiv R$, $\rho_A^0 = \rho_B^0 \equiv \rho^0$, $k_A = k_B \equiv k$, $F_A = F_B \equiv F$, and that sequential entry is chosen. Then, the MNE will first enter A if (and only if) $r_A \geq r_B$.

Proof of Lemma A1: Define $\Delta\Pi^{seq} \equiv \Pi^{seq}(AB) - \Pi^{seq}(BA)$. Then,

$$\Delta\Pi^{seq} = \frac{1}{2k}\left[\rho_A\theta^h R(1+\delta) + \delta\rho_A\theta^h\left(\frac{\left(\rho_B^{+-}\theta^h R\right)^2}{2k} - F\right)\right]^2 - F$$

$$-\frac{1}{2k}\left[\rho_B\theta^h R(1+\delta) + \delta\rho_B\theta^h\left(\frac{\left(\rho_A^{-+}\theta^h R\right)^2}{2k} - F\right)\right]^2 + F \geq 0, \text{ or}$$

$$\left(\rho_A\theta^h\left[R(1+\delta) + \delta\left(\frac{\left(\rho_B^{+-}\theta^h R\right)^2}{2k} - F\right)\right] + \rho_B\theta^h\left[R(1+\delta) + \delta\left(\frac{\left(\rho_A^{-+}\theta^h R\right)^2}{2k} - F\right)\right]\right) \times$$

$$\left(\rho_A\theta^h\left[R(1+\delta) + \delta\left(\frac{\left(\rho_B^{+-}\theta^h R\right)^2}{2k} - F\right)\right] - \rho_B\theta^h\left[R(1+\delta) + \delta\left(\frac{\left(\rho_A^{-+}\theta^h R\right)^2}{2k} - F\right)\right]\right) \geq 0.$$

As the first term is always positive, the sign of $\Delta\Pi^{seq}$ is determined by

$$\theta^h R(1+\delta)(\rho_A - \rho_B) + \delta\rho_A\theta^h \quad \frac{\left(\rho_B^{+-}\theta^h R\right)^2}{2k} - F\Bigg) - \delta\rho_B\theta^h \quad \frac{\left(\rho_A^{-+}\theta^h R\right)^2}{2k} - F\Bigg)\Bigg)$$

, which, as $\rho_A^{-+} = \frac{\rho_A}{\rho_B}\rho_B^{+-}$, can be rewritten as

$$\theta^h(\rho_A - \rho_B)\left[R + \delta\left(R - \frac{\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\theta^h R\right)^2}{2k} - F\right)\right]. \tag{12}$$

If we can show that the squared bracket of (12) is always positive, then $\text{sgn}(\Delta\Pi^{seq}) = \text{sgn}(\rho_A - \rho_B)$.

As we assume that $X \leq 1$ even if a type is known to be high and that entry is optimal for the high type, $k \geq \theta^h R(1+\delta)$ and $\frac{(\theta^h R)^2(1+\delta)^2}{2k} > F$ respectively. Then,

$$\left[R + \delta\left(R - \frac{\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\theta^h R\right)^2}{2k} - F\right)\right] \geq \left[R + \delta\left(R - \frac{\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\theta^h R\right)^2}{2k} - \frac{(\theta^h R)^2(1+\delta)^2}{2k}\right)\right]$$

$$\geq \left[R + \delta\left(R - \frac{\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\theta^h R\right)^2}{2\theta^h R(1+\delta)} - \frac{(\theta^h R)^2(1+\delta)^2}{2\theta^h R(1+\delta)}\right)\right] = R\left[1 + \delta\left(1 - \theta^h\frac{\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\right)^2}{2(1+\delta)} - \frac{\theta^h(1+\delta)}{2}\right)\right]$$

$$\geq R\left[1 - \frac{\delta}{1+\delta}\theta^h\frac{\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\right)^2}{2}\right], \text{ as } 1 - \frac{\theta^h(1+\delta)}{2} \geq 0.$$

Let us first consider the case (A) where $\rho_A - \rho_B \geq 0$. Now,

$$\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\right)^2 = \frac{\rho_A}{\rho_B}\left(\rho_B + (1-r_B)r_{AB}\frac{\rho^0}{\rho_A}(1-\rho_A)\right)^2$$

$$= \rho_A\rho_B + 2(1-r_B)r_{AB}\rho^0(1-\rho_A) + \frac{1}{\rho_A\rho_B}\left((1-r_B)r_{AB}\rho^0(1-\rho_A)\right)^2$$

$$= \rho_A\rho_B + 2(1-r_B)r_{AB}\rho^0(1-\rho_A) + \frac{1}{\rho_B\rho_A}\left((1-r_B)r_{AB}\rho^0(1-\rho_A)\right)^2,$$

and note that

$\rho_A\rho_B \leq 1$, $2(1-r_B)r_{AB}\rho^0\,(1-\rho_A) \leq 2$, and

$$\frac{1}{\rho_B\rho_A}\left((1-r_B)r_{AB}\rho^0\,(1-\rho_A)\right)^2 = \frac{(\rho^0)^2}{(r_B+(1-r_B)\rho^0)(r_A+(1-r_A)\rho^0)}\left((1-r_B)r_{AB}\,(1-\rho_A)\right)^2$$
$$\leq \frac{(\rho^0)^2}{(1-r_B)\rho^0(1-r_A)\rho^0}\left((1-r_B)r_{AB}\,(1-\rho_A)\right)^2$$
$$= (1-r_B)(1-\rho^0)r_{AB}^2\,(1-\rho_A) \leq 1.$$

Thus, $\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\right)^2 \leq 4$, and

$$\left[1-\frac{\delta}{1+\delta}\theta^h\frac{\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\right)^2}{2}\right] \geq \left[1-\frac{\delta}{1+\delta}2\theta^h\right]$$
$$= \frac{1}{1+\delta}\left[1-\delta\theta^h+\delta-\delta\theta^h\right] \geq 0.$$

Now consider case (B) where $\rho_A - \rho_B < 0$. Then,

$$\left[1-\frac{\delta}{1+\delta}\theta^h\frac{\frac{\rho_A}{\rho_B}\left(\rho_B^{+-}\right)^2}{2}\right] \geq \left[1-\frac{\delta}{1+\delta}\theta^h\frac{\left(\rho_B^{+-}\right)^2}{2}\right] \geq \left[1-\frac{\delta}{1+\delta}\theta^h\frac{1}{2}\right] \geq 0.$$

To have a better idea, we now allow for general parameter values and analyze $\frac{d\Delta\Pi^{seq}}{dr_A}$ (recall that $\Delta\Pi^{seq} = \Pi^{seq}(AB) - \Pi^{seq}(BA)$):

$$\frac{d\Delta\Pi^{seq}}{dr_A} = X_A^{AB}\frac{d\rho_A}{dr_A}\theta^h R_A(1+\delta) + \delta X_A^{AB}\frac{d\rho_A}{dr_A}\theta^h\left(\frac{(\rho_B^{+-}\theta^h R_B)^2}{2k_B} - F_B\right)$$
$$+\delta X_A^{AB}\rho_A\theta^h X_B^{AB}\frac{d\rho_B^{+-}}{dr_A}\theta^h R_B - \delta X_B^{BA}\rho_B\theta^h X_A^{BA}\frac{d\rho_A^{-+}}{dr_A}\theta^h R_A,$$

where X_A^{AB} is the capacity chosen in A under sequential entry starting in A. The first term describes increased profits in A, while the second term covers the increased likelihood of entry into B. The third term is negative, as ρ_B^{+-} decreases with r_A. Finally, the fourth term captures gone profits when A is entered as the second country. The expression can be rewritten as

$$\delta\theta^h R_A(1-\rho_A^0)\left[X_A^{AB} - \delta X_B^{BA}\theta^h X_A^{BA}(r_B+(1-r_B)(1-r_{AB})\rho_B^0)\right]$$
$$+\theta^h(1-\rho_A^0)X_A^{AB}\left[R_A + \delta\left(\frac{(\rho_B^{+-}\theta^h R_B)^2}{2k_B} - F_B - (1-r_B)r_{AB}\frac{\rho_A^0}{\rho_A}\frac{\rho_B^{+-}\theta^h R_B}{k_B}\theta^h R_B\right)\right].$$

Taking the term in squared brackets of the first line gives

$$\left(X_A^{AB} - X_B^{BA}\delta\theta^h X_A^{BA}(r_B+(1-r_B)(1-r_{AB})\rho_B^0)\right)$$
$$= \frac{\rho_A\theta^h R_A(1+\delta)+\delta\rho_A\theta^h\left(\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B} - F_B\right)}{k_A}$$
$$- \frac{\rho_B\theta^h R_B(1+\delta)+\delta\rho_B\theta^h\left(\frac{\left(\rho_A^{+-}\theta^h R_A\right)^2}{2k_A} - F_A\right)}{k_B}\delta\theta^h\frac{\rho_A^{+-}\theta^h R_A}{k_A}(r_B+(1-r_B)(1-r_{AB})\rho_B^0)$$
$$\geq \frac{\rho_A\theta^h R_A(1+\delta)}{k_A} - \frac{\rho_B\theta^h R_B(1+\delta)+\delta\rho_B\theta^h\left(\frac{\left(\rho_A^{+-}\theta^h R_A\right)^2}{2k_A} - F_A\right)}{k_B}\delta\theta^h\frac{\rho_A^{+-}\theta^h R_A}{k_A}(r_B+(1-r_B)(1-r_{AB})\rho_B^0)$$
$$\geq \left(\frac{\rho_A\theta^h R_A(1+\delta)}{k_A} - \delta\theta^h\frac{\rho_A^{+-}\theta^h R_A}{k_A}(r_B+(1-r_B)(1-r_{AB})\rho_B^0)\right)$$

(since $k \geq \theta^h R(1+\delta)$ and $\frac{(\theta^h R)^2(1+\delta)^2}{2k} > F$)

$$= \frac{1}{k_A}\theta^h R_A\rho_A\left((1+\delta) - \delta\theta^h\frac{\rho_B^{+-}}{\rho_B}\left(\rho_B - (1-r_B)r_{AB}\rho_A^0\right)\right)$$
$$= \frac{1}{k_A}\theta^h R_A\rho_A\left((1+\delta) - \delta\theta^h\rho_B^{+-} + \delta\theta^h\frac{\rho_B^{+-}}{\rho_B}(1-r_B)r_{AB}\rho_A^0\right) \geq 0.$$

The term in squared brackets of the second line equals

$$\left[R_A + \delta\frac{(\theta^h R_B)^2}{k_B}\rho_B^{+-}\left(\frac{\rho_B^{+-}}{2} - (1-r_B)r_{AB}\frac{\rho_A^0}{\rho_A}\right) - \delta F_B\right]$$
$$\geq \left[R_A - \delta(1-r_B)r_{AB}\frac{\rho_A^0}{\rho_A}\frac{\rho_B^{+-}\theta^h R_B}{k_B}\theta^h R_B\right] \geq \left[R_A - \frac{\delta}{(1+\delta)}(1-r_B)r_{AB}\frac{\rho_A^0}{\rho_A}\rho_B^{+-}\theta^h R_B\right].$$

As R_B is multiplied with terms that are all smaller than 1, the last term can only be negative if R_B is much larger than R_A. Therefore, $\frac{d\Delta\Pi^{seq}}{dr_A}$ will generally be positive.

Considering simultaneous entry, we can establish the following Lemma.

Lemma A2: There exists a value r_A^* such that for $r_A \geq r_A^*$, sequential entry is never chosen.

Proof of Lemma A2: Assume without loss of generality that if sequential entry is chosen, the MNE starts in A, and define $\Delta\Pi = \Pi^{seq} - \Pi^{sim}$.
Rewriting gives
$\Delta\Pi = \Pi_A^{seq} - \Pi_A^{sim} + \delta X_A^{seq} \rho_A \theta^h \Pi_B^{seq} - \Pi_B^{sim}$.
Furthermore,
$\Delta\Pi_A \equiv \Pi_A^{seq} - \Pi_A^{sim} = -\left(\delta\rho_A\theta^h\right)^2 \frac{\left(\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B} - F_B\right)^2}{2k_A} \leq 0$, as the capacity in A under sequential entry is too high if just profits in A are considered. Furthermore, $\lim_{r_A \to 1} \rho_B^{+-} = \rho_B$. For $r_A \to 1$, Π_B^{seq} approaches a value smaller or equal than Π_B^{Sim}. As $\delta X_A^{seq} \rho_A \theta^h < 1$, $\delta X_A^{seq} \rho_A \theta^h \Pi_B^{seq} - \Pi_B^{sim}$ is negative for $r_A = 1$. By continuity, the desired value r_A^* exists.

Note that this Lemma does not imply that for $r_A < r_A^*$, sequential entry is always optimal. This might or might not be the case, depending on parameter values.

Hypothesis 3: *Provided that market A is entered in period 1 but B is not, a higher value of proximity between A and B, r_{AB}, should increase the probability that the MNE enters B in period 2.*

Proof: The profits in B given sequential entry is chosen equal $\left(\frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B} - F_B\right)^2$. They are increasing in ρ_B^{+-}, which itself increases in r_{AB}.

Proofs of Propositions underlying Hypotheses R1-R5

Hypothesis R1: *Firms where sequential entry is observed are on average more successful than firms where isolated or simultaneous entry is observed in otherwise identical markets.*

Proof: The proof associated with Hypothesis R1 follows from the discussion in section 5.

Hypothesis R2: *Firms where sequential entry is observed have on average a larger capacity than firms where only isolated or simultaneous entry is observed in otherwise identical markets.*

Proof: Follows from the definition of sequential entry and Proposition 2. Furthermore, note the belief in markets entered in later stages is ceteris paribus higher and that these markets also serve as "first" countries for later stages. Under sequential entry, capacities should thus be higher along the whole investment path.

Hypothesis R3: *For firms where simultaneous or isolated entry is observed, the more successful ones* *should on average exhibit a lower capacity.*

Proof: The proof associated with Hypothesis R3 follows from the discussion in section 5.

Hypothesis R4: *If a country is entered at later expansion phases, the minimum market size necessary to enter should be smaller. Moreover, the maximum fixed costs making entry just profitable should be higher.*

Proof: Here, we neglect the restrictions imposed by having a model with only two periods. This restriction per se decreases expected profit streams for countries entered later. Since the expected time horizon will not automatically differ for different entry phases, effects induced by the reduced time horizon should not be emphasized too much.

Concerning the minimum requirements for market size R_j, compare profits when a country is entered under isolated or simultaneous and when it is entered – as second investment – under sequential entry. Without loss of generality, assume that this country is A. In the first case, the requirement for entry is

$$R_A \geq \frac{\sqrt{2k_A F_A}}{\rho_A \theta^h (1+\delta)}. \tag{13}$$

In the second case, entry occurs if and only if

$$R_A \geq \frac{\sqrt{2k_A F_A}}{\rho_A^{-+} \theta^h}. \tag{14}$$

Since $\rho_A^{-+} \geq \rho_A$, the right hand side of (13) is larger than the right hand side of (14), abstracting from the longer time horizon in the first case.

Equivalently, we show that the threshold with respect to F_j is larger in the second than in the first case. If A is entered first under sequential entry, the condition for entry equals
$\Pi^{seq} = \frac{1}{2k_A}\left[\rho_A \theta^h R_A(1+\delta) + \delta \rho_A \theta^h \left(\frac{\left(\rho_B^{+-} \theta^h R_B\right)^2}{2k_B} - F_B\right)\right]^2 - F_A \geq 0.$
Since the thresholds now also depend on characteristics in B, it is not possible to make a general statement. However, let us assume that both countries are identical and only differ in R_j, giving respective thresholds $R_A \geq \frac{\sqrt{2kF}}{\rho^{-+}\theta^h}$ (if entered as second under sequential entry) and $R_A \geq \frac{\sqrt{2kF} - \delta\rho\theta^h \left(\frac{\left(\rho^{+-}\theta^h R_B\right)^2}{2k} - F\right)}{\rho\theta^h(1+\delta)}$ (if entered as first under sequential entry). R_B still plays a role in determining the relevant thresholds. However, we can claim that if both countries are identical except their values of R_j and sequential entry is chosen, the one with a higher R_j is always entered first, completing the argument. To see this point take $\Delta\Pi^{seq}$ defined as the difference when A is entered first and when B is entered first under sequential entry. It equals
$\Delta\Pi^{seq} = \frac{1}{2k}\left[\rho\theta^h R_A(1+\delta) + \delta\rho\theta^h \left(\frac{\left(\rho^{+-}\theta^h R_B\right)^2}{2k} - F\right)\right]^2$
$-\frac{1}{2k}\left[\rho\theta^h R_B(1+\delta) + \delta\rho\theta^h \left(\frac{\left(\rho^{+-}\theta^h R_A\right)^2}{2k} - F\right)\right]^2$

$= \frac{1}{2k}\left\{\left[\rho\theta^h R_A(1+\delta)+\delta\rho\theta^h\left(\frac{\left(\rho^{+-}\theta^h R_B\right)^2}{2k}-F\right)\right]+\left[\rho\theta^h R_B(1+\delta)+\delta\rho\theta^h\left(\frac{\left(\rho^{+-}\theta^h R_A\right)^2}{2k}-F\right)\right]\right\}\times$
$\left\{\left[\rho\theta^h R_A(1+\delta)+\delta\rho\theta^h\left(\frac{\left(\rho^{+-}\theta^h R_B\right)^2}{2k}-F\right)\right]-\left[\rho\theta^h R_B(1+\delta)+\delta\rho\theta^h\left(\frac{\left(\rho^{+-}\theta^h R_A\right)^2}{2k}-F\right)\right]\right\}.$
As the first line of the previous expression is always positive, it is sufficient to look at the last line. It equals
$(R_A - R_B)\rho\theta^h\left[(1+\delta)-\delta\left(\frac{\left(\rho^{+-}\theta^h\right)^2}{2k}\right)(R_A+R_B)\right]$. Thus, it remains to show that the term in squared brackets is always positive. Then, the sign of $\Delta\Pi^{seq}$ is only determined by the sign of $(R_A - R_B)$.
Thus,
$\left[(1+\delta)-\delta\left(\rho^{+-}\theta^h\right)^2\frac{(R_A+R_B)}{2k}\right] \geq \left[(1+\delta)-\delta\left(\rho^{+-}\theta^h\right)^2\frac{2\max\{R_A,R_B\}}{2k}\right]$
$\geq \left[(1+\delta)-\delta\left(\rho^{+-}\theta^h\right)^2\frac{2\max\{R_A,R_B\}}{2\theta^h\max\{R_A,R_B\}(1+\delta)}\right]$
as $k_j \geq \theta^h R_j(1+\delta)$
$= \left[(1+\delta)-\delta\left(\rho^{+-}\right)^2\theta^h\frac{1}{(1+\delta)}\right] \geq 0$

Equivalently, we can show that if both countries are identical but only differ with respect to their fixed entry costs, the one with a higher level of F_j should be entered later.

Hypothesis R5: *If fixed entry costs in one country are ceteris paribus higher, it is more likely that this country is chosen as second under sequential entry. Furthermore, the relative profitability of sequential compared to simultaneous entry increases if the fixed entry costs in the second target country are larger. Thus, fixed costs should on average be higher for countries entered at later stages.*

Proof: First, we derive $\Delta\Pi^{Seq}$, the difference between profits under sequential entry when A and when B is chosen first. Here, $\frac{\partial\Delta\Pi^{seq}}{dF_B} = -\delta X_A^{AB}\rho_A\theta^h + 1 > 0$ and $\frac{\partial\Delta\Pi^{seq}}{dF_A} = -1 + \delta X_B^{BA}\rho_B\theta^h < 0$. Second, we derive the difference between profits under sequential and simultaneous entry and get $\frac{d\Delta\Pi}{dF_B} = -\delta X_A^{Seq}\rho_A\theta^h + 1 > 0$. Finally, it helps to establish that if fixed costs are very small, sequential entry can never be optimal. Note that when sequential entry is chosen, entry into B after a failure in A can not be optimal (otherwise, the firm could increase expected profits by choosing simultaneous entry). Thus, the belief, ρ_B^{--}, i.e., the belief in B after a failure in A must satisfy $\frac{\left(\rho_B^{--}\theta^h R_B\right)^2}{2k_B} - F_B \leq 0$. As $\rho_B^{--} > 0$, F_B needs to be sufficiently large to make this condition hold.

Appendix II – Correlated Learning

The ex-ante joint beliefs for being in one of the four potential states (θ^h,θ^h), $(\theta^h,0)$, $(0,\theta^h)$, or $(0,0)$ are characterized by the following Corollary.

Lemma A3: The ex-ante probabilities of being in state (θ_A,θ_B) equal
$\text{Prob}(\theta^h,\theta^h) \equiv p^{hh} = \left(r_B + (1-r_B)[r_{AB}\frac{\rho_A^0}{\rho_A} + (1-r_{AB})\rho_B^0]\right)\rho_A$
$\text{Prob}(\theta^h,0) \equiv p^{hl} = (1-r_B)\left(1 - [r_{AB}\frac{\rho_A^0}{\rho_A} + (1-r_{AB})\rho_B^0]\right)\rho_A$
$\text{Prob}(0,\theta^h) \equiv p^{lh} = \left(r_B + (1-r_B)(1-r_{AB})\rho_B^0\right)(1-\rho_A)$
$\text{Prob}(0,0) \equiv p^{ll} = (1-r_B)\left(1-(1-r_{AB})\rho_B^0\right)(1-\rho_A)$

Proof: As $\text{Prob}(\theta^h, \theta^h) = \text{Prob}(\theta_B^h \mid \theta_A^h)\text{Prob}(\theta_A^h)$, we need $\text{Prob}(\theta_B^h \mid \theta_A^h)$. Taking $\rho_B = r_B + (1 - r_B)[r_{AB}\rho_A^0 + (1 - r_{AB})\rho_B^0]$, we obtain
$\text{Prob}(\theta_B^h \mid \theta_A^h) = r_B + (1 - r_B)[r_{AB}\text{E}[\rho_A^0 \mid \theta_A^h] + (1 - r_{AB})\rho_B^0]$.
Bayes' rule can be used to compute $\text{E}[\rho_A^0 \mid \theta_A^h]$, and we get
$\text{E}[\rho_A^0 \mid \theta_A^h] = \frac{\rho_A^0(r_A+(1-r_A)\cdot 1)}{\rho_A^0(r_A+(1-r_A)\cdot 1)+(1-\rho_A^0)(r_A+(1-r_A)\cdot 0)} = \frac{\rho_A^0}{\rho_A}$.
Equivalently,
$\text{E}[\rho_A^0 \mid \theta_A^l] = \frac{\rho_A^0[1-(r_A+(1-r_A))]}{\rho_A^0[1-(r_A+(1-r_A))]+(1-\rho_A^0)[1-(r_A+(1-r_A)\cdot 0)]} = 0$, proving the Corollary.

Updating occurs for each of the potential outcome realizations $(Y_A, Y_B) \in \{(R_A, R_B),\ (R_A, 0),\ (0, R_B),\ (0, 0)\}$:

1. $(Y_A, Y_B) = (R_A, R_B)$

   $\rho_A^{++} = \rho_B^{++} = 1$

2. $(Y_A, Y_B) = (R_A, 0)$

   $\rho_A^{+-} = 1$

   $\rho_B^{+-} = (p^{hh})^{+-} + (p^{lh})^{+-} = \frac{p^{hh}X_A\theta^h(1-X_B\theta^h)}{p^{hh}X_A\theta^h(1-X_B\theta^h)+p^{hl}X_A\theta^h} + 0 = \frac{p^{hh}(1-X_B\theta^h)}{(1-X_B\theta_B^h)p^{hh}+p^{hl}}$

3. $(Y_A, Y_B) = (0, R_B)$

   $\rho_A^{-+} = \frac{p^{hh}(1-X_A\theta^h)}{(1-X_A\theta^h)p^{hh}+p^{lh}}$

   $\rho_B^{-+} = 1$

4. $(Y_A, Y_B) = (0, 0)$

   $\rho_A^{--} = \frac{p^{hh}(1-X_A\theta^h)(1-X_B\theta^h)+p^{hl}(1-X_A\theta^h)}{(1-X_A\theta^h)(p^{hh}(1-X_B\theta^h)+p^{hl})+(p^{lh}(1-X_B\theta^h)+p^{ll})}$

   $\rho_B^{--} = \frac{p^{hh}(1-X_A\theta^h)(1-X_B\theta^h)+p^{lh}(1-X_B\theta^h)}{(1-X_A\theta^h)(p^{hh}(1-X_B\theta^h)+p^{hl})+(p^{lh}(1-X_B\theta^h)+p^{ll})}$

Observing (R_A, R_B), both beliefs jump to 1, i.e., $\rho_A^{++} = \rho_B^{++} = 1$. If a success is only realized in country A but not in B, implying $(R_A, 0)$, $\rho_A^{+-} = 1$, while $\rho_B^{+-} = (p^{hh})^{+-} + (p^{lh})^{+-} = \frac{p^{hh}X_A\theta^h(1-X_B\theta^h)}{p^{hh}X_A\theta^h(1-X_B\theta^h)+p^{hl}X_A\theta^h} + 0 = \frac{p^{hh}(1-X_B\theta^h)}{(1-X_B\theta_B^h)p^{hh}+p^{hl}}$. Conversely, the realization $(0, R_B)$ gives $\rho_A^{-+} = \frac{p^{hh}(1-X_A\theta^h)}{(1-X_A\theta^h)p^{hh}+p^{lh}}$ and $\rho_B^{-+} = 1$.
Finally, after a double failure, beliefs fall to
$\rho_A^{--} = \frac{p^{hh}(1-X_A\theta^h)(1-X_B\theta^h)+p^{hl}(1-X_A\theta^h)}{(1-X_A\theta^h)(p^{hh}(1-X_B\theta^h)+p^{hl})+(p^{lh}(1-X_B\theta^h)+p^{ll})}$ and
$\rho_B^{--} = \frac{p^{hh}(1-X_A\theta^h)(1-X_B\theta^h)+p^{lh}(1-X_B\theta^h)}{(1-X_A\theta^h)(p^{hh}(1-X_B\theta^h)+p^{hl})+(p^{lh}(1-X_B\theta^h)+p^{ll})}$.

The case we are interested in is where entry initially occurs only in one country, say A. This is covered by setting $X_B = 0$ and taking into account that a "failure" there occurs with probability 1. If the MNE only enters A and observes a success, the belief in B becomes
$\rho_B^{+-} = \left(r_B + (1 - r_B)[r_{AB}\frac{\rho_A^0}{\rho_A} + (1 - r_{AB})\rho_B^0]\right) > \rho_B$. Recall that ρ_B^{-+} is not of interest as B is never entered after a failure in A.

Starting out by investing in B and observing a success there yields
$\rho_A^{-+} = \left(r_B + (1 - r_B)[r_{AB}\frac{\rho_A^0}{\rho_A} + (1 - r_{AB})\rho_B^0]\right)\frac{\rho_A}{\rho_B} = \rho_B^{+-}\frac{\rho_A}{\rho_B}$.

Finally, beliefs also follow a martingale here; to see this, take the expected change in the belief in A for arbitrary investment levels X_A and X_B. Keeping in mind that $\rho_A = p^{hh} + p^{hl}$, we have
$\mathrm{E}[\rho_{At+1} \mid \rho_{At}] = p^{hh}X_A X_B\theta^h\theta^h\rho_A^{++} + \left[p^{hh}X_A\theta^h(1 - X_B\theta^h) + p^{hl}X_A\theta^h\right]\rho_A^{+-}$
$+ \left[p^{hh}(1 - X_A\theta^h)X_B\theta^h + p^{lh}X_B\theta^h\right]\rho_A^{-+}$
$+ \left[p^{hh}(1 - X_A\theta^h)(1 - X_B\theta^h) + p^{hl}(1 - X_A\theta^h) + p^{lh}(1 - X_B\theta^h) + p^{ll}\right]\rho_A^{--}$
$= X_A\theta^h\rho_A + (1 - X_A\theta^h)(p^{hh} + p^{hl}) = \rho_A$.

Appendix III – Adjustable Capacity

Assume that the capacity can be adjusted upwards in the second period. We assume that the cost function is a function of the total capacity, i.e., the marginal investment cost for the first capacity unit in period 2 equals the marginal cost for the last capacity unit in the first period. Generally, the option to adjust the capacity later will allow the firms to increase investments in period 2 if a success was observed in $t = 1$. After a failure, nothing changes. Obviously, the capacity in the first period will be smaller than without the adjustment option. What we show here is that sequential entry is still associated with a higher investment level in country A. All other main results will hold.

Isolated Entry

As the MNE only considers entry into one country, we can omit the country subscript. Define X_1 as the first period and $X_2 = X_1 + \Delta X$ as the total second-period-capacity following a success. Furthermore, define $\Delta K(X_2) = K(X_2) - K(X_1)$ as the costs of the capacity increase.
We first have to determine the capacity adjustment in the second period after a success has been observed. Generally, expected profits then equal $X_2R\theta^h - \Delta K(X_2) = X_2R\theta^h - \left(k\frac{X_2^2}{2} - k\frac{X_1^2}{2}\right)$. The gives an optimal capacity level $X_2 = \frac{R\theta^h}{k}$ and implies second-period profits $\Pi_2^+ = \frac{\left(R\theta^h\right)^2}{2k} + k\frac{X_1^2}{2}$. As $\rho^- < \rho$, the capacity does not get adjusted after a failure, yielding expected second-period profits $\Pi_2^- = \rho_2^- X_1\theta^h R = \frac{\rho(1-X_1\theta^h)}{\rho(1-X_1\theta^h)+(1-\rho)}X_1\theta^h R$.

This allows us to state

Lemma A4: Assume the capacity can be adjusted in the second period. Then, the first-period investment level under isolated entry equals $X_1 = \frac{\left(k+2\delta\rho\theta^h R\theta^h\right) - \sqrt{(k+2\delta\rho\theta^h R\theta^h)^2 - 3\delta\rho\theta^h\rho\theta^h R\left(\delta R(\theta^h)^2 + 2k(1+\delta)\right)}}{3\delta\rho\theta^h k}$.

Proof: Total expected profits are equal to $\Pi^{iso} = X_1\rho R\theta^h - K(X_1) + \delta\left(\rho X_1\theta^h\Pi_2^+ + (1 - \rho X_1\theta^h)\Pi_2^-\right)$.
Substituting allows us to state the first-order condition:
$\rho R\theta^h - kX_1 + \delta\left(\rho\theta^h\frac{\left(R\theta^h\right)^2}{2k} + 3\rho\theta^h k\frac{X_1^2}{2} + \rho\theta^h R - 2\rho X_1\theta^h R\theta^h\right) = 0$, with
$X_1 = \frac{\left(k+2\delta\rho\theta^h R\theta^h\right) \pm \sqrt{(k+2\delta\rho\theta^h R\theta^h)^2 - 3\delta\rho\theta^h\rho\theta^h R\left(\delta R(\theta^h)^2 + 2k(1+\delta)\right)}}{3\delta\rho\theta^h k}$.

The second order condition then guarantees that the stated level is a maximum, while the other level constitutes a minimum.

Sequential Entry

For sequential entry, we can show that the resulting capacity in A is higher than under isolated entry. Under sequential entry, the situation in country B is identical to the case without the option to adjust one's capacity; so $X_B^{Seq} = \frac{\rho_B^{+-}\theta^h R_B}{k_B}$, which yields expected profits in B, $\Pi_B^{seq} = \frac{\left(\rho_B^{+-}\theta^h R_B\right)^2}{2k_B} - F_B$. Furthermore, the considerations in A in the second period are equivalent to isolated entry. A success yields a second-period capacity $X_{2A} = \frac{R_A\theta^h}{k_A}$ associated with profits $\Pi_{2A}^+ = \frac{\left(R_A\theta^h\right)^2}{2k_A} + k_A\frac{X_{1A}^2}{2}$. A failure leaves the capacity unchanged and gives second-period profits $\Pi_2^- = \rho_{2A}^- X_{1A}\theta^h R_A = \frac{\rho_A(1-X_{1A}\theta^h)}{\rho_A(1-X_{1A}\theta^h)+(1-\rho_A)}X_{1A}\theta^h R_A$. Now we can state

Lemma A5: Assume the capacity can be adjusted in the second period. Then, the first-period investment level in the first country entered under sequential entry equals

$$X_{1A}^{Seq} = \frac{2\delta\rho_A\theta^h\theta^h R_A + k_A - \sqrt{(2\delta\rho_A\theta^h\theta^h R_A + k_A)^2 - 4\frac{3}{2}\delta\rho_A\theta^h k_A\rho_A\theta^h\left(R_A(1+\delta) + \delta\frac{\left(R_A\theta^h\right)^2}{2k_A} + \delta\Pi_B^{seq}\right)}}{3\delta\rho_A\theta^h k_A}.$$

Proof: Total profits equal

$$\Pi^{Seq} = X_{1A}^{Seq}\rho_A\theta^h R_A - k_A\frac{(X_{1A}^{Seq})^2}{2} - F_A + \delta X_{A1}^{Seq}\rho_A\theta^h\left(\frac{\left(R_A\theta^h\right)^2}{2k_A} + k_A\frac{X_{1A}^2}{2} + \Pi_B^{seq}\right)$$

$+\delta\rho_A(1 - X_{1A}\theta^h)X_{1A}\theta^h R_A$, which implies the first order condition

$$(X_{A1}^{Seq})^2\frac{3}{2}\delta\rho_A\theta^h k_A - X_{1A}^{Seq}\left(2\delta\rho_A\theta^h\theta^h R_A + k_A\right) + \rho_A\theta^h\left(R_A(1+\delta) + \delta\frac{\left(R_A\theta^h\right)^2}{2k_A} + \delta\Pi_B^{seq}\right) = 0$$

and potential capacity levels

$$X_{1A}^{Seq} = \frac{2\delta\rho_A\theta^h\theta^h R_A + k_A \pm \sqrt{(2\delta\rho_A\theta^h\theta^h R_A + k_A)^2 - 4\frac{3}{2}\delta\rho_A\theta^h k_A\rho_A\theta^h\left(R_A(1+\delta) + \delta\frac{\left(R_A\theta^h\right)^2}{2k_A} + \delta\Pi_B^{seq}\right)}}{3\delta\rho_A\theta^h k_A}.$$

The second order condition guarantees that the stated level is a maximum. This allows us to constitute

Lemma A6: Assume the capacity can be adjusted in the second period. Then, the first-period investment level in the first country entered under sequential entry is higher there than when this country was entered in isolation.

Proof: $X_{1A}^{seq} \geq X_{1A}^{iso}$ is satisfied as long as $2k_A\delta\Pi_B^{seq} > 0$, which obviously is the case.

Simultaneous Entry

With adjustable capacity, we can not always state that simultaneous and isolated entry lead to identical outcomes. Now, a success in A could induce a capacity adjustment in B, even after a failure in B was observed. Thus, the capacity under simultaneous entry might be higher than under isolated entry. Yet, it will never be as high as under sequential entry. The additional value of more capacity increases in ρ_B^{+-}, while ρ_B^{+-} decreases the capacity in B. Thus, ρ_B^{+-} is highest under sequential entry where $X_{1B} = 0$ in the first period. As the higher capacity in A under sequential entry is induced by the extra profits expected in B, it is obvious that sequential entry is always associated with a higher capacity than simultaneous entry. The reason is that these expected extra profits are highest when no previous investments in B occured (otherwise, the updated belief would be lower).

Appendix IV – Tables and Figures

Table 1: Descriptive Statistics (Mean Values)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Dependent Variable:					
Location Decision	0.013	0.012	0.011	0.012	0.013
Host-country Variables:					
Tax	0.301	0.292	0.291	0.282	0.276
log GDP	25.750	25.572	25.575	25.762	25.882
InvestFree	60.860	59.292	58.100	57.850	57.518
InvestCost	26.782	28.383	31.163	28.539	21.810
CPI	4.756	4.676	4.562	4.695	4.895
Bilateral Variables:					
log Distance to parent	8.103	8.131	8.158	8.106	8.025
log Distance to 1st		8.407	8.427	8.373	8.273
log Distance to 2nd			8.461	8.418	8.312
log Distance to 3rd				8.441	8.400
log Distance to 4th					8.498
Border to parent	0.104	0.093	0.090	0.097	0.102
Border to 1st		0.032	0.031	0.033	0.037
Border to 2nd			0.030	0.032	0.039
Border to 3rd				0.034	0.034
Border to 4th					0.041
Language same as parent	0.026	0.023	0.023	0.024	0.026
Language same as 1st		0.137	0.132	0.129	0.111
Language same as 2nd			0.137	0.129	0.095
Language same as 3rd				0.124	0.137
Language same as 4th					0.133
Colony of parent	0.025	0.023	0.023	0.024	0.026
Colony of 1st		0.052	0.048	0.048	0.223
Colony of 2nd			0.049	0.043	0.198
Colony of 3rd				0.048	0.194
Colony of 4th					0.210
Same country as 1st		0.012	0.011	0.014	0.017
Same country as 2nd			0.010	0.013	0.016
Same country as 3rd				0.011	0.007
Same country as 4th					0.013
GTA with parent	0.221	0.199	0.206	0.231	0.261
GTA with 1st		0.173	0.181	0.202	0.242
GTA with 2nd			0.192	0.201	0.262
GTA with 3rd				0.196	0.241
GTA with 4th					0.179
Observations	*1,164,529*	*402,359*	*199,168*	*90,716*	*74,876*

Table 2: Sequential Location Decision (Basic Results)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Host-country Variables:					
Tax	-1.626***	-1.619***	-1.613***	-3.346***	0.742
	(0.191)	(0.343)	(0.468)	(0.698)	(0.704)
log GDP	0.836***	0.757***	0.664***	0.749***	0.524***
	(0.010)	(0.016)	(0.022)	(0.032)	(0.031)
InvestFree	0.017***	0.010***	0.011***	0.009***	0.018***
	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)
InvestCost	-0.008***	-0.010***	-0.012***	-0.010***	-0.009***
	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
CPI	0.042***	0.003	-0.002	-0.038	-0.021
	(0.006)	(0.012)	(0.017)	(0.026)	(0.028)
Bilateral Variables:					
log Distance to parent	-0.522***	-0.251***	-0.033	0.011	0.315***
	(0.016)	(0.030)	(0.043)	(0.069)	(0.076)
log Distance to 1st		-0.538***	-0.358***	-0.217***	-0.267***
		(0.021)	(0.030)	(0.050)	(0.063)
log Distance to 2nd			-0.347***	-0.216***	-0.148***
			(0.030)	(0.047)	(0.057)
log Distance to 3rd				-0.367***	-0.455***
				(0.048)	(0.058)
log Distance to 4th					-0.272***
					(0.044)
Border to parent	0.535***	0.371***	0.318***	0.474***	0.272**
	(0.028)	(0.051)	(0.074)	(0.108)	(0.114)
Border to 1st		-0.156**	0.016	-0.130	-0.568***
		(0.059)	(0.084)	(0.152)	(0.180)
Border to 2nd			0.199**	-0.064	0.352**
			(0.083)	(0.142)	(0.146)
Border to 3rd				-0.078	0.630***
				(0.141)	(0.148)
Border to 4th					-0.258*
					(0.150)
					continued

Table 2: Sequential Location Decision (Basic Results)
(continued)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Language same as parent	0.378***	0.251***	0.297***	0.031	-0.486***
	(0.033)	(0.062)	(0.092)	(0.150)	(0.167)
Language same as 1st		0.087	-0.321***	-0.129	0.424***
		(0.057)	(0.087)	(0.132)	(0.154)
Language same as 2nd			-0.276***	-0.059	-0.438***
			(0.087)	(0.130)	(0.166)
Language same as 3rd				-0.188	-0.554***
				(0.132)	(0.149)
Language same as 4th					0.163
					(0.127)
Colony of parent	0.361***	0.339***	0.624***	-0.208	0.565***
	(0.047)	(0.085)	(0.118)	(0.196)	(0.204)
Colony of 1st		0.429***	0.611***	0.273*	-0.113
		(0.066)	(0.096)	(0.157)	(0.198)
Colony of 2nd			0.653***	0.247*	0.132
			(0.095)	(0.150)	(0.205)
Colony of 3rd				0.525***	-0.076
				(0.140)	(0.193)
Colony of 4th					0.019
					(0.146)
Same country as 1st		-0.226**	-0.282*	0.105	-0.415
		(0.100)	(0.162)	(0.256)	(0.340)
Same country as 2nd			-0.551***	-0.006	-0.061
			(0.170)	(0.223)	(0.251)
Same country as 3rd				-0.107	-0.680*
				(0.249)	(0.370)
Same country as 4th					-0.214
					(0.273)
					continued

Table 2: Sequential Location Decision (Basic Results)
(concluded)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
GTA with parent	0.073**	-0.221***	-0.188**	-0.543***	-0.088
	(0.029)	(0.055)	(0.081)	(0.128)	(0.149)
GTA with 1st		0.535***	0.333***	0.330***	0.153
		(0.051)	(0.078)	(0.119)	(0.148)
GTA with 2nd			0.410***	0.350***	0.223*
			(0.078)	(0.121)	(0.126)
GTA with 3rd				0.316***	-0.310**
				(0.116)	(0.136)
GTA with 4th					0.471***
					(0.120)
Pseudo R2	0.2258	0.2819	0.2706	0.2553	0.2255
Observations	1,164,529	402,359	199,168	90,716	74,876
Location decisions	15,165	4,694	2,249	1,099	958
Years between decisions		1.999	1.611	1.478	1.326

Notes: Conditional logit model. If the MNE has chosen two (or more) locations in phase $p-1$, we use the greater investment (measured in fixed assets) as reference for the investment in phase p. Robust standard errors reported in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively. *Location decisions* reports the actual number of location decisions made (Location decision = 1). *Years between decisions* are the average years between the respective (sequential) location decisions made by the multinationals in the sample. Control variables are taken from different sources. *Tax* is the statutory tax rate of a host country. The tax data is collected from databases provided by the International Bureau of Fiscal Documentation (IBFD) and tax surveys provided by Ernst&Young, PwC, and KPMG. *log GDP* measures the real GDP at constant U.S. dollars of the year 2000 and is taken from the World Bank's World Development Indicators 2009. The investment freedom index *InvestFree* is taken from the Heritage Indicators database. The index can take on values between 0 and 100; higher values are associated with more investment freedom. *InvestCost* is from World Bank's Doing Business Database and measures the cost of starting a business relative to income per capita. *CPI* (Corruption Perception Index) is published annually by Transparency International. It ranks countries in terms of perceived levels of corruption, as determined by expert assessments and opinion surveys. The scores range from 10 (country perceived as virtually corruption free) to 0 (country perceived as almost totally corrupt). *log Distance* is the log of the distance (in kilometer) between the most populated cities in the host country and the country of the previous investment. As to the bilateral variables for the first investment, we use Germany as the reference country. *Border* is a common border indicator, *Language* a common language indicator, *Colony* a former colony indicator, *Same country* a dummy indicating whether the host country and the country of the previous investment are the same. *GTA* is an indicator for the existence of a general trade agreement between the host country and the country of the previous investment. The bilateral variables are either taken from the Centre d'Études Prospectives et d'Informations Internationales (*log Distance, Border, Language, Colony, Same country*), or from the World Trade Organization (*GTA*).

Table 3: Marginal Effects of Continuous Variables

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Host-country Variables:					
Tax	-.0200	-.0173	-.0168	-.0372	.0088
log GDP	.0103	.0081	.0069	.0083	.0062
InvestFree	.0002	.0001	.0001	.0001	.0002
InvestCost	-.0001	-.0001	-.0001	-.0001	-.0002
CPI	.0005	.0000	-.0000	-.0004	-.0001
Bilateral Variables:					
log Distance to parent	-.0064	-.0027	-.0003	.0001	.0037
log Distance to 1st		-.0057	-.0037	-.0024	-.0032
log Distance to 2nd			-.0036	-.0024	-.0017
log Distance to 3rd				-.0041	-.0054
log Distance to 4th					-.0032

Notes: Marginal effects correspond to Table 2 (Basic Results). The values shown are the average marginal effects. The latter are obtained as $p_j(x)/\partial x_{jk} = p_j(x)[1 - p_j(x)]\beta_k$, where p_j is the response probability given by Equation 11.

Table 4: Sequential Location Decision (Sensitivity I)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Host-country Variables:					
Tax	-1.279***	-1.599***	-1.637***	-3.331***	0.909
	(0.193)	(0.343)	(0.469)	(0.698)	(0.708)
log GDP	0.584***	0.667***	0.606***	0.764***	0.617***
	(0.012)	(0.020)	(0.027)	(0.042)	(0.041)
InvestFree	0.016***	0.010***	0.011***	0.008***	0.020***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
InvestCost	-0.006***	-0.009***	-0.011***	-0.010***	-0.009***
	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)
CPI	0.011*	-0.009	-0.010	-0.035	-0.017
	(0.006)	(0.012)	(0.017)	(0.026)	(0.027)
Bilateral Variables:					
log Distance to parent	-0.469***	-0.232***	-0.021	0.008	0.299***
	(0.016)	(0.029)	(0.043)	(0.069)	(0.075)
log Distance to 1st		-0.530***	-0.354***	-0.219***	-0.283***
		(0.020)	(0.031)	(0.050)	(0.061)
log Distance to 2nd			-0.345***	-0.217***	-0.151***
			(0.029)	(0.046)	(0.056)
log Distance to 3rd				-0.368***	-0.461***
				(0.047)	(0.057)
log Distance to 4th					-0.271***
					(0.043)
Border to parent	0.313***	0.293***	0.268***	0.485***	0.343***
	(0.028)	(0.052)	(0.076)	(0.109)	(0.115)
Border to 1st		-0.148**	0.031	-0.133	-0.564***
		(0.059)	(0.084)	(0.151)	(0.179)
Border to 2nd			0.194**	-0.061	0.341**
			(0.083)	(0.142)	(0.146)
Border to 3rd				-0.078	0.628***
				(0.141)	(0.146)
Border to 4th					-0.275*
					(0.150)

continued

Table 4: Sequential Location Decision (Sensitivity I)
(continued)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Language same as parent	0.326***	0.254***	0.304***	0.027	-0.515***
	(0.032)	(0.061)	(0.092)	(0.150)	(0.169)
Language same as 1st		0.092	-0.321***	-0.132	0.361**
		(0.056)	(0.087)	(0.132)	(0.157)
Language same as 2nd			-0.281***	-0.059	-0.419**
			(0.086)	(0.130)	(0.166)
Language same as 3rd				-0.186	-0.557***
				(0.132)	(0.148)
Language same as 4th					0.157
					(0.126)
Colony of parent	0.441***	0.355***	0.627***	-0.205	0.601***
	(0.046)	(0.084)	(0.117)	(0.196)	(0.206)
Colony of 1st		0.378***	0.587***	0.282*	0.017
		(0.067)	(0.097)	(0.157)	(0.202)
Colony of 2nd			0.636***	0.252*	0.192
			(0.095)	(0.150)	(0.202)
Colony of 3rd				0.531***	-0.041
				(0.139)	(0.191)
Colony of 4th					0.055
					(0.147)
Same country as 1st		-0.213**	-0.281*	0.105	-0.450
		(0.101)	(0.162)	(0.255)	(0.340)
Same country as 2nd			-0.549***	-0.008	-0.063
			(0.171)	(0.222)	(0.249)
Same country as 3rd				-0.108	-0.693*
				(0.248)	(0.368)
Same country as 4th					-0.218
					(0.271)

continued

Table 4: Sequential Location Decision (Sensitivity I)
(concluded)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
GTA with parent	0.063**	-0.214***	-0.179**	-0.546***	-0.135
	(0.028)	(0.054)	(0.080)	(0.128)	(0.148)
GTA with 1st		0.519	0.325***	0.332***	0.183
		(0.051)	(0.078)	(0.119)	(0.146)
GTA with 2nd			0.401***	0.351***	0.222*
			(0.078)	(0.121)	(0.124)
GTA with 3rd				0.317***	-0.284**
				(0.115)	(0.134)
GTA with 4th					0.472***
					(0.118)
StockInv	0.428***	0.173***	0.117***	-0.034	-0.229***
	(0.015)	(0.028)	(0.041)	(0.060)	(0.077)
Pseudo R2	0.2307	0.2827	0.2710	0.2553	0.2265
Observations	1,164,529	402,359	199,168	90,716	74,876
Location decisions	15,165	4,694	2,249	1,099	958
Years between decisions		1.999	1.611	1.478	1.326

Notes: Conditional logit model. *Sensitivity I*: All estimations additionally include the stock of all German investments in country j prior to firm i's investment, *StockInv.* If the MNE has chosen two (or more) locations in phase $p-1$, we use the greater investment (measured in fixed assets) as reference for the investment in phase p. Robust standard errors reported in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively. *Location decisions* reports the actual number of location decisions made (Location decision = 1). *Years between decisions* are the average years between the respective (sequential) location decisions made by the multinationals in the sample. Control variables are taken from different sources. *Tax* is the statutory tax rate of a host country. The tax data is collected from databases provided by the International Bureau of Fiscal Documentation (IBFD) and tax surveys provided by Ernst&Young, PwC, and KPMG. *log GDP* measures the real GDP at constant U.S. dollars of the year 2000 and is taken from the World Bank's World Development Indicators 2009. The investment freedom index *InvestFree* is taken from the Heritage Indicators database. The index can take on values between 0 and 100; higher values are associated with more investment freedom. *InvestCost* is from World Bank's Doing Business Database and measures the cost of starting a business relative to income per capita. *CPI* (Corruption Perception Index) is published annually by Transparency International. It ranks countries in terms of perceived levels of corruption, as determined by expert assessments and opinion surveys. The scores range from 10 (country perceived as virtually corruption free) to 0 (country perceived as almost totally corrupt). *log Distance* is the log of the distance (in kilometer) between the most populated cities in the host country and the country of the previous investment. As to the bilateral variables for the first investment, we use Germany as the reference country. *Border* is a common border indicator, *Language* a common language indicator, *Colony* a former colony indicator, *Same country* a dummy indicating whether the host country and the country of the previous investment are the same. *GTA* is an indicator for the existence of a general trade agreement between the host country and the country of the previous investment. The bilateral variables are either taken from the Centre d'Études Prospectives et d'Informations Internationales (*log Distance, Border, Language, Colony, Same country*), or from the World Trade Organization (*GTA*).

Table 5: Sequential Location Decision (Sensitivity II)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Host-country Variables:					
Tax	-1.626***	-1.622***	-1.498***	-3.418***	0.688
	(0.190)	(0.343)	(0.469)	(0.689)	(0.725)
log GDP	0.836***	0.757***	0.663***	0.754***	0.536***
	(0.009)	(0.016)	(0.022)	(0.032)	(0.031)
InvestFree	0.017***	0.010***	0.009***	0.008***	0.015***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
InvestCost	-0.008***	-0.010***	-0.012***	-0.011***	-0.008***
	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)
CPI	0.042***	0.003	-0.007	-0.042	-0.023
	(0.006)	(0.011)	(0.016)	(0.026)	(0.028)
Bilateral Variables:					
log Distance to parent	-0.521***	-0.251***	-0.042	0.047	0.352***
	(0.016)	(0.029)	(0.044)	(0.069)	(0.078)
log Distance to 1st		-0.538***	-0.375***	-0.228***	-0.249***
		(0.020)	(0.030)	(0.051)	(0.066)
log Distance to 2nd			-0.296***	-0.219***	-0.246***
			(0.032)	(0.048)	(0.057)
log Distance to 3rd				-0.358***	-0.363***
				(0.046)	(0.058)
log Distance to 4th					-0.297***
					(0.048)
Border to parent	0.535***	0.371***	0.281***	0.462***	0.315***
	(0.027)	(0.051)	(0.074)	(0.108)	(0.114)
Border to 1st		-0.161***	0.143*	-0.139	-0.509***
		(0.059)	(0.086)	(0.151)	(0.176)
Border to 2nd			-0.083	-0.025	0.208
			(0.098)	(0.142)	(0.155)
Border to 3rd				0.006	0.641***
				(0.137)	(0.145)
Border to 4th					-0.302**
					(0.147)

continued

Table 5: Sequential Location Decision (Sensitivity II)
(continued)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Language same as parent	0.377***	0.251***	0.339***	0.052	-0.429***
	(0.033)	(0.062)	(0.093)	(0.148)	(0.166)
Language same as 1st		0.092	-0.373***	-0.113	0.389***
		(0.056)	(0.089)	(0.134)	(0.149)
Language same as 2nd			-0.202**	-0.188	-0.605***
			(0.084)	(0.130)	(0.164)
Language same as 3rd				-0.191	-0.483***
				(0.138)	(0.146)
Language same as 4th					0.395***
					(0.118)
Colony of parent	0.361***	0.339***	0.659***	-0.146	0.509**
	(0.046)	(0.085)	(0.119)	(0.195)	(0.204)
Colony of 1st		0.427***	0.716***	0.346**	-0.194
		(0.066)	(0.101)	(0.160)	(0.186)
Colony of 2nd			0.443***	0.224	-0.007
			(0.096)	(0.150)	(0.201)
Colony of 3rd				0.382	0.234
				(0.146)	(0.190)
Colony of 4th					-0.248*
					(0.149)
Same country as 1st		-0.225**	-0.452***	0.046	-0.448
		(0.100)	(0.165)	(0.252)	(0.341)
Same country as 2nd			-0.084	0.347	0.417
			(0.162)	(0.222)	(0.255)
Same country as 3rd				-0.220	-0.606**
				(0.249)	(0.315)
Same country as 4th					-0.104
					(0.275)

continued

Table 5: Sequential Location Decision (Sensitivity II)
(concluded)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
GTA with parent	0.073***	-0.220***	-0.186**	-0.574***	-0.072
	(0.028)	(0.054)	(0.082)	(0.130)	(0.155)
GTA with 1st		0.534***	0.309***	0.333***	0.206
		(0.051)	(0.078)	(0.122)	(0.151)
GTA with 2nd			0.418***	0.299**	0.209
			(0.081)	(0.122)	(0.133)
GTA with 3rd				0.364***	-0.107
				(0.115)	(0.137)
GTA with 4th					0.152
					(0.132)
Pseudo R2	0.2258	0.2819	0.2646	0.2564	0.2271
Observations	1,164,529	402,240	198,531	91,212	75,348
Location decisions	15,165	4,693	2,242	1,105	964
Years between decisions		1.998	1.613	1.485	1.332

Notes: Conditional logit model. *Sensitivity II*: If the MNE has chosen two (or more) locations in phase $p-1$, we use the greater investment (measured in *total assets* rather than in fixed assets) as reference for the investment in phase p. Robust standard errors reported in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively. *Location decisions* reports the actual number of location decisions made (Location decision = 1). *Years between decisions* are the average years between the respective (sequential) location decisions made by the multinationals in the sample. Control variables are taken from different sources. *Tax* is the statutory tax rate of a host country. The tax data is collected from databases provided by the International Bureau of Fiscal Documentation (IBFD) and tax surveys provided by Ernst&Young, PwC, and KPMG. *log GDP* measures the real GDP at constant U.S. dollars of the year 2000 and is taken from the World Bank's World Development Indicators 2009. The investment freedom index *InvestFree* is taken from the Heritage Indicators database. The index can take on values between 0 and 100; higher values are associated with more investment freedom. *InvestCost* is from World Bank's Doing Business Database and measures the cost of starting a business relative to income per capita. *CPI* (Corruption Perception Index) is published annually by Transparency International. It ranks countries in terms of perceived levels of corruption, as determined by expert assessments and opinion surveys. The scores range from 10 (country perceived as virtually corruption free) to 0 (country perceived as almost totally corrupt). *log Distance* is the log of the distance (in kilometer) between the most populated cities in the host country and the country of the previous investment. As to the bilateral variables for the first investment, we use Germany as the reference country. *Border* is a common border indicator, *Language* a common language indicator, *Colony* a former colony indicator, *Same country* a dummy indicating whether the host country and the country of the previous investment are the same. *GTA* is an indicator for the existence of a general trade agreement between the host country and the country of the previous investment. The bilateral variables are either taken from the Centre d'Études Prospectives et d'Informations Internationales (*log Distance, Border, Language, Colony, Same country*), or from the World Trade Organization (*GTA*).

Table 6: Sequential Location Decision (Sensitivity III)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Host-country Variables:					
Tax	-1.626***	-1.623***	-2.346***	-3.411**	1.710
	(0.191)	(0.343)	(0.795)	(1.678)	(2.936)
log GDP	0.836***	0.758***	0.760***	0.794***	0.525***
	(0.010)	(0.016)	(0.036)	(0.085)	(0.112)
InvestFree	0.017***	0.010***	0.010***	0.009	0.027**
	(0.001)	(0.001)	(0.003)	(0.006)	(0.013)
InvestCost	-0.008***	-0.010***	-0.012***	-0.012*	-0.007
	(0.001)	(0.001)	(0.004)	(0.007)	(0.012)
CPI	0.042***	0.003	-0.046*	-0.067	-0.160
	(0.006)	(0.011)	(0.027)	(0.056)	(0.117)
Bilateral Variables:					
log Distance to parent	-0.521***	-0.252***	-0.099	0.144	0.634*
	(0.016)	(0.030)	(0.069)	(0.167)	(.337)
log Distance to 1st		-0.538***	-0.429***	-0.280**	-1.064***
		(0.020)	(0.045)	(0.114)	(0.227)
log Distance to 2nd			-0.348***	-0.368***	0.243
			(0.049)	(0.112)	(0.315)
log Distance to 3rd				-0.248**	-0.732***
				(0.125)	(0.150)
log Distance to 4th					-0.271
					(0.222)
Border to parent	0.535***	0.372***	0.345***	0.240	0.739
	(0.027)	(0.051)	(0.122)	(0.284)	(0.534)
Border to 1st		-0.160***	-0.109	0.517	-0.645
		(0.059)	(0.145)	(0.349)	(0.846)
Border to 2nd			-0.163	-0.749*	0.896
			(0.156)	(0.400)	(0.637)
Border to 3rd				0.320	0.383
				(0.316)	(0.571)
Border to 4th					0.292
					(0.638)
					continued

Table 6: Sequential Location Decision (Sensitivity III)
(continued)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
Language same as parent	0.377***	0.251***	0.225	0.601	-0.974
	(0.033)	(0.061)	(0.155)	(0.380)	(0.808)
Language same as 1st		0.091	-0.278**	-1.018***	-0.215
		(0.056)	(0.137)	(0.335)	(0.673)
Language same as 2nd			0.163	-0.023	0.122
			(0.130)	(0.340)	(0.623)
Language same as 3rd				-0.109	0.337
				(0.340)	(0.583)
Language same as 4th					-1.121
					(0.719)
Colony of parent	0.361***	0.339***	0.650***	0.686*	0.434
	(0.046)	(0.085)	(0.181)	(0.409)	(0.947)
Colony of 1st		0.427***	0.437***	0.844**	0.851
		(0.066)	(0.162)	(0.350)	(0.670)
Colony of 2nd			0.320**	0.314	-1.620
			(0.154)	(0.389)	(1.223)
Colony of 3rd				-0.136	-1.049
				(0.381)	(1.187)
Colony of 4th					-0.365
					(0.773)
Same country as 1st		-0.224**	-0.296	-2.006**	-0.292
		(0.100)	(0.267)	(0.875)	(0.930)
Same country as 2nd			0.033	-0.009	-1.211
			(0.245)	(0.578)	(1.216)
Same country as 3rd				1.030**	-1.531
				(0.439)	(1.360)
Same country as 4th					1.958*
					(1.056)

continued

Table 6: Sequential Location Decision (Sensitivity III)
(concluded)

	Foreign Investment of the MNE:				
	1st	*2nd*	*3rd*	*4th*	*5th*
GTA with parent	0.073***	-0.220***	-0.278**	-0.157	-0.004
	(0.028)	(0.054)	(0.127)	(0.296)	(0.697)
GTA with 1st		0.534***	0.256**	-0.290	-0.821
		(0.051)	(0.125)	(0.278)	(0.551)
GTA with 2nd			0.305**	0.293	-0.055
			(0.128)	(0.308)	(0.600)
GTA with 3rd				0.367	-0.396
				(0.281)	(0.565)
GTA with 4th					0.440
					(0.606)
Pseudo R2	0.2258	0.2819	0.2845	0.2695	0.3331
Observations	1,164,529	402,256	79,677	15,999	4,688
Location decisions	15,165	4,693	885	190	60
Years between decisions		1.999	1.821	2.034	1.90

Notes: Conditional logit model. *Sensitivity III*: We only include sequential investments if the MNE has chosen only one location in phase $p-1$. In such cases, we have precise information on the reference investments in phase p. Robust standard errors reported in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively. *Location decisions* reports the actual number of location decisions made (Location decision = 1). *Years between decisions* are the average years between the respective (sequential) location decisions made by the multinationals in the sample. Control variables are taken from different sources. *Tax* is the statutory tax rate of a host country. The tax data is collected from databases provided by the International Bureau of Fiscal Documentation (IBFD) and tax surveys provided by Ernst&Young, PwC, and KPMG. *log GDP* measures the real GDP at constant U.S. dollars of the year 2000 and is taken from the World Bank's World Development Indicators 2009. The investment freedom index *InvestFree* is taken from the Heritage Indicators database. The index can take on values between 0 and 100; higher values are associated with more investment freedom. *InvestCost* is from World Bank's Doing Business Database and measures the cost of starting a business relative to income per capita. *CPI* (Corruption Perception Index) is published annually by Transparency International. It ranks countries in terms of perceived levels of corruption, as determined by expert assessments and opinion surveys. The scores range from 10 (country perceived as virtually corruption free) to 0 (country perceived as almost totally corrupt). *log Distance* is the log of the distance (in kilometer) between the most populated cities in the host country and the country of the previous investment. As to the bilateral variables for the first investment, we use Germany as the reference country. *Border* is a common border indicator, *Language* a common language indicator, *Colony* a former colony indicator, *Same country* a dummy indicating whether the host country and the country of the previous investment are the same. *GTA* is an indicator for the existence of a general trade agreement between the host country and the country of the previous investment. The bilateral variables are either taken from the Centre d'Études Prospectives et d'Informations Internationales (*log Distance, Border, Language, Colony, Same country*), or from the World Trade Organization (*GTA*).

Table 7: Simultaneous vs. Sequential Entry

Maximum # of investments:	**(2)**	**(3)**	**(4)**
	Dependent variable is mean investment size measured as fixed assets		
Sequential entry	2307.46***	1843.77***	948.83*
	(827.86)	(635.99)	(586.86)
R2	0.9279	0.9167	0.8444
Observations	*1,812*	*2,543*	*2,954*
	Dependent variable is mean investment size measured as total assets		
Sequential entry	23596.19***	18705.9**	14582.75*
	(6560.25)	(7739.50)	(8228.59)
R2	0.9409	0.8581	0.7853
Observations	*1,812*	*2,543*	*2,954*
	Dependent variable is mean sales-to-total-asset ratio		
Sequential entry	.253***	.216***	.208***
	(.069)	(.056)	(.052)
R2	0.0425	0.0386	0.0392
Observations	*1,812*	*2,543*	*2,954*
Share of sequential entries	.582	.629	.651

Notes: Maximum # of investments is the maximum number of foreign entities per firm which have been established (one-plant firms are not considered). For example, the column denoted by (3) indicates that firms have established 3 or 2 investments. The dummy variable *Sequential entry* indicates whether the investments have been established simultaneously (Sequential entry = 0) or sequentially (Sequential entry = 1). Robust standard errors reported in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively. For reasons of comparability of firms, the above coefficients are obtained from cross-section OLS regressions which also control for a firm's total sum of investments and averages of all country controls used in the conditional logit regressions above.

Table 8: Signals for one-plant & simultaneous-plant units

Dependent variable is fixed assets of an affiliate		
	one-plant units	*one-plant & all simultaneous-plant units*
Sales/Fixed Assets	-2.510*	-0.234*
	(1.417)	(0.127)
Observations	*6,130*	*7,477*
Sales/Total Assets	-957.711**	-686.600**
	(494.686)	(329.551)
Observations	*6,765*	*8,357*

Notes: OLS estimation. Robust standard errors reported in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively. For reasons of comparability, the above coefficients are obtained from cross-section OLS regressions which also control for averages of all country controls used in the conditional logit regressions above.

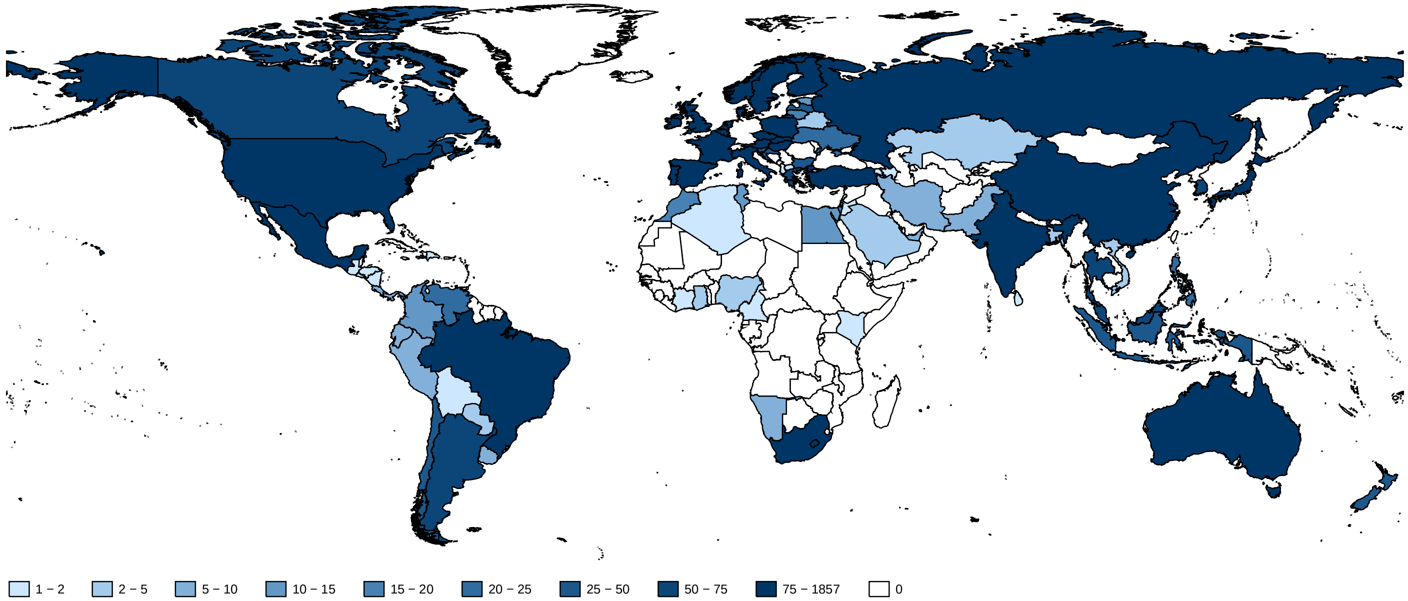

Figure 1: Location of First Investment

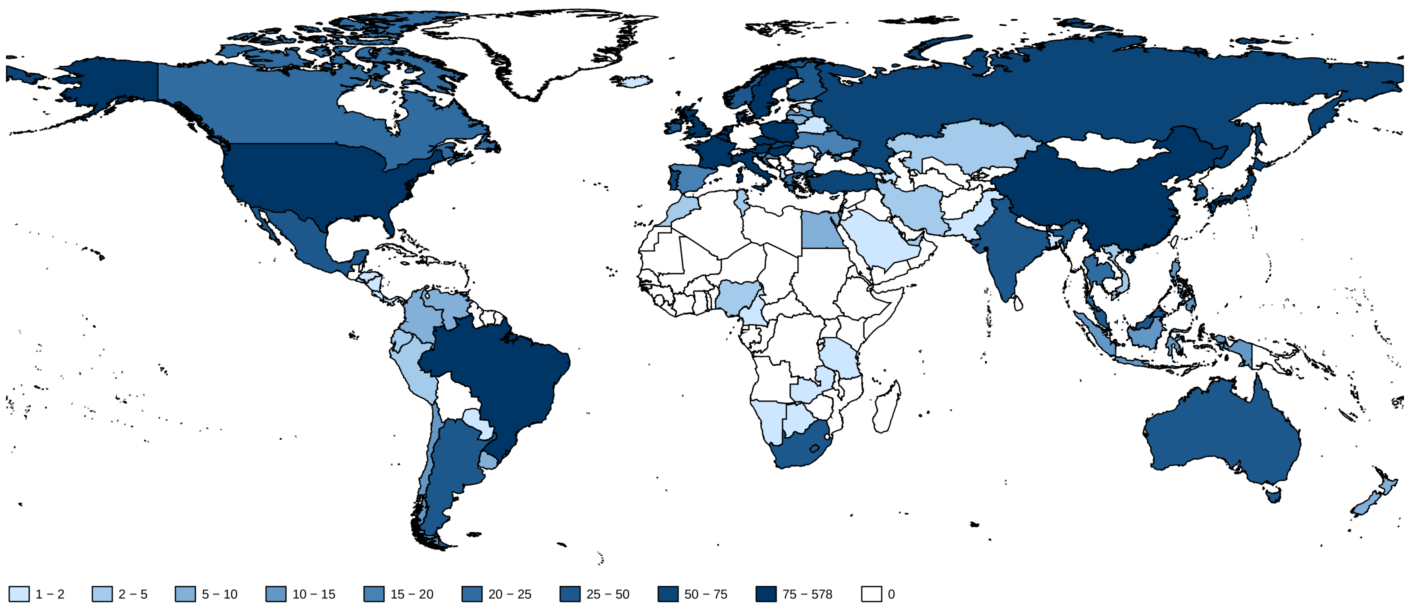

Figure 2: Location of Second Investment

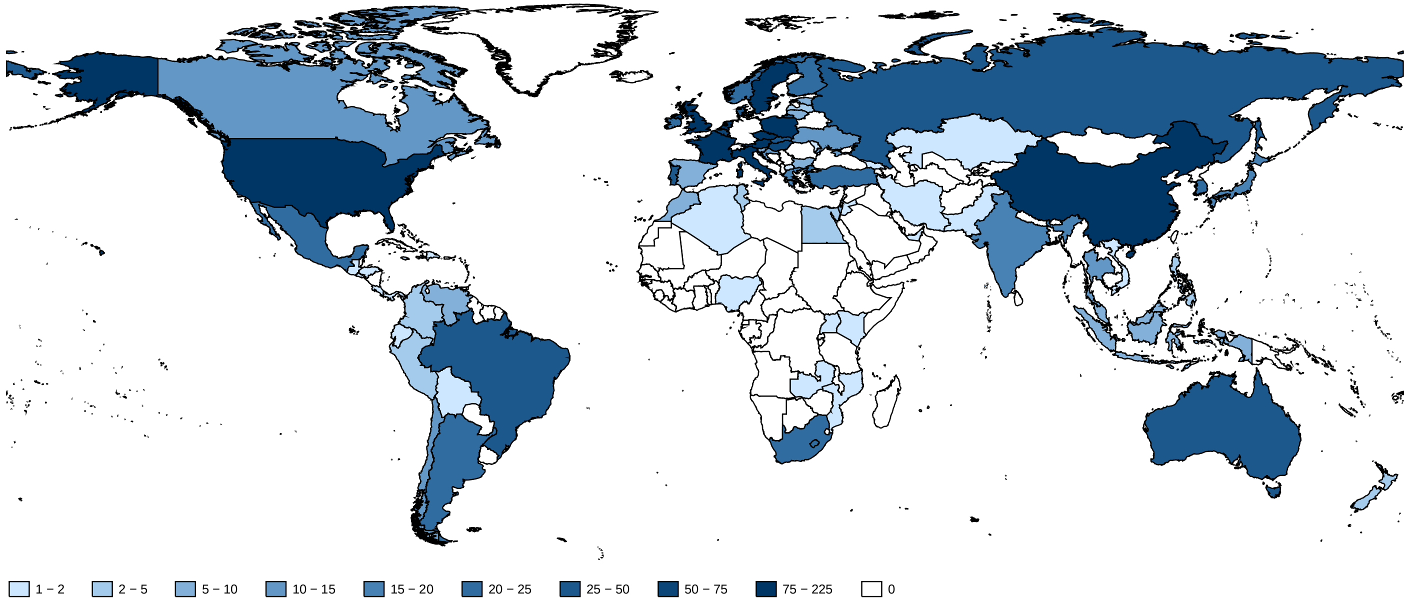

Figure 3: Location of Third Investment

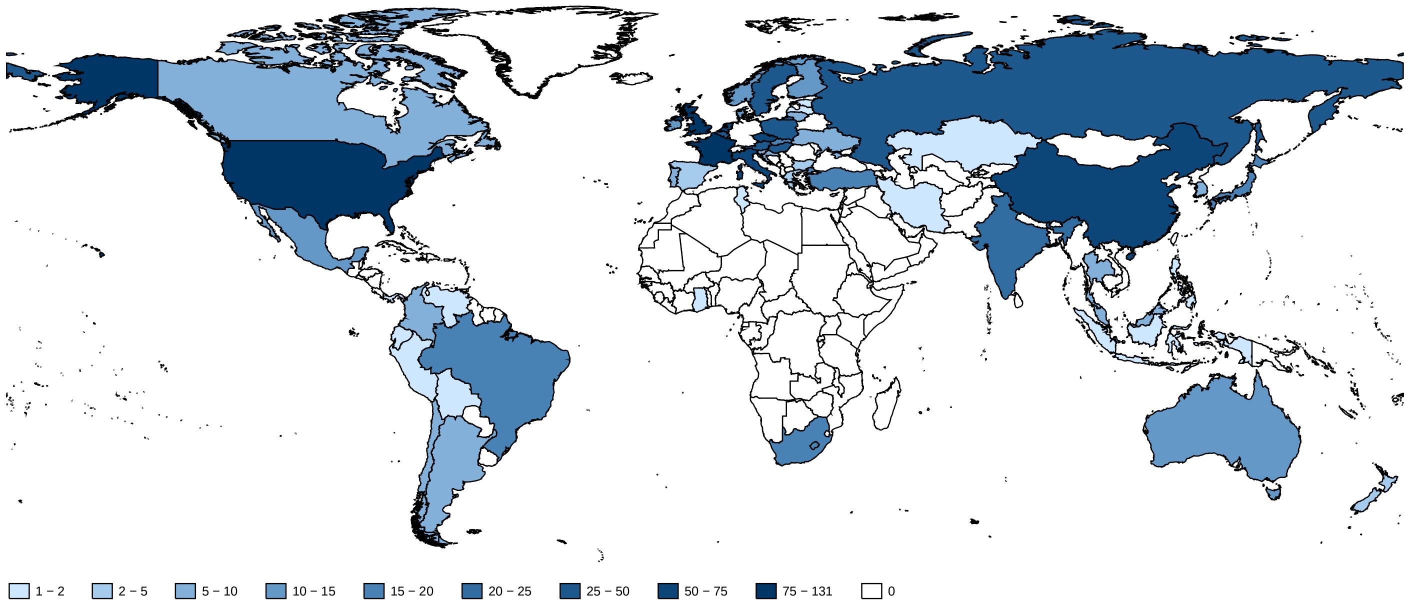

Figure 4: Location of Fourth Investment

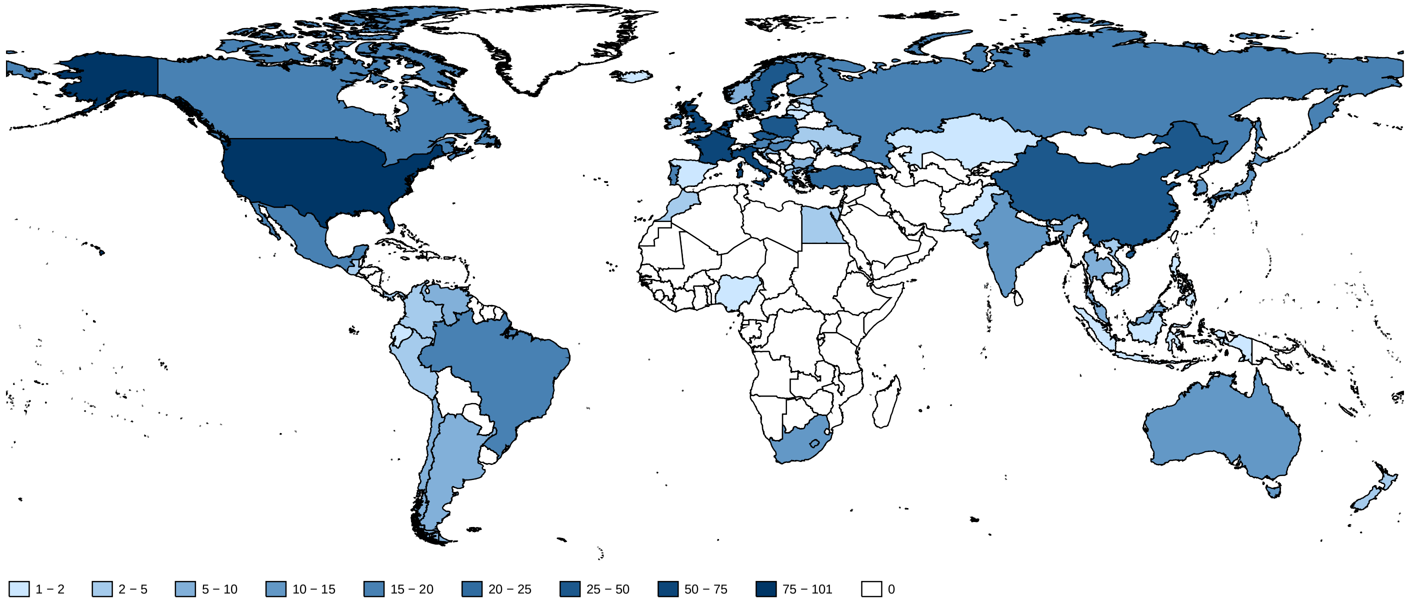

Figure 5: Location of Fifth Investment