

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Gumpert, Anna

Conference Paper The organization of knowledge in multinational firms

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Multinational Firms and International Trade, No. F06-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Gumpert, Anna (2014) : The organization of knowledge in multinational firms, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik -Session: Multinational Firms and International Trade, No. F06-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at: https://hdl.handle.net/10419/100332

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU

The organization of knowledge in multinational firms

Anna Gumpert*

March 1, 2014

Abstract

Employees must learn about firm technologies to use them in production. Within multinational firms, knowledge can be acquired centrally, by managers at headquarters, or locally, by production workers. Local knowledge acquisition increases with the bilateral communication costs with central management, and decreases with local knowledge acquisition costs. This mechanism explains why multinationals foreign sales and their probability of entry decrease in the distance of a country from the multinationals home country, and why multinationals pay higher wages than comparable domestic firms. The selection into foreign destinations and the foreign productivity distribution of German multinationals are consistent with the models predictions.

^{*}University of Munich, Seminar for Comparative Economics, Akademiestrasse 1, D-80799 Munich, anna.gumpert@econ.lmu.de. I am grateful to Pol Antràs and Esteban Rossi-Hansberg for helpful discussions and suggestions throughout the development of this project, and to Monika Schnitzer for her constant encouragement and support. I thank Arnaud Costinot, Carsten Eckel, Inga Heiland, Elhanan Helpman, Chen Li, James Markusen, Marc Melitz, Alexander Tarasov, and Martin Watzinger for their comments. This paper was partly written while I was a visitor at the Economics Department of Harvard University. I thank the Economics Department for its hospitality, and am particularly thankful to Pol Antràs for rendering my visit possible. The empirical analyses were conducted during visits to the research centre of the Deutsche Bundesbank. The hospitality of the Bundesbank, the constructive support of the staff as well as access to its Microdatabase Direct investment (MiDi) are gratefully acknowledged. The project has benefited from financial support through the Deutsche Forschungsgemeinschaft (German Science Foundation) under SFB-Transregio 15 and GRK 801.

1 Introduction

Knowledge is commonly considered as one of the main determinants of growth and productivity. The spatial diffusion of knowledge is limited, suggesting that a sizeable fraction of knowledge is tacit in nature (see Keller, 2004, for a survey). Multinational firms manage production processes that span different countries and involve employees from diverse economic and cultural contexts. Spatial and intercultural frictions hamper communication and information flows within multinational firms' boundaries (Ambos and Ambos, 2009; De Long and Fahey, 2000; UNCTAD, 2004, ch. IV A. 2).¹ Organizing knowledge is thus a particular challenge for multinational firms, and likely to be a driving force of their performance and investment patterns.

This paper analyzes the organization of knowledge within multinational firms and demonstrates that it helps understand major stylized facts of multinational firms' behavior. The organization of knowledge explains why the foreign sales of a multinational firm tend to decrease in the distance of a foreign country to the home country of the multinational firm (e.g. Keller and Yeaple, 2013), why the probability of multinational entry in a foreign country decrease in the distance of the country to the multinational firm's home country, and why multinational firms pay higher wages than comparable domestic firms both in the home and the host country (e.g. Aitken et al., 1996; Girma and Görg, 2007; Heyman et al., 2007; Hijzen et al., 2013).

As in much of the previous literature (e.g. Irarrazabal et al., 2013; Markusen and Maskus, 2001), knowledge is a non-rival factor within the firm. The paper additionally assumes that knowledge can only be fruitfully employed in production if it is learned by a worker. This assumption captures the fact that technologies are typically only useful if someone knows how to utilize them. Based on a stylized model of the organizational structure of a multinational firm, the paper determines the optimal allocation of knowledge within the firm given internationally heterogeneous market sizes, wages, labor productivities, knowledge acquisition costs and communication costs between the employees at the production locations and the headquarters.

The paper shows that the optimal amount of knowledge allocated by the firm to foreign employees in the host country increases with the communication costs between managers in the multinational's headquarters in the home country and the employees in the host country. Lower foreign labor productivities, higher foreign wages, higher knowledge acquisition costs and larger market size discourage the allotment of expertise. Correspondingly, the marginal costs of foreign production increase in the communication costs between the home country and the host country. At the same time, firms mitigate cost increases due to higher wages and knowledge acquisition costs or lower labor productivities by reallocating knowledge.

Due to the impact of the communication costs between the home and the host country on a firm's marginal costs, a multinational firm's sales and its probability of entry are predicted to be geographically concentrated and decreasing in distance, i.e., the investment patterns of MNEs are predicted to exhibit gravity. The concentration is higher in sectors with more complex production processes, as found in Bahar (2013). As a multinational firm re-allocates knowledge compared to a domestic firm, the model also explains why the remuneration of workers at multinational firms is higher than the remuneration of workers at domestic firms both in the home and the host country (e.g. Heyman et al., 2007). These multinational wage premiums

¹The business economics and management literature discusses and evaluates various strategies to address the frictions in within-firm communication faced by internationally active firms (e.g. Foss and Pedersen, 2002; Ghemawat, 2007; Hansen and Løvås, 2004; Lagerström and Andersson, 2003).

vary with home and host country characteristics, as in Girma and Görg (2007), because the organization of knowledge in multinational firms is sensitive to these characteristics. The paper thus contributes to the literature by providing a single mechanism that is apt to explain different empirical phenomena that have been separately analyzed in the literature so far (Irarrazabal et al., 2013; Keller and Yeaple, 2013, for analyses of gravity; e.g. Egger and Kreickemeier, 2013; Malchow-Møller et al., 2013, for analyses of MNE wage premiums).

To evaluate the model systematically, the empirical analysis derives and tests predictions on within-firm differences in performance across different foreign countries and the selection of firms into foreign investment. Consistent with the theoretical hypotheses, German multinational firms tend to be relatively more productive in countries that are larger, that are characterized by lower bilateral communication costs with Germany, higher labor productivities, lower knowledge acquisition costs and lower wages. The findings are robust to the inclusion of additional determinants of foreign performance as well as alternative ways of measuring the model parameters. The cut-off productivity levels for entry are found to decrease in country size and increase in bilateral communication costs.

The results of the paper have important implications for investment promotion purposes. The paper recommends that investment promotion policies should not only focus on improving the investment climate inside a country, but also on ameliorating bilateral communication facilities with targeted source countries of FDI, for example by improving language training and communication infrastructures.

The paper builds on two distinct strands of the literature. In developing the model of the optimal organizational structure, the paper uses ingredients from the literature of firms as communication networks and knowledge hierarchies (e.g. Bolton and Dewatripont, 1994; Garicano, 2000; Radner, 1993). This literature studies how firms organize in response to differences in knowledge acquisition costs and in the cost of communication between agents, how hierarchical structures emerge, and how knowledge is efficiently allocated within a firm. Garicano (2000)'s knowledge hierarchies framework has been applied to various settings, and its main predictions have been confirmed empirically (e.g. Bloom et al., 2009; Garicano and Hubbard, 2009). Caliendo and Rossi-Hansberg (2012) and Caliendo et al. (2012) demonstrate theoretically and empirically that the model framework is useful to understand the labor market consequences of international trade. To the best of our knowledge, this paper is the first to apply the knowledge hierarchy framework to a context with several establishments.

Concerning the international investment aspects, the paper draws on the large literature on multinational firms. Much of this literature has associated multinational activity predominantly with its advantages for a firm, such as savings of trade costs with the host market (e.g. Markusen, 1984; Helpman et al., 2004), savings of trade costs with third markets (e.g. Ekholm et al., 2007; Tintelnot, 2012), savings of factor input costs (e.g. Helpman, 1984; Antras, 2003; Antras and Helpman, 2004), or a combination of these (e.g. Carr et al., 2001; Grossman et al., 2006; Yeaple, 2003). In line with recent evidence on U.S. multinational firms indicating that most foreign affiliates are market-seeking, the theory section of this paper assumes a horizontal motive of foreign activities (Ramondo et al., 2013).

Within the literature on multinational firms, the strand most relevant to this paper is concerned with the role of headquarter inputs for local affiliate production. Keller and Yeaple (2013) estimate the cost of transferring knowledge within multinational firms using data on U.S. manufacturing FDI. In their model, firms face a trade-off between the costs associated with disembodied and embodied knowledge transfer from their headquarters. Firms can either produce intermediaries locally but subject to efficiency losses due to communication frictions with the headquarters, or import intermediaries from home subject to transport costs. The communication friction varies across sectors, but does not vary across space. This paper analyzes a complementary margin of multinational firms' behavior: it focuses on the choice between employing domestic expertise from the headquarters or hiring foreign expertise in the host market. The mechanism explains the gravity of foreign direct investment in sales, as does the model in Keller and Yeaple (2013), but is applicable across sectors and also generates multinational wage premiums consistent with empirical evidence in the literature. Yeaple (2013) develops a model featuring managerial expertise as a scarce input within a multinational firm to explain empirical regularities in the expansion patterns of U.S. multinational multi-product firms. Similarly, Irarrazabal et al. (2013) assume that headquarter inputs, such as managerial oversight or marketing services, are required in the production process of affiliates and demonstrate that this assumption helps to explain the gravity of FDI in sales and the probability of entry. Defever (2012) assumes that smooth communication between foreign affiliates and the headquarters acts as a productivity shifter in a Cobb–Douglas production function and finds that geographical proximity to existing affiliates is a relevant consideration for the location decision of a multinational firm, above and beyond other locational characteristics. Bahar (2013) demonstrates that the probability of firms' conducting horizontal FDI decreases in their knowledge intensity, and that firms tend to prefer proximate destinations for the location of knowledge intensive FDI. He argues that iceberg type knowledge transfer costs that increase in the knowledge intensity of activities are appropriate for explaining this finding. All papers share their focus on the gravity of FDI and their theoretical approach, extending the basic framework in Helpman et al. (2004) with productivity-shifting mechanisms. This implies that the marginal costs of production are constant, country-specific and independent across countries. This paper is distinct in modeling the organizational structure of multinational firms. It contributes to the literature by providing a coherent rationale for both the gravity and multinational wage premiums, and by demonstrating how the performance of multinational firms may be interdependent across markets.

The following section develops the theoretical model. It is divided into three subsections. The first subsection analyzes the cost-minimization problem at the firm level. The second subsection investigates the internationalization decision of the firm in a partial equilibrium framework. The last subsection outlines the general equilibrium implications. Section 3 discusses the empirical strategy and Section 4 describes the data used in the empirical analysis. Section 5 presents the empirical results. Section 6 examines how far the empirical results discriminate between the mechanism proposed in this paper and alternative models that may generate the observed empirical patterns. The last section concludes.

2 Theory

The model economy consists of two countries, the home country j = 0 and the foreign country j = 1. The countries are populated by N_j agents that are each endowed with one unit of time.

There is an unbounded mass of potential firms in each country. Each firm can produce a differentiated product. Production requires labor and knowledge. To enter, firms have to pay a

fixed entry cost of f units of domestic labor which is thereafter sunk. Upon entry, each entrant i draws a firm-specific knowledge level \bar{z}_i from a known distribution $G(\bar{z})$ which is symmetric in the two countries. Mathematically, knowledge is an interval that ranges from zero to a firm-specific upper bound \bar{Z}_i . \bar{z}_i denotes the length of a firm's knowledge interval $[0, \bar{Z}_i]$ (i.e. its Lebesgue measure). Empirically, the knowledge level \bar{z}_i corresponds to the state of a firm's technology.

The problem of each firm is to determine the set of countries it would like to serve and the set of production locations given its knowledge level \bar{z}_i . In addition, firms choose the cost-minimizing allocation of knowledge within the firm. The following section first analyzes the cost-minimization problem for different sets of production locations, then determines the optimal set of countries and production locations, and finally investigates the entry decision.

2.1 Cost minimization

2.1.1 Assumptions

To simplify the exposition, subsection 2.1 focuses on the cost-minimization problem of a single firm located in the home country and characterized by the knowledge level \bar{z} . The firm consists of a headquarters in country j = 0 and production plants in country j = 0 and, potentially, in j = 1 as well. The headquarters are composed of a number n_h of managers, and the production plants are made up of a number n_j of production workers.

Production is assumed to be a problem solving process that is based on labor and knowledge (as in Garicano, 2000). The firm can employ labor for production in both countries j = 0 and j = 1. For each unit of labor employed in production, a mass 1 of problems is realized. Transforming labor input into output requires that the problems be solved. A firm is able to solve a problem if it is realized within the firm's knowledge interval. The problems are distributed according to a problem probability distribution function from the exponential family:

$$f(z) = \lambda e^{-\lambda z}$$

where $z \in [0, \infty)$ refers to the domain of possible problems and $\lambda > 0$ denotes the problem arrival rate. A higher value of λ implies that the mass of the probability distribution is concentrated close to zero. Intuitively, this means that the production process is more predictable as problems in the tail of the probability distribution occur with lower probability.

Consistent with the previous literature (e.g. Irarrazabal et al., 2013), knowledge is assumed to be non-rival within the firm and can thus be used at all locations. Unlike previous papers, knowledge is only useful if it has been learnt by an employee. The underlying idea is that firm knowledge by itself is insufficient for production. Knowledge is only useful if there is someone who knows how to employ it.

Knowledge can either be acquired locally by production workers, or centrally by managers. Employees can communicate with each other and leverage the potentially different knowledge sets. Communication is costly. As is standard in the literature (e.g. Bolton and Dewatripont, 1994; Garicano, 2000; Bloom et al., 2009), communication costs are borne by the employee who is asked a question: the employee has to spend time listening to the problems. Garicano (2000) demonstrates that at optimum, only production workers spend their time in production and managers use their time solely for communication. Both production workers and managers are optimally characterized by knowledge levels that are uniform within each group and different between the two groups. Uniform knowledge levels reduce communication time by diminishing the time spent searching for a competent contact. Workers know that only managers may know solutions to problems that they themselves cannot solve, and that it does not matter which manager they approach. To minimize the probability that costly communication is necessary, the knowledge level of production workers covers the solution to the more frequent problems, whereas managers know the solution to problems that occur more rarely.

The production process works as follows. During each unit of time that they spend in production, the production workers scrutinize which of the problems are covered by their knowledge interval. Problems covered by the knowledge interval are solved immediately and output is produced. The production workers communicate all problems that are not covered by their knowledge interval to the managers. The amount of time that managers spend listening to the problems depends on whether the production workers are located in the same or another country. The managers in country j spend $\theta_{kj} \in (0, 1)$ units of time listening to problems delegated by production workers in country k. It is generally assumed that $\theta_{10} \ge \theta_{00}$ and $\theta_{11} = \theta_{00}, \theta_{01} = \theta_{10}$ capture the fact that frictions in international communication may be more severe than frictions in communication within a country. The managers solve all problems covered by their knowledge interval. Any problems that are not covered by the knowledge intervals of either the workers or the managers remain unsolved.

The output q_j of one unit of labor input can be calculated as the value of the cumulative distribution function times labor productivity. Labor productivity A_j varies across countries and measures the maximum amount of output that can be produced with one unit of labor input and infinite knowledge.

$$q_j = A_j (1 - e^{-\lambda \bar{z}})$$

 n_j workers are correspondingly able to produce $q_j = n_j A_j (1 - e^{-\lambda \bar{z}})$ units of output. A higher value of \bar{z} implies that more infrequent problems can be solved, so the resulting product is more sophisticated.

Knowledge acquisition is costly: employees have to hire teachers to teach them (as in Caliendo and Rossi-Hansberg, 2012). Similar to the modelling of communication costs, teaching costs are fully borne by the teachers. Teachers have to spend $c_j x$ units of time to teach a knowledge interval of length x. All agents receive wage w_j per unit of time they spend working. Correspondingly, teachers receive remuneration $w_j c_j x$. The firm in turn has to remunerate its employees for the time they spend in production and for their knowledge acquisition expenses.

The problem of the firm is to design the optimal knowledge acquisition process. As outlined above, Garicano (2000) shows that only production workers spend their time in production and that they optimally learn the solutions to problems that occur more frequently in the production process. The knowledge interval of workers correspondingly starts at 0, where the mass of the problem density is highest, and ranges to a country specific upper bound Z_j , j = 0, 1. z_j denotes the length of the knowledge interval of production workers $[0, Z_j]$. The managers learn to solve infrequent problems. It is never optimal for the firm not to learn part of its knowledge interval $[0, \overline{Z}]$.² More knowledge enables the firm to produce more output with a given amount

²All managers are assumed to have the same knowledge level. Large multinational corporations may afford specialized divisions at their headquarters that are responsible for certain countries only and that possess specialized knowledge. It is possible to accomodate this consideration in the model by allowing the firm to choose

of labor input and, as will be shown below, knowledge decreases marginal costs. The upper bound of managerial knowledge and the upper bound of the knowledge interval of the firm thus coincide. The knowledge interval of managers ranges from a lower bound Z_h to \overline{Z} . z_h denotes the length of this interval $[Z_h, \overline{Z}]$. The firm chooses the knowledge levels z_j and z_h as well as the number of production workers n_j and the number of managers n_h . By choosing z_j and z_h , the firm determines the upper bound of the workers' knowledge interval(s) Z_j and the lower bound of the managerial knowledge in the headquarters Z_h .

2.1.2 The optimization problem

The firm strives to minimize the costs of producing a given quantity q_j in country j. The costs are composed of the cost for personnel at the headquarters and at the different production locations. Each employee is remunerated with the country specific wage w_j per unit of time spent working for the firm and for knowledge acquisition costs $w_j c_j z_k$, k = h, j.

The firm optimally chooses the number of production workers $\{n_j\}_{j=0}^1$ employed at the different production locations, the country specific knowledge level of the production knowledge $\{z_j\}_{j=0}^1$, the number of managers employed in the headquarters n_h , and the knowledge level of those managers z_h .

$$C(q_0, w_0, q_1, w_1) = \min_{\{n_j, z_j\}_{j=0}^1, n_h, z_h} \sum_{j=0}^1 n_j w_j (1 + c_j z_j) + n_h w_0 (1 + c_0 z_h)$$
(1)

s.t.
$$n_j A_j (1 - e^{-\lambda \bar{z}}) \ge q_j \qquad \forall j$$
 (2)

$$z_j \ge \bar{z} - z_h \quad \forall j \tag{3}$$

$$n_h \ge \sum_{j=0}^{1} n_j \theta_{j0} e^{-\lambda z_j} \tag{4}$$

$$n_h \ge 0, \, z_h \ge 0, z_h \le \bar{z} \tag{5}$$

$$n_j \ge 0, \, z_j \ge 0, z_j \le \bar{z} \qquad \forall j$$

$$\tag{6}$$

In the cost minimization problem, the production quantities $\{q_j\}_{j=0}^1$ are taken as given, but they will be endogenized in subsection 2.2, as will be the set of production locations of the firm. Wages $\{w_j\}_{j=0}^1$ will be endogenized in subsection 2.3. The problem arrival rate λ , communication costs $\{\theta_{j0}\}_{j=0}^1$, labor productivities $\{A_j\}_{j=0}^1 > 0$ and knowledge acquisition costs $\{c_j\}_{j=0}^1 > 0$ are exogenous parameters determined by the state of a country's economic development and institutions.

In choosing the optimal values of $\{n_j\}_{j=0}^1$, n_h , $\{z_j\}_{j=0}^1$ and z_h the firm has to take four types of constraints into account:

- Eq. (2): The firm has to produce a total output $n_j A_j (1 e^{-\lambda \bar{z}})$ of at least q_j units.
- Eq. (3): The knowledge intervals of the production workers $[0, Z_j]$ and the managers $[Z_h, \overline{Z}]$

eight endogenous variables: $\{z_j, n_j, z_h^j, n_h^j\}_{l=0}^1$, where n_h^j is the number of managers responsible for production in country j and z_h^j is their knowledge level. Such a modified model does not generate all of the predictions of the main model in the paper. In particular, residual multinational wage premiums in the home country of the multinational firm do not arise. Multinational firms probably have specialized departments in their headquarters. As upper management is responsible for worldwide operations, managers have to be able to address issues brought up from anywhere in the corporation at least at some level of seniority. This is captured by the assumption of non-tailored managerial knowledge. A formal analysis of the modified model is sketched in Appendix A.5.

have to jointly cover the full knowledge interval $[0, \overline{Z}]$ of the firm. This is ensured if the lengths of the production workers' knowledge interval z_j and the managers' knowledge interval z_h add up to at least the knowledge level of the firm \overline{z} .

- Eq. (4): The firm has to hire a sufficient number of managers such that the managers are able to listen to all problems brought up to them. The amount of problems is calculated as the sum of the country specific masses of problems n_j times the probability that the solution is *not* found by the production workers in j, $e^{-\lambda z_j}$. This term is multiplied by the communication costs θ_{j0} to obtain the required number of managers.
- Eq. (5, 6): All choice variables are restricted to be positive. Employees' knowledge cannot exceed the total knowledge of the firm.

Assumption 1. The exogenous parameters \bar{z}, λ and $\{c_j, A_j, \theta_{j0}\}_{j=0}^1$, as well as those exogenous parameters that are contained in $\{q_j\}_{j=0}^1$ and $\{w_j\}_{j=0}^1$, fulfil the following parameter restriction:

$$\frac{\frac{q_{\bar{j}}}{A_{\bar{j}}}\theta_{\bar{j}0}w_0(c_0+\lambda)}{\frac{q_{\bar{j}}}{A_{\bar{j}}}w_{\bar{j}}c_{\bar{j}}-\frac{c_0}{\lambda}\frac{q_{\hat{j}}}{A_{\hat{j}}}w_{\hat{j}}c_{\hat{j}}} \le e^{\lambda\bar{z}} \le \frac{\frac{q_{\bar{j}}}{A_{\bar{j}}}w_{\bar{j}}(\lambda(1+c_{\bar{j}}\bar{z})+c_{\bar{j}})+\frac{q_{\bar{j}}}{A_{\bar{j}}}\lambda\theta_{\bar{j}0}w_0+\frac{q_{\hat{j}}}{A_{\hat{j}}}w_{\hat{j}}(\lambda(1+c_{\hat{j}}\bar{z})+c_{\hat{j}})}{\frac{q_{\bar{j}}}{A_{\bar{j}}}w_{\bar{j}}c_{\bar{j}}}, \quad (7)$$

where \bar{j} denotes the country with the binding knowledge constraint $z_{\bar{j}} = \bar{z} - z_h$ and \hat{j} denotes the country where the knowledge constraint is not binding.

The parameter restriction ensures that constraints (3) and (4) are innocuous in the sense that the firm is never forced to adopt a knowledge level \bar{z} that exceeds the level it would adopt if it were free to choose its total knowledge level, and that the firm is never forced to hire managers even though it would prefer to hire only production workers. Given Assumption 1, the marginal costs of production are strictly decreasing in the overall knowledge level, so covering the full knowledge interval is always optimal for the firm. The parameter restriction is formally derived in Appendix A.1.1.

The Lagrangian equation is given by

$$\mathcal{L} = \sum_{j=0}^{1} n_j w_j (1 + c_j z_j) + n_h w_0 (1 + c_0 z_h) + \sum_{j=0}^{1} \xi_j \left[q_j - n_j A_j (1 - e^{-\lambda \bar{z}}) \right] + \sum_{j=0}^{1} \phi_j \left[\bar{z} - z_h - z_j \right] + \kappa \left[\sum_{j=0}^{1} n_j \theta_{j0} e^{-\lambda z_j} - n_h \right].$$

The Lagrangian multiplier ξ_j denotes the marginal costs of production. The other Lagrangian multipliers do not have intuitive interpretations.

The first order conditions are detailed in Appendix A.1.2. The optimal number of production workers is determined by the quantity constraint:

$$n_j = \frac{q_j}{A_j(1 - e^{-\lambda \bar{z}})}.$$
(8)

The optimal number of managers results from the constraint on the number of managers:

$$n_h = \sum_{j=0}^{1} n_j \theta_{j0} e^{-\lambda z_j} = \sum_{j=0}^{1} \frac{q_j \theta_{j0} e^{-\lambda z_j}}{A_j (1 - e^{-\lambda \bar{z}})}.$$
(9)

The optimal knowledge levels of the production workers $\{z_j\}_{j=0}^1$ may be different due to asymmetries in the country characteristics. The knowledge constraint is binding for at least one country:

$$z_j = \bar{z} - z_h. \tag{10}$$

Intuitively, if the knowledge constraint were non-binding for both countries, the overlap of knowledge at the headquarters and all production locations would remain unused. This cannot be optimal.

If the knowledge constraint is non-binding, the knowledge level of the production workers is determined by

$$e^{-\lambda z_j} = \frac{w_j c_j}{\lambda \theta_{j0} w_0 (1 + c_0 z_h)}.$$
 (11)

The knowledge constraint is binding whenever the marginal costs of additional local knowledge exceed the marginal benefit of using fewer services from the headquarters, i.e. whenever the firm would actually like to choose knowledge level $\frac{1}{\lambda} \ln \left[\frac{\lambda \theta_{j0} w_0(1+c_0 z_h)}{w_j c_j} \right] < \bar{z} - z_h$. From $\frac{1}{\lambda} \ln \left[\frac{\lambda \theta_{\bar{j}0} w_0(1+c_0 z_h)}{w_{\bar{j}} c_{\bar{j}}} \right] < \bar{z} - z_h < \frac{1}{\lambda} \ln \left[\frac{\lambda \theta_{j0} w_0(1+c_0 z_h)}{w_j c_{\bar{j}}} \right]$, it follows that the country characteristics of the country with binding constraint $z_{\bar{j}} = \bar{z} - z_h$ and non-binding constraint $z_{\hat{j}} = \bar{z} - z_h$ are related by

$$\theta_{\overline{j}0}w_{\hat{j}}c_{\hat{j}} < \theta_{\hat{j}0}w_{\overline{j}}c_{\overline{j}}.$$

The knowledge constraint is ceteris paribus more likely to be binding in the home country due to the lower communication costs, and in the country with higher wages and higher knowledge acquisition costs.

Managerial knowledge is implicitly determined by

$$\sum_{j=0}^{1} \left[\mathbf{1}(z_j > \bar{z} - z_h) c_0 \frac{q_j w_j c_j}{A_j} + \mathbf{1}(z_j = \bar{z} - z_h) \lambda (1 + c_0 z_h) \frac{q_j}{A_j} \left(\theta_{j0} e^{-\lambda(\bar{z} - z_h)} w_0 (c_0 + \lambda(1 + c_0 z_h)) - w_j c_j \right) \right] = 0.$$
(12)

The indicator function $\mathbf{1}(\cdot)$ determines whether the constraint $z_j = \bar{z} - z_h$ is binding or not.

If the firm only produces in the domestic country, z_0 , n_0 and n_h are determined by the constraints (2)-(4). Managerial knowledge is implicitly defined by the simpler condition

$$\theta_{00}e^{-\lambda(\bar{z}-z_h)}(c_0+\lambda(1+c_0z_h))-c_0=0.$$
(13)

The first order condition (13) equates the marginal benefit of increasing z_h , i.e. the savings in knowledge acquisition costs of production workers, n_jc_0 , and the marginal costs of increasing z_h . These are composed of the costs of increasing managerial knowledge, $n_j\theta_{00}e^{-\lambda(\bar{z}-z_h)}c_0$, and the increase in the demand for management services, $n_j\theta_{00}e^{-\lambda(\bar{z}-z_h)}\lambda(1+c_0z_h)$. The number of production workers n_j cancels from the equation. The optimal knowledge allocation in a domestic firm is only determined by the within-country communication costs θ_{00} , the knowledge acquisition costs c_0 , the problem arrival rate λ and the firm's total knowledge \bar{z} . The first order condition (12) analogously equates the marginal benefit and the marginal costs of increasing z_h . The marginal benefit consists of lower knowledge acquisition costs of the production workers in the country with the binding knowledge constraint, $\mathbf{1}(z_j = \bar{z} - z_h)\frac{q_j}{A_j}w_jc_j$. The marginal costs consist of the costs of increasing managerial knowledge $\mathbf{1}(z_j > \bar{z} - z_h)c_0 \frac{q_j w_j c_j}{A_j \lambda(1+c_0 z_h)} + \mathbf{1}(z_j = \bar{z} - z_h)\frac{q_j}{A_j}\theta_{j0}e^{-\lambda(\bar{z}-z_h)}w_0c_0$ and the increase in the demand for managers $\mathbf{1}(z_j = \bar{z} - z_h)\frac{q_j}{A_j}\theta_{j0}e^{-\lambda(\bar{z}-z_h)}w_0\lambda(1+c_0 z_h)$. The optimal knowledge allocation in a multinational firm is more complex, because the firm cannot simultaneously choose the knowledge level that would be optimal if it produced in only one of the two countries (unless all parameters in the foreign country coincide with the parameters in the domestic country). As the marginal costs and the marginal benefit of varying z_h depend on the number of production workers at the two locations, which in turn depends on the production quantities, the firm takes the amount of output at the two locations into account in choosing the optimal level of managerial knowledge.

The marginal costs of production are given by

$$\xi_{j} = \frac{1}{A_{j}(1 - e^{-\lambda\bar{z}})} \left[w_{j}(1 + c_{j}z_{j}) + \frac{1}{\lambda}w_{j}c_{j} \right] \qquad \text{for } z_{j} > \bar{z} - z_{h}; \quad (14)$$
$$= \frac{1}{A_{j}(1 - e^{-\lambda\bar{z}})} \left[w_{j}(1 + c_{j}(\bar{z} - z_{j})) + w_{0}(1 + c_{0}z_{j})\theta_{j}e^{-\lambda(\bar{z} - z_{h})} \right] \qquad \text{for } z_{j} = \bar{z} - z_{h}. \quad (15)$$

$$= \frac{1}{A_j(1 - e^{-\lambda \bar{z}})} \left[w_j(1 + c_j(\bar{z} - z_h)) + w_0(1 + c_0 z_h)\theta_{j0}e^{-\lambda(\bar{z} - z_h)} \right] \quad \text{for } z_j = \bar{z} - z_h. \quad (15)$$

As $\xi_j > 0 \quad \forall q_j$, the cost function is strictly increasing in q_j .

2.1.3 Comparative statics

Proposition 1. The optimal choices of the firm vary with total firm knowledge, the problem arrival rate and country characteristics.

Choices/ Parameters	\overline{z}	λ	$ heta_{j0}$	c_j	w_j	A_j	q_j
# production workers n_j	-	-	0	0	0	-	+
Production knowledge z_j , MNE, $z_j = \bar{z} - z_h$	+	+/-	+	-	-	+	-
Production knowledge z_j , MNE, $z_j > \bar{z} - z_h$	+	-	+	-	-	+	-
Production knowledge z_j , domestic firm	+	+/-	+	-	0	0	0
# managers n_h , multinational firm (MNE)	-	-	+/-	+	+	-	+
# managers n_h , domestic firm	-	-	+/-	+	0	-	+
Managerial knowledge z_h , MNE, $z_j = \bar{z} - z_h$	+	+/-	-	+	+	-	+
Managerial knowledge z_h , MNE, $z_j = \bar{z} - z_h$	+	+/-	0	-	-	+	-
Managerial knowledge z_h , domestic firm	+	+/-	-	+	0	0	0

Table 1: Comparative statics

The table presents the effect of the parameters on the endogenous choices of the firm. + denotes positive effects, - denotes negative effects, +/- refers to ambiguous effects and 0 refers to independence. "Domestic firm" refers to firms that *produce* only domestically; the category includes exporters.

Proof. See Appendix A.1.3.

The effects of \bar{z} , A_j and q_j on the number of production workers are intuitive. A higher value of λ implies that the same output can be produced with less labor input as a higher fraction of problems occurs close to zero, so the number of workers decreases.

The knowledge level of the production workers increases in the total knowledge of the firm \bar{z} . The problem arrival rate λ has an indirect ambiguous effect via managerial knowledge on the knowledge level of production workers. In addition, it directly and negatively affects the knowledge level of production workers in a multinational firm if the knowledge constraint is

not binding: less knowledge is required to produce a given level of output. This direct effect dominates the indirect effect, which is the only force that is present if the knowledge constraint is binding, be it in a multinational or domestic firm. Higher communication costs entail an incentive to assign workers more knowledge to increase the amount of problems that do not need to be communicated to the headquarters. Higher learning costs c_j and, in the case of a multinational firm, higher wages w_j increase the remuneration for every single worker, so it is optimal to decrease their knowledge to mitigate cost increases. If the labor productivity A_j in country j decreases, or a larger quantity q_j is to be produced, more workers need to be hired, each of whom receives $w_j(1 + c_j z_j)$. A multinational firm can mitigate this cost increase by decreasing the knowledge level of the production workers in country j and adjusting the optimal knowledge allocation within its organization.

Wages, labor productivity and the production quantity do not affect the production workers' optimal knowledge level for domestic firms. Labor productivity and the production quantity scale the "breadth" of the domestic organization. An increase in the production quantity (or a decrease in labor productivity) leads to a proportional increase in the number of workers, which in turn causes a proportional increase in the number of managers. Changing the allocation of knowledge within a domestic firm does not affect this requirement. Similarly, wages scale the total costs of production. The effect of knowledge acquisition costs and communication costs is different. The firm faces a trade-off: allocating more (less) knowledge to the production workers increases (decreases) the total costs at the production level, but decreases (increases) the costs that accrue due to communication between production workers and managers. The optimal amount of knowledge allocated to the production workers therefore decreases in the knowledge acquisition costs with the headquarters θ_{00} also in a firm that produces only domestically. All the results for domestic firms are in line with the findings for single-establishment firms in Bloom et al. (2009).

The effect of total firm knowledge, the production quantity and labor productivity on the number of managers is straightforward. The problem arrival rate affects the number of managers in two distinct ways. It has a direct, negative effect on the number of managers because it decreases the number of workers. It additionally has an indirect, ambiguous effect on the number of managers via knowledge. The direct effect always dominates the indirect effect. Communication costs have an ambiguous effect on the optimal number of managers. On the one hand, an increase in communication costs implies that managers have to spend more time to accomodate a given number of problems, so the number of managers increases. On the other hand, higher communication costs lead to higher knowledge levels at the production locations. The number of problems referred to the headquarters decreases, as does the number of managers. The number of managers increases in the knowledge acquisition costs and, in case of multinational firms, in wages because these lead to a decrease in production workers' knowledge. This implies that the number of problems sent to headquarters increases, so more managers have to be hired to deal with them.

Managerial knowledge increases in the total knowledge of the firm. The effect of the problem arrival rate λ is ambiguous. A higher value of λ decreases the probability that the solution is not found by the production workers for a given value of z_h . This sets an incentive to increase managerial knowledge z_h to save production costs at the level of the production workers. At the same time, a higher value of λ implies that the number of managers responds more strongly to changes in z_j . More managers need to be hired if $z_j = \bar{z} - z_h$ is decreased, which dampens the positive effect of λ on z_h . If the knowledge constraint is binding, country characteristics have exactly the opposite effect on managerial knowledge as they have on production knowledge, because the sum of the two knowledge levels has to cover the full knowledge interval. Country characteristics (except communication costs) have an identical effect on managerial knowledge as on production knowledge if the knowledge constraint is not binding. This result may seem counterintuitive at first. Take the effect of wages as an example. Higher wages imply that the firm decreases the level of knowledge of production workers. Consequently, the number of problems sent to headquarters increases. This entails an incentive to decrease the marginal costs of using the headquarters, which is achieved by decreasing heaquarter knowledge. This is possible, as the knowledge constraint is not binding. The effect of the other characteristics can be derived analogously.

2.2 Profit maximization

There are many firms in each country j = 0 and j = 1. Each firm *i* produces a distinct variety and is characterized by a firm-specific knowledge level \bar{z}_i . It solves the cost minimization problem analyzed in the previous subsection. The analysis is conducted for a firm in the domestic country j = 0, and the analogous results apply to firms in the foreign country j = 1.

Consumers in each country have CES preferences:

$$U(x_j(\bar{z})) = \left(\int_{\Omega_j} x_j(\bar{z}_i)^{\frac{\sigma-1}{\sigma}} M_j \mu(\bar{z}) d\bar{z}\right)^{\frac{\sigma}{\sigma-1}}$$
(16)

where Ω_j is the set of varieties available in country j, M_j is the mass of firms, $\mu(\bar{z})$ denotes the density of knowledge levels of the firms in country j, $\sigma > 1$ is the elasticity of substitution and $x_j(\bar{z}_i)$ is the individual consumption level of the variety produced by firm i with knowledge input \bar{z}_i in country j. Preferences are assumed to be symmetric across countries. The mass of firms M_j and the density of their knowledge levels $\mu(\bar{z})$ are determined in the next subsection.

The total demand is given by the product of the number of customers N_j and the individual demands:

$$q_j(\bar{z}_i) = N_j x_j(\bar{z}_i). \tag{17}$$

Utility maximization subject to the individual's budget constraint implies that the demand function for product i is given by

$$p(\bar{z}_i) = q_j(\bar{z}_i)^{-\frac{1}{\sigma}} Q_j^{\frac{1}{\sigma}} P_j^{\frac{\sigma-1}{\sigma}}$$

$$\tag{18}$$

where Q_j is the consumption basket in country j and P_j denotes the price index. We normalize the domestic price index P_0 to 1.

Firms maximize profits. Each firm chooses whether and how to serve the foreign country, and determines the profit maximizing quantities q_0 and q_1 . Firms can access the foreign country via exporting or foreign direct investment. They incur fixed costs f_X for the former and f_I for the latter option, where $f_I > f_X$. Exporting from country k to country j additionally entails iceberg transport costs $\tau_{kj} > 1$, with $\tau_{01} = \tau_{10}$ and $\tau_{00} = \tau_{11} = 1$. In the following, optimal quantities are characterized by the mode superscripts D for domestic firms, X for exporters, and I for multinational firms.

The profit maximization problem of a multinationally active firm is given by

$$\max_{q_0^I, q_1^I} \quad \pi^I(\bar{z}_i, w_0, w_1) = \sum_{j=0}^1 p_j(q_j^I(\bar{z}_i)) q_j^I(\bar{z}_i) - C(\bar{z}_i, q_0^I(\bar{z}_i), w_0, q_1^I(\bar{z}_i), w_1).$$
(19)

Optimal prices are a constant mark-up over marginal costs:

$$p_j(\bar{z}_i) = \frac{\sigma}{\sigma - 1} \xi_j(\bar{z}_i, q_0^I(\bar{z}_i), w_0, q_1^I(\bar{z}_i), w_1).$$
(20)

The marginal costs ξ_j are a function of $\{q_j^I\}_{j=0}^1$ through z_h and z_j . The optimal quantities are thus implicitly defined by

$$q_j^I(\bar{z}_i) = Q_j P_j^{\sigma-1} \left(\frac{\sigma}{\sigma-1} \xi_j(\bar{z}_i, q_0^I(\bar{z}_i), w_0, q_1^I(\bar{z}_i), w_1) \right)^{-\sigma}.$$
 (21)

An exporting firm solves an analogously defined profit maximization problem:

$$\max_{q_0^X, q_1^X} \quad \pi^X(\bar{z}_i, w_0) = \sum_{j=0}^1 p_j(q_j^X(\bar{z}_i)) q_j^X(\bar{z}_i) - C(\bar{z}_i, q_0^X(\bar{z}_i) + \tau_{01} q_1^X(\bar{z}_i), w_0).$$
(22)

Optimal prices are a constant mark-up over marginal costs including transport costs τ_{01} if applicable. The marginal costs are constant. The optimal quantities are given by

$$q_0^X(\bar{z}_i) = Q_0 \left(\frac{\sigma}{\sigma - 1} \xi_0(\bar{z}_i, w_0)\right)^{-\sigma}; \qquad q_1^X(\bar{z}_i) = Q_1 P_1^{\sigma - 1} \left(\frac{\sigma}{\sigma - 1} \tau_{01} \xi_0(\bar{z}_i, w_0)\right)^{-\sigma}.$$
 (23)

The optimal production quantity of a purely domestically active firm is determined by similar considerations.

Firms can either be active purely domestically, produce domestically and export to the foreign country or produce both at home and abroad. The profits of these three options are given by

$$\pi_0^D(\bar{z}_i, w_0) = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1}\right)^{1 - \sigma} Q_0 \xi_0(\bar{z}_i, w_0)^{1 - \sigma};$$
(24)

$$\pi_0^X(\bar{z}_i, w_0) = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1}\right)^{1 - \sigma} \left(Q_0 + Q_1 P_1^{\sigma - 1} \tau_{01}^{1 - \sigma}\right) \xi_0(\bar{z}_i, w_0)^{1 - \sigma};$$
(25)

$$\pi_0^I(\bar{z}_i, w_0, w_1) = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1} \right)^{1 - \sigma} \left(Q_0 \xi_0(\bar{z}_i, q_0^I(\bar{z}_i), w_0, q_1^I(\bar{z}_i), w_1)^{1 - \sigma} + Q_1 P_1^{\sigma - 1} \xi_1(\bar{z}_i, q_0^I(\bar{z}_i), w_0, q_1^I(\bar{z}_i), w_1)^{1 - \sigma} \right).$$
(26)

+
$$Q_1 P_1^{\sigma^{-1}} \xi_1(\bar{z}_i, q_0^I(\bar{z}_i), w_0, q_1^I(\bar{z}_i), w_1)^1$$

Firms start exporting if

$$p_1(q_1^X(\bar{z}_i))q_1^X(\bar{z}_i) - \tau_{01}\xi_0(\bar{z}_i, w_0)q_1^X(\bar{z}_i) \ge w_0 f^X$$
(27)

and become multinationals if

$$\pi_0^I(\bar{z}_i, w_0, w_1) - w_0 f^I \ge \pi_0^X(\bar{z}_i, w_0) - w_0 f^X.$$
(28)

The exporting decision only depends on foreign profits because the firm is able to produce additional output without adjusting its organizational structure. Domestic profits are unaffected by the exporting decision. Foreign investment entails reallocations to the optimal allotment of knowledge inside the firm because the firm has to balance domestic and foreign requirements. This implies that domestic performance is affected by the FDI decision, so total net profits domestic and foreign net profits—in case of FDI have to exceed the total net profits in case of exporting. Unlike conventional models of FDI, the model thus features an interdependence between the marginal costs of production and the profits across countries.

2.2.1 Firm performance and modes of market access

The quantities sold in the domestic country in case of exporting and only domestic activity are equal due to the constancy of the marginal costs in $q = \{q_0^D(\bar{z}_i), q_0^X(\bar{z}_i) + \tau_{01}q_1^X(\bar{z}_i)\}$ if the firm has a single production location. Quantities sold domestically by a multinational firm are lower than domestically sold quantities if the firm produced only domestically:

$$q_0^D(\bar{z}_i) = q_0^X(\bar{z}_i) \ge q_0^I(\bar{z}_i).$$
(29)

This result arises because a multinational firm cannot tailor its headquarters to domestic needs (unless the knowledge constraint is binding at both locations). Correspondingly, domestic marginal costs increase in the quantity of foreign production (see Appendix A.2), and domestic output and profits are lower in the case of FDI than in the case of exporting or domestic activity.

In the foreign country, quantities sold in the case of FDI exceed export quantities.

$$q_1^I(\bar{z}_i) > q_1^X(\bar{z}_i) \tag{30}$$

Otherwise, the difference in fixed costs and the sales foregone in the domestic country would not be worthwhile. FDI therefore only takes place if $\xi_1(\bar{z}_i, q_0^I, w_0, q_1^I, w_1) < \tau_{01}\xi_0(\bar{z}_i, w_0)$. This is more likely to be the case for high τ_{01} and low θ_{10} , low w_1 , low c_1 , high A_1 and a large market size N_1 (see Proposition 2 and section 2.3.2 below).

2.2.2 Firm performance and foreign country characteristics

An exporting firm sells larger quantities in the domestic country than in the foreign country by $\tau_{01} > 1$ and $\sigma > 1$.

$$q_0^X(\bar{z}_i) > q_1^X(\bar{z}_i) = \tau_{01}^{-\sigma} q_0^X(\bar{z}_i)$$

Foreign sales decrease in transport costs, so country attractiveness as export destination decreases with τ_{01} .

A multinational firm's foreign performance varies with country characteristics.

Proposition 2. Foreign marginal costs ξ_1 are increasing in the communication costs with the headquarters θ_{10} , with foreign wages w_1 and foreign learning costs c_1 . Foreign marginal costs ξ_1 are decreasing in foreign labor productivity A_1 and the production quantity q_1 .

Proof. See Appendix A.2.

A multinational firm has lower foreign marginal costs in foreign countries with lower bilateral

communication costs to the multinational's home country, with lower wages, lower learning costs or higher labor productivity.

It is important to note that the marginal costs of production ξ_1 are decreasing in the foreign production quantity q_1 , but increasing in the domestic production quantity q_0 (see Appendix A.2). The firm chooses the optimal allocation of knowledge in a way that benefits affiliates with larger output. Equation (12) foreshadows this implication: the firm takes the size of its output at the different affiliates into account in choosing the optimal headquarter knowledge level z_h .

Proposition 2 implies that foreign output and foreign sales are higher in countries with lower communication costs θ_{10} , lower wages w_1 and lower knowledge acquisition costs c_1 as well as higher labor productivity A_1 (see Appendix A.2). The effects of the communication costs, wages, the knowledge acquisition costs and the labor productivity work through two channels. Take the communication costs as an example. Higher communication costs have a direct, positive effect on the marginal costs of production, which exerts a negative effect on foreign output and foreign sales. Higher communication costs also have an indirect, positive effect on the marginal costs of production as the firm takes output into account in choosing the optimal allocation of knowledge. Due to the lower production quantity, the affiliate is relatively less important for the firm. The firm adjusts the allocation of knowledge. A lower production quantity due to higher communication costs thus implies that the marginal costs of production increase even further, depressing foreign output and foreign sales. In summary, the foreign country is more attractive for FDI if the communication costs with the home country, wages and the knowledge acquisition costs are lower and the labor productivity is higher due to the positive effect of these variables on foreign output and sales.

2.3 General equilibrium

2.3.1 Closed economy

Each potential entrant can produce a differentiated product. Upon paying the fixed entry cost f in units of domestic labor, firms draw a firm-specific knowledge level \bar{z}_i . The distribution of potential knowledge levels $G(\bar{z})$ is known and defined on the interval $[0, \bar{z}_{max}]$, where \bar{z}_{max} is the highest knowledge level that fulfils Assumption 1. Production additionally entails a fixed cost of production f^D paid in domestic labor. Firms with adverse knowledge draws cannot profitably cover these fixed costs, and so exit immediately.

The general equilibrium conditions determine the cut-off knowledge level for entry, \bar{z}^* , the number of firms M, wage w and total income Q. Recall that the price index P is normalized to unity. The parameters λ, A, c, θ and N are exogenous. Country subscripts are suppressed for simplicity of exposition.

A single firm's variable profits are given by $\pi^D(\bar{z}_i) = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} Q\xi(\bar{z}_i, w)^{1-\sigma}$. The marginal costs of production ξ are constant in the production quantity q and strictly decreasing in the knowledge level \bar{z}_i (for a proof, see Appendix A.3.1). Profits are thus strictly increasing in the knowledge level \bar{z}_i .

The least productive active firm is indifferent between entering and remaining inactive: its variable profits are equal to the fixed costs of production. The zero cut-off profit condition

determines the knowledge level \bar{z}^* of the marginal entrant.

$$wf^{D} = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1}\right)^{1 - \sigma} Q\xi(\bar{z}^{*}, w)^{1 - \sigma}$$
(31)

We assume that f^D is sufficiently low such that $\bar{z}^* < \bar{z}_{max}$ exists. The resulting density of the knowledge levels of the active firms is given by $\mu(\bar{z}) = \frac{g(\bar{z})}{1-G(\bar{z}^*)}$. As profits are strictly increasing in knowledge, the cut-off is unique. The cut-off increases with the knowledge acquisition costs c, wages w and communication costs θ and decreases in labor productivity A (for a proof, see Appendix A.3.1).

Potential entrants enter up to the point where the expected net value of entry is zero: entry occurs until the expected value of profits is equal to the sunk cost of entry. The *free entry* condition determines wages w:

$$wf = \int_{\bar{z}^*}^{\bar{z}_{max}} \frac{1}{\sigma} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} Q\xi(\bar{z},w)^{1-\sigma} dG(\bar{z})$$
(32)

A unique equilibrium exists as the zero cut-off profit condition is strictly increasing in \bar{z}^* and the free entry condition is strictly decreasing in \bar{z}^* in the w, \bar{z}^* -plane (for a proof, see Appendix A.3.1).

In equilibrium, labor and goods markets clear. Labor supply N equals labor demand. Labor is used to cover the fixed cost of entry, the fixed costs of investment, and the demands for labor in production, for management, and for teaching. Labor demand for production, management and teaching can be calculated by setting wages equal to 1 in the cost function C(q, w). The *labor market clearing condition* determines the number of firms M.

$$N = M\left(f + \frac{1}{1 - G(\bar{z}^*)} \int_{\bar{z}^*}^{\bar{z}_{max}} (f^D + C(q, 1)) dG(\bar{z})\right)$$
(33)

The goods market clearing condition determines the total income wN.

$$wN = Q \tag{34}$$

2.3.2 Open economy

Firms can either sell their product only domestically, export it to the foreign market or conduct foreign direct investment. Exporting and FDI entail additional fixed costs f^X and f^I to be paid in units of domestic labor. As in Helpman et al. (2004), we assume that $f^I > \tau_{10}^{\sigma-1} f^X > f^D$. f^X and f^I are sufficiently low such that exporters and foreign investors co-exist. For simplicity, we further assume that the foreign and the domestic country are symmetric with respect to the knowledge acquisition costs $c_1 = c_0$, the labor productivity $A_1 = A_0$ and the number of agents $N_1 = N_0$. This implies that the knowledge constraint (3) is binding in the home country and slack in the foreign country, so $z_0 = \bar{z} - z_h$ and $z_1 > \bar{z} - z_h$.

The general equilibrium conditions determine the cut-off knowledge level for activity \bar{z}^* , the cut-off knowledge level for exporting \bar{z}^X and for FDI \bar{z}^I , the number of firms $\{M_j\}_{j=0}^1$, wages $\{w_j\}_{j=0}^1$ and total income $\{Q_j\}_{j=0}^1$. We derive conditions for the home country; analogous conditions hold for the foreign country.

As above, the least productive active firm is indifferent between entering and remaining

inactive. The variable domestic profit is equal to the fixed costs of production. The corresponding zero cut-off profit condition $w_0 f^D = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} Q_0 \xi_0(\bar{z}^*, w_0)^{1-\sigma}$ determines the knowledge level \bar{z}^* of the marginal entrant.

The marginal exporter is indifferent between exporting and not exporting: the variable foreign export profits are equal to the fixed costs of exporting. The second zero cut-off profit condition determines the exporting cut-off \bar{z}^X .

$$w_0 f^X = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1}\right)^{1 - \sigma} Q_1 P_1^{\sigma - 1} (\tau_{01} \xi_0(\bar{z}^X, w_0))^{1 - \sigma}$$
(35)

Domestic and export profits are strictly increasing in \bar{z}_i as the marginal costs $\xi_0(\bar{z}_i, w_0)$ are strictly decreasing in \bar{z}_i . The marginal costs of the marginal exporter are lower than the marginal costs of the marginal entrant if the market sizes are similar, as in Melitz (2003):

$$\xi_0(\bar{z}^X, w_0) = \left(\frac{Q_1 f^D}{Q_0 f^X}\right)^{\frac{1}{\sigma-1}} \frac{P_1}{\tau_{01}} \xi_0(\bar{z}^*, w_0) \quad <\xi_0(\bar{z}^*, w_0) \quad \text{for } Q_0 \approx Q_1 P_1^{\sigma-1}.^3$$

Due to the symmetry of home and foreign in the exogenous parameters λ , θ_{j0} , c_j , A_j , N_j and the symmetry in the distribution of knowledge levels $G(\bar{z})$, in equilibrium, $w_1 = w_0$, $Q_1 = Q_0$ and $P_1 = P_0 = 1$. Consequently, the exporting cut-off knowledge level is higher than the zero cut-off knowledge level: $\bar{z}^X > \bar{z}^*$. An increase in τ_{01} implies that the exporting cut-off knowledge level \bar{z}^X increases: it is profitable to export to more distant destinations only for firms with lower marginal costs.

The marginal multinational firm is indifferent between exporting and foreign direct investment. The net total export profits are equal to the net total profits earned from FDI. The multinational cut-off \bar{z}^I is determined by the third *zero cut-off profit condition*, where we use $w_1 = w_0, Q_1 = Q_0$ and $P_1 = P_0 = 1$.

$$\frac{1}{\sigma} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} Q_0 \left(\xi_0(\bar{z}^I, q_0^I(\bar{z}^I), w_0, q_1^I(\bar{z}^I), w_0)^{1-\sigma} + \xi_1(\bar{z}^I, q_0^I(\bar{z}^I), w_0, q_1^I(\bar{z}^I), w_0)^{1-\sigma}\right) - w_0 f^I = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \xi_0(\bar{z}^I, w_0)^{1-\sigma} Q_0(1+\tau_{01}^{1-\sigma}) - w_0 f^X$$
(36)

The total profits from FDI have to exceed the total profits from exporting plus the difference in the fixed costs. Both total FDI and total export profits increase in the knowledge level \bar{z} . As the fixed costs of foreign direct investment are higher than the fixed costs of exporting by a factor of more than $\tau_{01}^{\sigma-1}$, firms have to have a higher knowledge level to profitably carry out foreign investment: $\bar{z}^I > \bar{z}^X$. Domestic profits decrease in case of FDI as the headquarters are no longer tailored to domestic needs, but balance domestic and foreign requirements. Compared to a model with independent marginal costs of production, the productivity cut-off is thus shifted upwards (for a proof, see Appendix A.3.2).

Export profits decrease with τ_{01} and profits from FDI decrease θ_{10} . Consequently, the knowledge cut-off \bar{z}^I varies with these parameters.

³Strictly speaking, $\xi_0(\bar{z}^X, w_0) < \xi_0(\bar{z}^*, w_0)$ if $Q_1 P_1^{\sigma-1} < \tau_{01}^{\sigma-1} \frac{f^X}{f^D} Q_0$. The marginal exporter has lower marginal costs than the domestic entrant whenever $Q_1 P_1^{\sigma-1} \leq Q_0$ for sure, or when the foreign market size adjusted by the price index does not exceed the domestic market size by more than a multiple of $\tau_{01}^{\sigma-1} \frac{f^X}{f^D} > 1$.

Proposition 3. The knowledge cut-off \bar{z}^I increases in the bilateral communication costs θ_{10} between the host country and the multinational's home country, and decreases in the bilateral transportation costs τ_{01} between the home and the host country.

Proof. See Appendix A.3.2.

The depressing effect of the bilateral communication costs θ_{10} on the tendency to invest decreases in λ . A higher value of λ is associated with a more predictable production process. The productivity cut-off \bar{z}^I is thus even higher in sectors with less predictable, more complex production processes. This result is consistent with the findings in Bahar (2013) who shows that firms in knowledge intensive sectors are less likely to conduct horizontal foreign direct investment, and that given investment, high communication costs have a negative effect on the location probability of knowledge intensive activities.

Firms enter up to the point where the net value of entry is zero. The expected value of the profits has to be adjusted for export and multinational profits. The *free entry condition* is given by

$$w_0 f = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1} \right)^{1 - \sigma} Q_0 \left(\int_{\bar{z}^*}^{\bar{z}^I} \xi_0(\bar{z}, w_0)^{1 - \sigma} dG(\bar{z}) + \int_{\bar{z}^X}^{\bar{z}^I} (\tau_{01} \xi_0(\bar{z}, w_0))^{1 - \sigma} dG(\bar{z}) + \int_{\bar{z}^I}^{\bar{z}_{max}} \left(\xi_0(\bar{z}, q_0(\bar{z}), w_0, q_1(\bar{z}), w_0)^{1 - \sigma} + \xi_1(\bar{z}, q_0(\bar{z}), w_0, q_1(\bar{z}), w_0)^{1 - \sigma} \right) dG(\bar{z}) \right)$$
(37)

The *labor market clearing condition* is modified to include the demand for labor to cover the fixed costs of foreign activity and the demand for labor by foreign investors.⁴

$$N_{0} = M_{0} \left(f + \frac{1}{1 - G(\bar{z}^{*})} \left(\int_{\bar{z}^{*}}^{\bar{z}^{I}} f^{D} + C(q \in \{q_{0}^{D}, q_{0}^{X} + \tau_{01}q_{1}^{X}\}, 1) dG(\bar{z}) + \int_{\bar{z}^{X}}^{\bar{z}^{I}} f^{X} dG(\bar{z}) \right. \\ \left. + \int_{\bar{z}^{I}}^{\bar{z}_{max}} f^{D} + f^{I} + C(q_{0}, 1, q_{1}, 1) dG(\bar{z}) \right) \right)$$
(38)

The goods market clearing condition corresponds to the above equation, with the consumption basket adjusted for the available foreign goods. Due to symmetry, the trade balance condition is fulfilled.

2.4 Multinational wage premiums

A large empirical literature finds that foreign owned firms typically pay higher wages than domestic firms (e.g. Aitken et al., 1996; Lipsey, 2004; Hijzen et al., 2013). The foreign wage premium tends to be higher in developing countries (e.g. Aitken et al., 1996; Hijzen et al., 2013) and varies with the nationality of the acquirer (Girma and Görg, 2007). Likewise, parent companies of of MNEs pay higher wages than domestic firms (Heyman et al., 2007). Such wage premiums exist even after controlling for firm and industry characteristics, and after taking the endogeneity of foreign take-overs into account (Girma and Görg, 2007; Hijzen et al., 2013).⁵

 $^{^{4}}$ In principle, the demand for domestic labor by domestic multinational firms and the demand for domestic labor by multinational firms from the foreign country need to be included separately. By symmetry, these labor demands sum up to the total demand for labor by a domestic multinational firm.

⁵Part of the multinational wage premium is attributable to worker heterogeneity, which we abstract from. Analyses that focus on wage premiums for incumbent workers find considerably smaller wage premiums than analyses at the firm or establishment level (e.g. Heyman et al., 2007). However, the wage premium cannot be

These empirical patterns can be explained by the predictions on the wage setting behavior of multinational firms in the present model. Multinational wage premiums occur via two different channels, a selection effect and residual multinational wage premiums.

Only firms with higher knowledge level \bar{z} become multinationals, as demonstrated in subsection 2.3.2. These firms pay higher wages to managers and production workers both in their home country and the foreign country due to the positive effect of \bar{z} on z_h , z_0 and z_1 (see Proposition 1). The resulting wage premium does not stem from multinationality per se, but from a firm characteristic—knowledge—that favors the selection of firms into becoming multinational and leads to higher wages. The channel thus works similarly to previous explanations that attribute multinational wage premiums to differences in firm characteristics between multinational and domestic firms, such as differences in labor demand volatility or closure rates (for a survey, see Malchow-Møller et al., 2013, who categorize these explanations as "heterogeneous firm" theories).

Residual multinational wage premiums, in contrast, are wage premiums paid by multinational firms compared to domestic firms with the same observable characteristics. Such wage premiums occur in the model under certain parameter conditions.

Proposition 4. Multinational firms pay higher wages to their production workers than domestic firms with the same marginal costs in their home country if $\theta_{00}w_1c_1 < \theta_{10}w_0c_0$. Multinational firms pay higher wages than domestic firms with the same marginal costs in the foreign country if $\theta_{00}w_1c_1 < \theta_{10}w_0c_0$ and $c_1 \ge c_0$.

Proof. See Appendix A.4.

To see this, take the firm at the cut-off knowledge level for foreign investment \bar{z}^I . The parameter condition $\theta_{00}w_1c_1 < \theta_{10}w_0c_0$ implies that the knowledge constraint (3) is binding in the home country and slack in the foreign country. A comparison of equations (12) and (13) demonstrates that for the firm with $\bar{z} = \bar{z}^I$, a higher z_h is optimal as domestic firm than as a multinational firm. This implies that the firm pays higher wages to its production workers in the home country as a multinational than as a domestic firm or exporter. The result holds more generally: given their knowledge level \bar{z} , firms pay higher wages to their domestic production workers as multinational than as domestic producers.⁶

As the marginal costs of production are decreasing in z_h given \bar{z} , firms have lower marginal costs in their domestic market as domestic producers than as multinational firms. A domestic producer with the same observed marginal costs as a multinational firm is therefore characterized by a lower knowledge level \bar{z} . The multinational wage premium in the domestic market is thus reinforced by the positive effect of \bar{z} on production knowledge.

Concerning the foreign country, a multinational firm with cut-off knowledge level \bar{z}^I pays higher wages than a foreign domestic firm with the same knowledge level by $z_1 > z_0 > \bar{z}^I - z_h^{Foreign}$, where z_0, z_1 denote the production worker knowledge levels of the multinational firm and $\bar{z}^I - z_h^{Foreign}$ refers to the production knowledge of a foreign domestic producer. The

fully attributed to worker heterogeneity (Malchow-Møller et al., 2013).

⁶Given the assumptions in the paper, all firms with knowledge above the knowledge cut-off \bar{z}^{I} select into FDI and all firms with knowledge below the cut-off into exporting or domestic activity. Multinational firms and domestic producers are thus generally characterized by dissimilar knowledge levels, with exception of those firms around the cut-off \bar{z}^{I} . A simple extension of the model with uncertain fixed costs of foreign market entry, $\tilde{f}^{I} = f^{I} + \epsilon_{i}, \epsilon_{i} \sim N(0, \nu^{2})$, allows multinationals and domestic firms with similar knowledge levels to arise.

first inequality follows from the knowledge constraint being slack in the foreign country. The second inequality results because managerial knowledge is increasing in c_j . Given \bar{z} and $c_1 \geq c_0$, production workers employed by domestic producers in the foreign country earn (weakly) lower wages than production workers employed by domestic producers in the home country. As multinationals from the home country pay higher wages to their domestic production workers than comparable domestic firms in the home country, the second inequality results.

As before, a foreign domestic producer is has lower marginal costs than a multinational firm with the same knowledge level. A foreign domestic producer thus has lower firm knowledge than a multinational firm with the same marginal costs, so the residual foreign multinational wage premium is reinforced by the impact of overall firm knowledge on wages.

With respect to the empirical relevance of this finding, it is necessary to assess to which country pairs the parameter restriction $\theta_{00}w_1c_1 < \theta_{10}w_0c_0$ applies. By assumption, there is a friction in cross-border communication compared to within country communication, i.e. $\theta_{10} > \theta_{00}$. The condition thus holds whenever the product of foreign wages and knowledge acquisition costs is lower than the product of the corresponding domestic variables adjusted by the communication friction, $w_1c_1 < \frac{\theta_{10}}{\theta_{00}}w_0c_0$. This includes the case $w_0c_0 \approx w_1c_1$, which can reasonably be argued to apply to FDI from developed countries to other developed countries. Concerning FDI from developing than in developed countries, knowledge acquisition costs are likely to be higher in developing than in developed countries, $c_1 > c_0$, and wages are typically much lower, $w_1 < w_0$. Wage premiums occur whenever the difference in wages adjusted by the communication friction outweighs the difference in knowledge acquisition costs. Note that in addition, the foreign wage premium is predicted to be the higher, the higher c_1 is. Consistent with the empirical evidence, MNE wage premiums are thus predicted to be stronger for developing than for developed countries.

As both communication costs between the home country of the multinational firm and the foreign country and relative wages and knowledge acquisition costs in the domestic and foreign play a role, the model additionally explains why wage premiums vary with the nationality of the acquirer. All implications result from the cost minimization problem of the firm and additional assumptions on worker preferences, as in Egger and Kreickemeier (e.g. 2013), are not necessary

3 Empirical analysis

The empirical analysis is based on the Microdatabase Direct investment (MiDi), a data set with balance sheet information on virtually the universe of foreign affiliates of German multinational firms from 1999 until 2010. The data set contains detailed balance sheet information on every affiliate, including its sales, its number of employees and its financial structure. Information on parent and affiliate sectors, mostly at the two digit level, is also provided. From 2002 onwards, information on the sales and the number of employees of the German investor is available.

The data allow testing the following predictions of the model:

Testable hypotheses.

 The foreign productivity of a given multinational firm increases in the size of the foreign market Q
_j and the labor productivity A
_j and decreases in the bilateral communication costs with Germany θ
_{j0}, foreign wages w
_j and knowledge acquisition costs c
j. 2. The cut-off productivity level for investment in a foreign country increases with the bilateral communication costs with Germany $\tilde{\theta}{i0}$, and decreases with the bilateral transport costs τ_{0i} .

The first hypothesis follows from Proposition 2 and the second hypothesis results from Proposition 3 by associating productivity with inverse marginal costs $\frac{1}{\xi_j}$, as is standard in the literature (e.g. Melitz, 2003; Helpman et al., 2004). A "~" is used to denote the empirical analog of a model parameter. The production quantity q_j is an endogenous model outcome, so market size is used to take into account the fact that demand and thus q_j will be higher in larger markets.

We use the natural log of sales over employees $\ln(sales/employees)$ to measure productivity. To be consistent with the level of analysis of the model, we aggregate the available information on the parent-country level, taking the degree of participation of the parent in the affiliate into account where applicable.

Recall that the model demonstrates that a firm's marginal cost depends on \bar{z}_i , the knowledge level of the firm, which also determines firm selection across countries. To be consistent with the model predictions, the empirical test of Hypothesis 1 focuses on within-parent variation in performance across countries, thus taking into account the fact that differences in \bar{z}_i may affect foreign performance. We estimate the following regression equation:

$$y_{ijt} = \beta_0 + \beta_1 \tilde{Q}_{jt} + \beta_2 \tilde{\theta}_{jt} + \beta_3 \tilde{c}_{jt} + \beta_4 \tilde{w}_{jt} + \beta_5 \tilde{A}_{jt} + \alpha_{it} + \epsilon_{ijt}$$
(39)

where y_{ijt} denotes the foreign labor productivity of firm *i* in country *j* and period *t*, α_{it} is a parent-year fixed effect, and ϵ_{ijt} is a firm-country-year specific error term. The parent-year fixed effects absorb the effect of \bar{z}_i , i.e., of parent characteristics that may influence performance across destinations. The regression results uncover the relation between deviations in the performance measure y_{ijt} from its parent-year specific mean and deviations in the country characteristics \tilde{x}_{jt} from their respective parent-year specific means, abstracting from parent-specific differences. To account for correlations across time, the standard errors are clustered at the level of the parent.

This approach helps to distinguish the predictions of the model from bias due to the selfselection of firms across countries. Large markets with high labor productivity and low factor costs are attractive investment destinations. The higher average productivity of firms in such markets could result from more productive multinationals selecting into those destinations. The empirical specification focuses on the performance of the same multinational firm across different countries and thus tests the model in a clean fashion.

It is still necessary to interpret the resulting estimates $\hat{\beta}$ with caution. The set of locations is a choice variable on the part of the firm, and does not vary exogenously. It is difficult to guarantee that our estimation conditions on all information available to the investor, so the results may still be biased due to unobservable firm–country-specific variables. One possibility for addressing this problem would be to analyze the effect of a large and unexpected shock in one of the regressors in a country on the outcome variables of firms which had already been active in that country before the shock occurred. We have unfortunately not been able to find such a shock, and thus have to stick with more suggestive empirical evidence. Estimating a Heckman selection model is an alternative option. However, such a specification does not allow maintaining a strict focus on within-parent variation across countries, as using firm–year fixed effects is not possible. In addition, the strong distributional assumptions of the Heckman selection model are not fulfilled in the data. These aspects are discussed in more detail in Appendix C.

To learn whether the selection pattern across countries is consistent with the predictions of the model, we construct the cut-off productivity level by country for each two digit parent sector group. The model predicts that firm performance across countries is interdependent and that foreign activity negatively affects domestic performance. We therefore employ two different strategies to approximate the cut-off productivity level. We employ the minimum domestic productivity of investors active in a country, which is subject to the caveat just mentioned. We additionally construct the global productivity of a firm as the log of global sales over the global number of employees and use the minimum worldwide productivity of investors in a country as measure for the cut-off productivity level. The estimation equation is:

$$y_{sjt} = \gamma_0 + \gamma_1 \tilde{Q}_{jt} + \gamma_2 \tilde{\theta}_{jt} + \gamma_3 \tilde{c}_{jt} + \gamma_4 \tilde{w}_{jt} + \gamma_5 \tilde{A}_{jt} + \delta_t + \delta_s + \upsilon_{sjt}$$
(40)

where y_{sjt} denotes the cut-off productivity level of firms in sector s in country j and period t, δ_t , δ_s are a year and sector fixed effect, and ϵ_{sjt} is a sector-country-year specific error term. Standard errors are clustered at the country level.

4 Data and descriptive statistics

The Microdatabase Direct investment (MiDi) is provided by the Bundesbank, the German central bank. We use the information on outward foreign direct investment by German companies. The database consists of a panel of yearly information on virtually the universe of foreign affiliates of German firms from 1999 until 2010. By the German Foreign Trade and Payment Regulation (Außenwirtschaftsverordnung), any resident who holds shares or voting rights of at least 10% in a company with a balance sheet total of more than 3 million euro is obliged to report information on the financial characteristics of these affiliates to the Bundesbank. Until 2002, information on stakes of at least 10% in a company with a balance sheet total of more than 5 million euro and stakes of at least 50% in a company with a balance sheet total of more than 0.5 million euro had to be reported (Lipponer, 2009). The same information has to be provided on branches or permanent establishments abroad if their operating assets exceed the reporting threshold. We clean the data so that all observations meet a uniform threshold: we keep affiliates with a balance sheet total of at least 5 million euro, but parent stakes of at least 10%, or with a balance sheet total between 3 and 5 million euro, but parent stakes of at least 50%.⁷

We drop observations on 20,016 affiliates of investors that are government institutions or private households, agriculture or mining companies and housing enterprises.⁸ The resulting data set contains 271,178 affiliate-year observations that correspond to 179,658 parent-country-year observations. Some affiliates are reported several times, because multiple investors hold participating interests in them. We augment the MiDi with information on country characteristics used as proxies for the model parameters. An overview of the model parameters, the empirical analogs, and the corresponding data source is given in Table 2.

 $^{^737,\!251}$ observations are thus dropped from the sample.

⁸Government institutions and private households are dropped because they are not multinational firms. Agriculture and mining companies are dropped because natural resources are decisive factors for their investments, but ignored in the theoretical and empirical analyses. Housing enterprises are dropped because they often report sales of zero, even though they are not small, which would lead to measurement error in our analysis.

Variables	Definition	Empirical analog	Data source
Endogeno	us outcomes		
ξ_j	Marginal costs of production	Inverse for eign labor productivity (ln (sales/# employees))	MiDi
Firm char	racteristics		
$ar{z}_i$	Knowledge level	Parent-year fixed effect α_{it}	n/a
Country c	haracteristics		
Q_j	Market size	GDP	IMF
		Market potential (sum of GDP weighted by distance)	Own calculation/ CEPII, IMF
$ heta_{j0}$	Bilateral communication costs	Flight time from Frankfurt to main city	www.weltinfo.com/service/flugzeiten.html,
			www.meine-flugzeit.de, main city: CEPII
		Linguistic distance	Own calculation/ Spolaore and Wacziarg (2009),
			CEPII, Ethnologue
		Share of population speaking English	Eurobarometer 52, 55, 237, 243, 386
			Crystal (2003)
		Common official language indicator	CEPII
		Time zone differences in minutes	www.timeanddate.com
c_j	Knowledge acquisition costs	Per pupil public expenditure on education	
		as $\%$ share of GDP per capita	World Bank (World DataBank)
		PISA maths, science scores	World Bank (World DataBank)
A_j	Labor productivity	GDP per employee	World Bank (World DataBank)
		Labor productivity per employee	OECD
w_j	Wage	Compensation per employee	OECD
$ au_{0j}$	Transport costs	Distance	CEPII
Additional	l controls		
	Social capital	Bilateral trust	Eurobarometer 46
	Investment climate	Statutory tax rate	IBFD
		Cost and time to start a business	World Bank (Doing Business)
		Cost, time and $\#$ of procedures to enforce contracts	World Bank (Doing Business)
		Cost, time and $\#$ of procedures to register property	World Bank (Doing Business)
		Rule of law, government effectiveness	World Governance Indicators

Table 2: Overview of model parameters and their empirical analogs

We use GDP data from the IMF to measure differences in demand across countries. We alternatively employ market potential calculated as the sum of GDP of all countries weighted by their distance to the host country (e.g. Chen and Moore, 2010). The bilateral communication costs between headquarter managers and foreign production workers are difficult to capture. We employ a number of different measures to approximate them. We focus on measures that approximate frictions in information and communication flows between people and refrain as far as possible from measures which could also capture differences in the flow of goods such as distance. The main measures are the duration of a flight from Frankfurt to the main city of the host country, the linguistic distance between German and the language(s) spoken in the host country, and the difference in time of day in minutes. We calculate the linguistic distance as a function of the number of common nodes of German and each language spoken by at least 20%of people in the host country according to the classification of languages provided by Ethnologue (Spolaore and Wacziarg, 2009) and take the simple average in case of several languages.⁹ The flight time is used to capture how quickly managers can travel to the affiliate and address potential issues in the production process on site. Despite technological advances, face-to-face communication is often indispensable to ensure successful production (e.g. UNCTAD, 2004). In principle, the flight time is correlated with geographic distance. However, there are no direct flights for a number of destinations, so the flight time contains important information on the accessibility of foreign countries beyond their geographic distance. The linguistic distance is based on the intuition that the closer a language is to one's mother tongue, the easier it is for most people to learn it and to express themselves precisely. Undoubtedly, international business communication likely often takes place in English (and we thus use the share of the population speaking English as measure for the communication costs in robustness checks below). Still, as a recent documentation by the Secretariat General Translation Directorate of the European Court of Auditors shows, non-native English speakers tend to develop their own English dialect that is strongly influenced by their native languages and often difficult to understand by native speakers (Gardner, 2013). The linguistic distance is therefore apt to capture frictions in communication despite the use of English in business contexts. The difference in time of day captures the fact that personnel at either location may have to work unconventional hours to facilitate communication if the affiliate is located in a different time zone. This may increase costs for the firm. We use the difference in the time of day in minutes and not the abosolute value of the deviation in time of day, because the time difference may have an asymmetric effect. The working day at the affiliate location starts before the parent's working day if the time difference is negative, but ends after the parent's working day if the time difference is positive. This may affect affiliate performance because in the former case, parents can be approached with problems at the same day, whereas affiliates may have to wait until the next day to have their problems addressed in the latter case. As robustness checks, we use the share of people speaking English and an indicator that German is among the official languages of the host country. The share of people speaking English is only available for a subset of countries and for a subset of years, but does not vary greatly over time. We therefore assign the available information for each country to all years and use the nearest value for each period.

Knowledge acquisition costs are measured using the expenditure per pupil as percentage

⁹Specifically, we use $\sqrt{\frac{7-\# \text{ common nodes}}{7}}$, slightly modifying the formula used in Spolaore and Wacziarg (2009), as 7 is the number of linguistic nodes of German and thus the maximum number of common nodes possible.

share of GDP per capita and PISA maths and science scores. The former is an input measure based on the idea that more public investment in education decreases the marginal costs of schooling. The latter are output measures that take into account the fact that large amounts of public funds per se do not imply that the educational system works efficiently. PISA scores are internationally comparable measures of how well the educational system is able to teach abilities to students. As they are only available for a relatively small subset of countries and years, we prefer the former measure. We use GDP per employee to measure labor productivity. This measure is imperfect, as it is does not reflect differences in hours worked, but is available for a large number of countries. We use labor productivity per employee for OECD countries as robustness check. To measure wages, we employ the compensation per employee provided by the OECD. However, these measures are only available for the majority of OECD countries, and the OECD countries are a non-representative sub-sample of all investment destinations of German firms. We therefore stick to GDP per employee as labor productivity measure and omit wages in most regressions below.¹⁰ The omission of wages is defendable, as wages are endogenously determined in general equilibrium and thus a function of the other covariates. Finally, transport costs are measured using data on bilateral distance. We take the logarithms of the covariates if their distribution in levels is skewed.

To take into account factors which may influence the variation in productivity across countries but are not explicitly included in the model, we also include data on bilateral trust and investment climate measures. We use data on bilateral trust from Germany towards other countries from the Eurobarometer. As investment climate measures, we use statutory tax rates as well as the indicators on the ease of enforcing contracts and of registering property from the World Bank Doing Business Indicators and indicators on the rule of law and government effectiveness from the World Governance Indicators. The cost and time to start a business are used as additional controls in the regressions on the cut-off productivity level.

Table 3 provides summary statistics of the variables used in the regression analysis.

 $^{^{10}}$ We have experimented with the Occupational Wages around the World Database by Freeman and Oostendoorp provided via the NBER (Oostendorp, 2013), a standardized database on wages by sector based on the ILO October Inquiries. However, its overlap with the MiDi database is even worse. Data are available for only 20% of observations.

Variable	Ν	Mean	SD
Foreign productivity	150, 570	5.583	1.189
Domestic productivity (firm level)	24,357	5.803	1.150
Number of countries per parent (firm level)	54,961	3.016	4.782
Log GDP	163,890	13.269	1.533
Log market potential	163, 686	6.457	1.431
Log flight time	165, 349	5.154	.959
Time difference to Germany, minutes	165, 662	0.198	3.588
Linguistic distance	163,805	0.803	0.232
Official language German	164,992	0.139	0.346
Per pupil public expenditure on education, % of GDP p.c.	128,653	22.978	3.982
PISA maths score	46,329	492.354	38.925
PISA science score	46,329	495.513	32.579
GDP per employee (in 1990 $$1,000$)	163,071	3.806	1.581
Labor productivity per employee, OECD	123, 239	11.532	1.390
Log distance	164,992	7.320	1.298
Log compensation per employee	113, 157	11.087	1.432
Bilateral trust from Germany, survey 1996	120,875	2.548	0.421
Log costs to enforce a contract	114,971	3.009	0.363
Log time to enforce a contract	114,971	6.162	0.439
# of procedures to enforce a contract	114,971	32.941	5.680
Log costs to register property	102, 230	1.074	1.010
Log time to register property	102, 645	3.416	1.050
# of procedures to register property	102,645	5.323	2.201
Costs to start a business	114,980	10.080	37.595
Time to start a business	114,980	24.525	26.056
Statutory tax rate	164, 110	28.933	7.672

Table 3: Summary statistics

Summary statistics for regression sample. *Variable definitions:* see Table 2. The number of observations varies due to differences in country coverage. Maximum possible number of observations: 165,760.

Figures 1 to 4 provide graphical evidence about the basic relations in the data. Each dot refers to one country. The upper panel scatters the cut-off productivity level by country, where the cut-off productivity level is measured using domestic productivity. The lower panel scatters the average deviation of foreign productivities from the parent–year specific mean. The average deviation of foreign productivities from the parent–year specific mean is calculated as follows. We calculate the mean of foreign productivities for each parent and year and substract it from the observed foreign productivities. We obtain the deviation in the foreign productivity per country from the parent–year specific mean. For every country, we calculate the average of these deviations. This implies that the scatter plots indicate the relationship between country characteristics and the relative performance of firms, i.e. whether firms tend to be more or less productive than average in countries with certain characteristics. The left figures scatter the respective variables against log GDP, the right figures scatter the residuals from a regression of the variables on GDP against log flight time. The relationship between performance and flight time are therefore robust to GDP differences.

Figure 1: Cut-off parent productivity vs. GDP, 2005

Figure 2: Cut-off parent productivity vs. flight time, 2005



Figure 1 scatters the minimum domestic productivity of investors active in a country versus the natural log of country GDP. Figure 2 scatters the residuals from a regression of the minimum domestic productivity on log GDP versus the natural log of the flight time from the main city in the host country to Germany. For data confidentiality, only countries where at least three investors are active are included.



Figure 4: Foreign productivity vs. flight time, 2005, deviations from parent mean



Figure 3 scatters the average of the deviations of foreign productivity (In foreign sales/foreign employees) in a country from the parent specific mean versus the log of host country GDP. Figure 4 scatters the residual from a regression of the average deviation of foreign productivity on log GDP versus the log of flight time between the host country and Germany. For data confidentiality, only countries where at least three investors are active are included. Two outliers, the Bahamas and Mauritius, commonly considered as tax havens, are dropped.

The model predicts that only more knowledge intensive and thus more productive firms invest in more challenging destinations (cf. Proposition 3). Figures 1 and 2 corroborate the conjecture that firm selection is non-random across countries: the cut-off productivity level as measured by the minimum domestic productivity level tends to be higher in smaller markets, and lower in countries that are quicker to reach.

Figures 3 and 4 show that firms tend to be relatively more productive in countries that are **Serie** larger and that are more quickly accessible. The relationships are reversed compared to the patterns displayed in Figures 1 and 2. Larger and less remote countries attract investors that are less productive in the home country, but foreign productivity tends to be higher in larger and less remote countries given domestic productivity. The pattern is consistent with higher \bar{z}_i firms' being able to invest in more difficult destinations, but all firms being more productive in larger markets with lower bilateral communication costs with the home country.

To run the regressions in the next section, we aggregate direct and indirect participation

interests per affiliate and further restrict the sample to majority owned affiliates.¹¹ The relationships predicted by the model hold whenever the parent is actively involved in the management of local production. This may not be the case when the majority of the affiliate is owned by other shareholders. The results are largely robust if we abandon this restriction. We report the regression results for the full sample in Appendix B.

5 Results

5.1 Within-firm differences in performance across countries

Table 4 presents the main results on Hypothesis 1. The table displays the number of parents and the number of country combinations contained in the regression sample along with the coefficients. The number of country combinations denotes the number of distinct combinations of countries that the investors included in the regression sample are active in. The number is decisive because the regression results are driven by the variation within firms across countries, so it is important to ensure that the number is sufficiently high. The number of country combinations often exceeds the number of parents, which implies that parents change their set of investment destinations over time.

Table 4 shows nine specifications. The first specification (column 1) includes log GDP and the measures of communication costs: log flight time in minutes, linguistic distance and the time difference from Germany in minutes. The second specification (column 2) adds knowledge acquisition costs measured by public expenditures on education per pupil as share of GDP per capita and labor productivity measured as GDP per employee. The third specification (column 3) also includes a measure for wages: the log compensation per employee. We separate the first from the other specifications to make transparent whether the measures of communication costs take up omitted country characteristics. The second and third specification are separately displayed because the wage data are only available for OECD countries, so the sample size decreases non-randomly once wages are included. Due to the selectivity of the sample including wages and the associated decrease in the number of observations, we omit wages from further regressions. Columns 4 to 9 include other potential determinants of cross-country differences in productivity that might be omitted variables in the main specification. These specifications will be detailed after discussing the main regression results. Table B.1 presents the regression results if all affiliates, including non-majority owned affiliates, are included in the regression sample. The regression coefficients are quantitatively similar, but less significant, which indicates that non-majority owned affiliates add noise to the sample. Table B.2 displays regression results for columns 4-9 including wages. The results are largely robust, but the number of observations is considerably lower. Recall that these regressions are based on a non-random subsample of countries.

We expect that parents are more productive in larger and more proximate countries with higher labor productivities and lower knowledge acquisition costs and wages.

 $^{^{11}24,\!060}$ observations are thus dropped from the sample.

	M	ain specification			Alternative determinants				
	1	2	3	4	5	6	7	8	9
Log GDP	0.140***	0.057***	0.081***	0.098***	0.072***	0.035***	0.049***	0.102***	0.053***
	(0.006)	(0.008)	(0.011)	(0.011)	(0.008)	(0.009)	(0.009)	(0.012)	(0.008)
Log flight time	-0.155^{***}	-0.044^{***}	-0.005	-0.028^{*}	-0.034^{**}	-0.057^{***}	-0.080^{***}	-0.055^{**}	-0.105^{***}
	(0.009)	(0.011)	(0.013)	(0.013)	(0.011)	(0.013)	(0.014)	(0.018)	(0.026)
Linguistic distance	-0.631^{***}	-0.071^{*}	-0.064^{+}	-0.084^{**}	-0.059^{*}	-0.147^{***}	-0.059^{+}	-0.069^{+}	-0.067^{*}
	(0.029)	(0.030)	(0.034)	(0.032)	(0.030)	(0.036)	(0.033)	(0.035)	(0.030)
Time difference to Germany	0.002	0.020***	0.036^{***}	0.042^{***}	0.019^{***}	0.024^{***}	0.019^{***}	0.046^{***}	0.019^{***}
	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)
Public expenditure per pupil,		0.016^{***}	0.018^{***}	0.013^{***}	0.017^{***}	0.018^{***}	0.019^{***}	0.017^{***}	0.015^{***}
% GDP per capita		(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)
GDP per employee		0.207^{***}	0.151^{***}	0.108^{***}	0.211^{***}	0.235^{***}	0.189^{***}	0.144^{***}	0.215^{***}
		(0.008)	(0.012)	(0.015)	(0.008)	(0.009)	(0.009)	(0.017)	(0.009)
Log compensation per employee			-0.013^{*}						
			(0.006)						
Trust, survey 1996				0.231^{***}				0.194^{***}	
				(0.032)				(0.038)	
Statutory tax rate					-0.008^{***}			-0.008^{***}	
					(0.001)			(0.002)	
Log costs to enforce contracts						0.029			
						(0.025)			
Log time to enforce contracts						-0.004		0.177^{***}	
						(0.022)		(0.019)	
# of procedures to enforce contracts						0.012^{***}			
						(0.002)			
Log costs to register property							-0.002		
							(0.007)		
Log time to register property							-0.079^{***}	-0.077^{***}	
							(0.008)	(0.008)	
# of procedures to register property							0.010^{**}		
							(0.003)		
Log distance									0.046^{**}
									(0.018)
Constant	5.026***	3.960^{***}	3.767^{***}	3.264^{***}	3.895^{***}	3.775^{***}	4.495^{***}	2.577^{***}	4.002^{***}
	(0.078)	(0.120)	(0.159)	(0.159)	(0.119)	(0.195)	(0.151)	(0.259)	(0.121)
R^2	0.092	$0.19\overline{4}$	0.136	0.138	$0.19\overline{6}$	0.204	0.208	0.159	$0.19\overline{5}$
Observations	149,223	117,519	91,600	99,061	117, 315	82,567	72,792	59,728	117,519
# parents	8,331	7,732	6,867	7,239	7,724	6,253	5,903	5,489	7,732
# country combinations	9,249	9,042	8,628	8,638	9,040	6,747	6,140	5,826	9,042

Table 4: Regression results: within-firm differences in productivity across countries

Standard errors in parentheses. + p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001. *Dependent variable:* foreign productivity defined as $\ln \frac{\text{foreign sales}}{\text{foreign employees}}$. Covariate definitions: see Table 2. Sample based on all foreign affiliates with degree of participation of at least 50%.

All signs of the coefficients are in line with the theoretical predictions. The estimated coefficients are generally highly significant, often at the 0.1% level. Parents tend to be relatively more productive in relatively larger countries. Higher communication costs—as measured by longer flight times and higher linguistic distance—generally have a negative impact on the outcome variables, though the effect is less significant once the sample is reduced to OECD countries. Positive time differences increase productivity. It is plausible to interpret this finding as an indication that a longer time lag between the incidence of a problem at a foreign location and the reaction of the parent is productivity decreasing. If affiliates work ahead of the parent, it is possible for them to send the issues to the headquarters at the beginning of the headquarters' office hours and have them addressed the same day; if the working day at affiliate locations begins after the parent's, problems that arise later during the day have to be postponed until the next day. In unreported regressions, we explore whether the relationship is non-linear, but do not find robust evidence for a U-shaped relationship. Relatively lower knowledge acquisition costs due to higher public investments in education and relatively higher labor productivity increases foreign performance, whereas higher wages decrease foreign productivity.

These findings are robust to the inclusion of potential omitted variables. The coefficients remain highly significant. Column 4 includes a measure of bilateral trust between Germany and the host countries calculated based on the survey question in the Eurobarometer from the year 1996 (see Bloom et al., 2012; Guiso et al., 2009). Bloom et al. (2012) consider trust as a form of social capital and argue that higher levels of bilateral trust between a multinational firm's headquarters and the affiliate increase the firm's productivity by allowing the firm to decentralize more easily. The regression results confirm that higher bilateral trust increases firm productivity. The impact of the communication cost measures is unchanged, even though Guiso et al. (2009) find that trust is positively associated with common linguistic roots. Communication costs and trust have distinct effects on firm performance.

Columns 5 to 8 include statutory tax rates, the cost, time and number of procedures required to register property and to enforce contracts as measures for differences in the investment climate across countries. The first three columns include the measures separately, column 8 includes trust, tax rates and the time to enforce contracts and register property. We choose time to enforce contracts and register property as investment climate measures because these indicators seem most comparable across countries and least likely to pick up differences in the income per capita, which is used as reference to obtain comparable costs measures in the Doing Business Indicators. Our findings are robust throughout specifications, and the investment climate measures affect productivity in intuitive ways. Higher taxation negatively affects foreign productivity. The cost and time to enforce contracts does not have a significant effect, but the number of procedures is positively associated with productivity. This finding is contrary to expectations, but may reflect the fact that given the time and cost of enforcing contracts, a higher number of steps in judicial proceedings need not reflect red tape, but may be attributable to the demand for carefully finding a fair solution. This effect may be taken up by the coefficient of the time to enforce contracts in column 8. The number of procedures required to register property has a similar effect. When it takes longer to register property, a firm's productivity in a country is negatively affected. Column 9 includes geographic distance in the main specification, a common regressor in the gravity literature. The main regression results are robust. Relatively more distant affiliates are predicted to be relatively more productive.

	1	2	3	4	5	6	7	8
Log GDP		0.034***	0.064***	0.058***	0.053***	0.199***	0.057***	0.058***
		(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Log flight time	-0.043^{***}	-0.090^{***}	-0.038^{***}	-0.051^{***}	-0.057^{***}	-0.015	-0.041^{***}	-0.042^{***}
	(0.011)	(0.012)	(0.011)	(0.010)	(0.009)	(0.013)	(0.012)	(0.012)
Linguistic distance	-0.068^{*}			-0.143^{***}	-0.151^{***}	-0.286^{***}	-0.045	-0.051^{+}
	(0.030)			(0.032)	(0.032)	(0.035)	(0.031)	(0.031)
Time difference to Germany	0.020***	0.025^{***}	0.020***	0.014^{***}	0.016^{***}	0.041^{***}	0.019^{***}	0.019^{***}
	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
Public expenditure per pupil,	0.016***	0.015^{***}	0.015^{***}			0.037^{***}	0.017^{***}	0.017^{***}
% GDP per capita	(0.002)	(0.002)	(0.002)			(0.002)	(0.002)	(0.002)
GDP per employee	0.209***	0.234^{***}	0.202***	0.188^{***}	0.198^{***}		0.190^{***}	0.191^{***}
	(0.008)	(0.010)	(0.008)	(0.009)	(0.008)		(0.010)	(0.010)
Log market potential	0.056***							
	(0.008)							
Share speaking English		0.001*						
0		(0.000)	0 4 0 6 * * *					
Official language German			0.126***					
			(0.022)	0.000***				
PISA Maths score				0.002****				
				(0.000)	0.001***			
PISA Science score					0.001			
					(0.000)	0.047		
Labor productivity per employee, OECD						-0.047		
Log compondation non amplance						(0.002)		
Log compensation per employee						(0.051)		
Pule of low						(0.001)	0.047**	
Itule of law							(0.047)	
Covernment effectiveness							(0.017)	0.042*
Government enectiveness								(0.042)
Constant	/ 335***	4 304***	3 806***	3 759***	3 961***	9 635***	3 9/1***	3 023***
Constant	(0.096)	(0.153)	(0.126)	(0.163)	(0.166)	(0.156)	(0.128)	(0.130)
$-R^2$	0.194	0.169	0.196	0.171	0.170	0.124	0.199	0.199
Observations	117.519	103.170	117.557	42.079	42.079	91.600	100.164	100.164
# parents	7,732	7,433	7.735	6.753	6.753	6.867	7.255	7.255
# country combinations	9.042	8,811	9.047	4.764	4.764	8.628	8.084	8.084

Table 5: Regression results: within-firm differences in productivity across countries

Standard errors in parentheses. $^+ p < 0.10$, $^* p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$ Dependent variables: productivity - ln (foreign sales/foreign employees). Covariate definitions: see Table 2.

5.1.1 Robustness checks

Table 5 replicates the regression results when alternative variables are used as measures for the model parameters.

We use the market potential instead of GDP (column 1), and the share of the population speaking English (column 2) and an indicator for German as an official language of a country (column 3) as alternative measures for linguistic distance. The baseline results are robust and significant at least at the 5% level. The alternative measures have the expected effects and are significant.

The results are likewise robust when PISA maths and science scores are employed as an alternative measure for the knowledge acquisition costs (columns 4, 5). As predicted by the model, lower knowledge acquisition costs, here measured using the output of the education system, increase productivity. The sample size is considerably smaller as the measures are only available for a subset of years.

We use labor productivity measures from the OECD as an alternative for GDP per employee and reinclude wages as the sample is restricted to OECD countries anyways (column 6). Labor productivity is calculated as the gross value added divided by the number of employees. As in the main regressions, flight time is insignificant once the sample is restricted to OECD countries. The size of the coefficient of linguistic distance increases by almost 100%. Both labor productivity and wages are insignificant. This finding is surprising, given that productivity is robustly significant in all other specifications. This may result from a combination of factors: on the one hand, GDP per employee is a coarse measure and may pick up effects of income per capita differences. On the other hand, the labor productivity measure is only available for OECD countries, which are relatively homogeneous. Thus, there may just be too little variation to estimate a meaningful effect.

Finally, we use alternative measures for the quality of the investment climate from the World Governance Indicators (columns 7, 8). We include the rule of law, a measure that is also used in Nunn (2007) and government effectiveness. The baseline results are robust. Linguistic distance is significant at the 15% and 10% level respectively. As Kaufmann et al. (2010) point out, the World Governance Indicators are survey measures, so the precision of the individual measures should not be overestimated. To take this issue into consideration, we additionally run unreported regressions including the upper and lower bounds of the confidence interval for the measures instead of the point estimates. Results are robust.

5.1.2 Relevance

To assess the relevance of our findings, Table 6 presents the effect of an increase in the independent variables by one standard deviation on foreign labor productivity expressed in standard deviations. For comparison, it also displays the coefficient from the main regression (Table 4, column 2) and the coefficient range. The coefficient range is taken as being from the second lowest to the second largest coefficient, to remove outlying coefficients.

The labor productivity level in a country as measured by GDP per employee is the quantitatively most important influencing factor of productivity at the firm level. The impact of market size, communication costs and knowledge acquisition costs on productivity differences across countries is similar in magnitude. An increase in market size, or a decrease in communi-

			Coefficie	nt range
Variables	Main coefficient	Effect in SD	Lower	Upper
Log GDP	0.057	0.073	0.035	0.140
Log flight time	-0.044	-0.035	-0.105	-0.015
Linguistic distance	-0.071	-0.014	-0.286	-0.051
Time difference	0.020	0.060	0.014	0.042
Per pupil public expenditure	0.016	0.054	0.015	0.019
GDP per employee	0.207	0.275	0.144	0.234

Table 6: Effect of covariates in standard deviations of dependent variable

Table lists coefficient from Table 4, column 2, the effect of an increase in the variable value by one standard deviation on the dependent variable expressed in standard deviations and the coefficient range. The coefficient range is the second lowest to the second largest coefficient to remove outliers. All coefficients are significant.

cation or knowledge acquisition costs by one standard deviation leads to an increase in foreign productivity by around 0.05 standard deviations.

5.2 Cut-off productivity levels

The regression results for the cut-off productivity level are shown in Table 7. The minimum domestic productivity of investors per sector in a country and year is used as cut-off productivity measure in the upper panel, and the minimum global productivity of investors per sector, country and year is used in the lower panel. Consistent with Hypothesis 2, the domestic productivity of the least productive investor in a country increases in the communication costs between the host country and the home country as measured by log flight time and linguistic distance. These results confirm the predictions of the model. The generally high significance levels, usually 0.1%, are reassuring, especially given that the number of observations is considerably smaller than in the regressions at the firm level. The signs and significance levels are robust to the inclusion of measures for the fixed costs of investment. The effect of flight time is still positive, but turns insignificant if distance is included. The effect is significant at the 0.1% level once the costs of starting a business are controlled for. Larger market size as measured by GDP decreases the cut-off productivity level required for investment. The other covariates do not have robust effects. Tables B.3 and B.4 in Appendix B demonstrate that the regression results are robust to the inclusion of additional regressors and largely robust to using alternative measures for the covariates, similar to the robustness checks in the previous subsection.

Table 7: Regression results: cut-off productivity level

Domestic productivity	1	2	3	4	5	6
Log GDP	-0.161^{***}	-0.157^{***}	-0.168^{*}	-0.165^{***}	-0.163^{***}	-0.182^{**}
	(0.017)	(0.028)	(0.059)	(0.031)	(0.028)	(0.046)
Log flight time	0.155***	0.210***	0.321^{***}	0.214^{***}	0.032	0.442***
0 0	(0.034)	(0.056)	(0.058)	(0.061)	(0.114)	(0.088)
Linguistic distance	0.443***	0.404***	0.404***	0.389***	0.404***	0.485***
	(0.105)	(0.097)	(0.092)	(0.106)	(0.111)	(0.109)
Time difference to Germany	0.004	0.009	0.005	0.008	0.007	0.001
This difference to definally	(0.004	(0.000)	(0.000)	(0.008)	(0.008)	(0.001)
CDP nor omployee	(0.007)	(0.003)	(0.012)	0.015	(0.008)	(0.003)
GDF per employee		-0.003	(0.046)	-0.013	(0.010)	(0.051)
		(0.035)	(0.050)	(0.040)	(0.039)	(0.051)
Per pupil public expenditure,		0.007	0.031	0.006	0.002	0.035
% GDP p.c.		(0.011)	(0.014)	(0.012)	(0.011)	(0.014)
Log compensation per employee			0.027			0.027^{+}
			(0.025)			(0.018)
Log costs of starting a business				-0.027		0.066^{**}
				(0.030)		(0.022)
Log time to start a business				0.010		0.049^{+}
				(0.034)		(0.030)
Log distance				· /	0.138	-0.066
0					(0.071)	(0.053)
Constant	7.728***	6.410***	4.961***	6.574^{***}	6.438***	4.318***
Comptant	(0.224)	(0.558)	(0.907)	(0.637)	(0.540)	(0.734)
R-squared	0.308	0.325	0.304	0.330	0.327	0.394
Observations	11 460	8 280	4 708	8,000	8 280	4 597
Soster dummios	11,409 V	0,200 V	4,700 V	8,000 V	0,200 V	4, 527 V
Sector dummes		I V	I V	I V	I V	I V
rear dummes	Y	Ŷ	Ŷ	<u> </u>	Ŷ	<u>Y</u>
Global productivity	1	2	3	4	5	6
Log GDP	-0.113***	-0.151***	-0.144*	-0.149***	-0.155^{***}	-0.139*
	(0.016)	(0.031)	(0.054)	(0.035)	(0.031)	(0.053)
Log flight time	0.154***	0.233^{***}	0.294^{***}	0.216^{***}	0.098	0.547^{***}
	(0.036)	(0.053)	(0.058)	(0.055)	(0.154)	(0.125)
Linguistic distance	0.202^{+}	0.414^{***}	0.491^{***}	0.388^{**}	0.409^{**}	0.492^{***}
	(0.153)	(0.118)	(0.101)	(0.125)	(0.134)	(0.123)
Time difference to Germany	-0.001	0.003	0.023^{+}	0.005	0.002	0.025
	(0.007)	(0.009)	(0.016)	(0.009)	(0.009)	(0.013)
GDP per employee		0.083	0.136^{*}	0.078^{+}	0.096	$0.143^{'}$
r · · · · · · ·		(0.044)	(0.063)	(0.052)	(0.049)	(0.068)
Public expenditure per pupil		0.013^{+}	0.035*	0.006	0.010	0.034*
% GDP p.c.		(0.010)	(0.013)	(0.010)	(0.010)	(0.014)
Log compensation per empleyee		(0.010)	_0.004	(0.010)	(0.010)	_0.007
Log compensation per employee			-0.004			-0.007
T . C			(0.025)	0.000		(0.019)
Log costs of starting business				-0.028		0.021
				(0.027)		(0.025)
Log time to start business				0.023		0.070^{*}
				(0.033)		(0.032)
Log distance					0.106	-0.175^{*}
					(0.114)	(0.080)
Constant	2.347***	4.191***	3.166^{**}	4.538^{***}	4.172***	3.365^{***}
	(0.256)	(0.527)	(0.815)	(0.591)	(0.518)	(0.784)
R^2	0.316	0.356	0.460	0.361	0.357	0.466
Observations	14.476	10.309	5.512	9,939	10.309	5.291
Sector dummies	1 1, 1, 0	-0,000	0,012	0,000	-0,000	
		V	V	V	V	Y
Voor dummios		Y V	Y V	Y	Y V	Y V

Standard errors in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001. Dependent variable: cut-off productivity - min(ln (domestic sales/domestic employees)) by country and sector group. Covariate definitions: see Table 2.

6 Discussion

The results on within-parent across-country differences in firm performance and the cut-off productivity levels of foreign destinations obtained in the empirical analysis are consistent with the model predictions. A given investor exhibits superior performance in large markets characterized by relatively low bilateral communication costs between the home country and the host country, comparatively low knowledge acquisition costs, and relatively high labor productivity. The cut-off productivity levels decrease in market size and increase in bilateral communication costs. This evidence is indirect: the data do not contain information on the allocation of knowledge across countries which are the focus of the model mechanism. It is therefore a key concern with respect to the empirical results whether they can really be attributed to the organization of knowledge in multinational firms.

Two recent papers study the relation between the geographical distribution of establishments and firm performance using U.S. data on national multi-establishment firms. Giroud (2013) finds that investment in a plant increases after a new airline route between the firm's headquarter and the plant location is introduced. Kalnins and Lafontaine (2013) demonstrate that greater distance of the establishment from the headquarters is associated with shorter establishment survival. Both articles attribute their findings to monitoring problems and information asymmetries between firm headquarters and establishments.

This paper proposes an alternative explanation to rationalize these empirical findings: new airline routes and lower distance decrease communication costs between headquarters and plants, which renders plants more productive and increases firms' investment incentives. At the same time, one could argue in the vein of Giroud (2013) that the empirical results in this paper are driven by monitoring problems within multinational firms. The effect of flight time can be rationalized using a model that features monitoring problems between production workers and headquarter managers, variation in cross-border monitoring costs, and heterogeneity in firms' monitoring technology, for example along the lines of Qian (1994). If monitoring costs are higher, a firm has to pay higher wages to implement the optimal effort level, so marginal costs of production are higher. Thus, such a model would likewise generate lower within-firm productivity in countries with higher cross-border monitoring costs. (A formal analysis of such a model is sketched in Appendix D.)

Neither of the two mechanisms is susceptible to a direct test with the data available. While we do not deny that monitoring is an important factor for efficient production, we are convinced of the empirical relevance of the mechanism based on the organization of knowledge proposed in this paper. Higher bilateral trust is likely to decrease monitoring costs. We find that communication costs are relevant even if bilateral trust is controlled for. In addition, we approximate communication costs not only by flight time, but also using linguistic distance. At least with respect to routine tasks, it is difficult to claim that monitoring problems are less easily mitigated if the linguistic distance between the home country and the host country is larger. A supervisor's assessment of a worker's performance depends on observing what the worker does, not what the worker claims to do.

Finally, our knowledge-based mechanism is consistent with empirical evidence on sectoral differences in the geographical concentration of foreign direct investment (Bahar, 2013, c.f. section 2.3.2) and explains *home* country labor market effects of multinational activity. Only firms with sufficiently high knowledge level $\bar{z}_i > \bar{z}^I$ select into foreign direct investment, and they tend

to pay higher wages than comparable domestic firms in the home country (c.f. Proposition 4). If monitoring were the only driver behind the empirical patterns presented in the previous section, the multinational firms should pay *lower* wages in the home country than comparable domestic firms: to overcome higher cross-border monitoring costs, firms would have to dispose of a better monitoring technology. Firms with better monitoring technology are able to implement optimal effort levels with lower wage payments. Multinational parents are therefore unambigously predicted to pay lower wages than their domestic counterparts in the home country according to a monitoring based model, which is at odds with the empirical evidence.

7 Conclusion

This paper studies the organization of knowledge in multinational firms. Multinational firms are predicted to optimally allocate more knowledge to foreign countries characterized by higher bilateral communication costs with the home country, higher labor productivity, lower wages and lower knowledge acquisition costs. The distribution of productivities across countries within German multinational firms is shown to be consistent with these predictions, as is the pattern of cut-off productivity levels. In addition, both the home and host country multinational wage premiums generated by the model are in line with the available empirical evidence.

The paper offers relevant insights for the design of investment promotion policies. Creating well-paid, relatively knowledge intensive new jobs is one of the main targets of investment promotion efforts (Javorcik, 2012). The results of this paper generally support the presumption that employment in multinational affiliates is likely to be more knowledge intensive and better paid than employment in domestic firms. In their efforts to reap these benefits, countries tend to focus on investing in targeted information campaigns and a good investment climate in terms of administration, governance and the education of their workforce. As this paper demonstrates, targeted foreign language training and good communication infrastructures may be equally relevant to foster FDI inflows, as they facilitate multinational enterprises' task of efficiently organizing across countries.

References

- Aitken, B., A. Harrison, and R. E. Lipsey (1996). Wages and foreign ownership A comparative study of Mexico, Venezuela, and the United States. *Journal of international Economics* 40(3), 345–371.
- Ambos, T. C. and B. Ambos (2009). The impact of distance on knowledge transfer effectiveness in multinational corporations. *Journal of International Management* 15(1), 1–14.
- Antras, P. (2003). Firms, contracts, and trade structure. The Quarterly Journal of Economics 118(4), 1375–1418.
- Antras, P. and E. Helpman (2004). Global Sourcing. *The Journal of Political Economy* 112(3), 552–580.
- Bahar, D. (2013). Heavier than Air? Knowledge Transmission within the Multinational Firm. mimeo, Harvard University.
- Bloom, N., L. Garicano, R. Sadun, and J. Van Reenen (2009). The distinct effects of information technology and communication technology on firm organization. NBER Working Paper No. 14975.
- Bloom, N., R. Sadun, and J. Van Reenen (2012). The organization of firms across countries^{*}. *The quarterly journal of economics* 127(4), 1663–1705.
- Bolton, P. and M. Dewatripont (1994). The firm as a communication network. *The Quarterly Journal of Economics* 109(4), 809–839.
- Caliendo, L., F. Monte, and E. Rossi-Hansberg (2012). The anatomy of French production hierarchies. *NBER Working Paper No. 18259.*
- Caliendo, L. and E. Rossi-Hansberg (2012). The Impact of Trade on Organization and Productivity. The Quarterly Journal of Economics 127(3), 1393–1467.
- Carr, D. L., J. R. Markusen, and K. E. Maskus (2001). Estimating the knowledge-capital model of the multinational enterprise. *The American Economic Review* 91(3), 693–708.
- Chen, M. X. and M. O. Moore (2010). Location decision of heterogeneous multinational firms. Journal of International Economics 80(2), 188–199.
- Crystal, D. (2003). English as a global language. Ernst Klett Sprachen.
- De Long, D. W. and L. Fahey (2000). Diagnosing cultural barriers to knowledge management. The Academy of Management Executive 14(4), 113–127.
- Defever, F. (2012). The spatial organization of multinational firms. Canadian Journal of Economics/Revue canadienne d'économique 45(2), 672–697.
- Egger, H. and U. Kreickemeier (2013). Why Foreign Ownership May Be Good For You. *Inter*national Economic Review 54(2), 693–716.
- Ekholm, K., R. Forslid, and J. R. Markusen (2007). Export-platform Foreign Direct Investment. Journal of the European Economic Association 5(4), 776–795.

- Foss, N. J. and T. Pedersen (2002). Transferring knowledge in mncs: The role of sources of subsidiary knowledge and organizational context. *Journal of International Management* 8(1), 49–67.
- Gardner, J. (2013). Misused English words and expressions in EU publications. European Court of Auditors Secretariat General Translation Directorate.
- Garicano, L. (2000). Hierarchies and the Organization of Knowledge in Production. *The Journal* of *Political Economy* 108(5), 874–904.
- Garicano, L. and T. N. Hubbard (2009). Specialization, firms, and markets: The division of labor within and between law firms. *Journal of Law, Economics and Organization* 25(2), 339–371.
- Ghemawat, P. (2007). Managing Differences: The Central Challenge of Global Strategy. *Har*vard Business Review 85(5), 59–68.
- Girma, S. and H. Görg (2007). Evaluating the foreign ownership wage premium using a difference-in-differences matching approach. *Journal of International Economics* 72(1), 97–112.
- Giroud, X. (2013). Proximity and investment: Evidence from plant-level data. *The Quarterly Journal of Economics* 128(2), 861–915.
- Grossman, G. M., E. Helpman, and A. Szeidl (2006). Optimal integration strategies for the multinational firm. *Journal of International Economics* 70(1), 216–238.
- Guiso, L., P. Sapienza, and L. Zingales (2009). Cultural biases in economic exchange? The Quarterly Journal of Economics 124(3), 1095–1131.
- Hansen, M. T. and B. Løvås (2004). How do multinational companies leverage technological competencies? moving from single to interdependent explanations. *Strategic Management Journal* 25(8-9), 801–822.
- Helpman, E. (1984). A simple theory of international trade with multinational corporations. The Journal of Political Economy 92(3), 451–471.
- Helpman, E., M. J. Melitz, and S. R. Yeaple (2004). Export Versus FDI with Heterogeneous Firms. *The American Economic Review* 94(1), 300–316.
- Heyman, F., F. Sjöholm, and P. G. Tingvall (2007). Is there really a foreign ownership wage premium? Evidence from matched employer–employee data. *Journal of International Eco*nomics 73(2), 355–376.
- Hijzen, A., P. S. Martins, T. Schank, and R. Upward (2013). Foreign-owned firms around the world: A comparative analysis of wages and employment at the micro-level. *European Economic Review 60*, 170–188.
- Irarrazabal, A., A. Moxnes, and L. D. Opromolla (2013). The margins of multinational production and the role of intrafirm trade. *The Journal of Political Economy* 121(1), 74–126.

- Javorcik, B. (2012). Does FDI Bring Good Jobs to Host Countries? Background Paper for the World Development Report 2013, World Bank.
- Kalnins, A. and F. Lafontaine (2013). Too Far Away? The Effect of Distance to Headquarters on Business Establishment Performance. American Economic Journal: Microeconomics, forthcoming.
- Kaufmann, D., A. Kraay, and M. Mastruzzi (2010). The Worldwide Governance Indicators: Methodology and Analytical Issues. *Policy Research Working Paper 5430*, World Bank.
- Keller, W. (2004). International technology diffusion. *Journal of Economic Literature* 42, 752–782.
- Keller, W. and S. R. Yeaple (2013). The gravity of knowledge. American Economic Review 103(4), 1414–44.
- Lagerström, K. and M. Andersson (2003). Creating and sharing knowledge within a transnational teamthe development of a global business system. *Journal of World Business* 38(2), 84–95.
- Lipponer, A. (2009). Microdatabase Direct Investment (MiDi) A Brief Guide. Technical report, Deutsche Bundesbank.
- Lipsey, R. E. (2004). Home-and host-country effects of foreign direct investment. In *Challenges to globalization: Analyzing the economics*, pp. 333–382. University of Chicago Press.
- Malchow-Møller, N., J. R. Markusen, and B. Schjerning (2013). Foreign Firms, Domestic Wages. The Scandinavian Journal of Economics 115(2), 292–325.
- Markusen, J. R. (1984). Multinationals, multi-plant economies, and the gains from trade. *Journal of International Economics* 16(3), 205–226.
- Markusen, J. R. and K. E. Maskus (2001). Multinational firms: reconciling theory and evidence. In Topics in empirical international economics: A festschrift in honor of Robert E. Lipsey, pp. 71–98. University of Chicago Press.
- Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71(6), 1695–1725.
- Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *The Quarterly Journal of Economics* 122(2), 569–600.
- Oostendorp, R. H. (2013). The Occupational Wages around the World (OWW) Database: Update for 1983- 2008. Background Paper for the World Development Report 2013, World Bank.
- Qian, Y. (1994). Incentives and loss of control in an optimal hierarchy. The Review of Economic Studies 61(3), 527–544.
- Radner, R. (1993). The organization of decentralized information processing. *Economet*rica 61(5), 1109–1146.

- Ramondo, N., V. Rappoport, and K. J. Ruhl (2013). Horizontal vs. vertical FDI: Revisiting evidence from US multinationals. *mimeo*, University of California at San Diego.
- Spolaore, E. and R. Wacziarg (2009). The diffusion of development. The Quarterly Journal of Economics 124 (2), 469–529.
- Tintelnot, F. (2012). Global production with export platforms. *mimeo*, Pennsylvania State University.
- UNCTAD (2004). World Investment Report 2004. The Shift Towards Services.
- Yeaple, S. R. (2003). The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment. *Journal of International Eco*nomics 60(2), 293–314.
- Yeaple, S. R. (2013). Scale, Scope, and the International Expansion Strategies of Multiproduct Firms. *mimeo*, Pennsylvania State University and NBER.

Appendix

To be completed.