

Margins of Multinational Labor Substitution*

Marc-Andreas Muendler[¶]

UC San Diego and CESifo

Sascha O. Becker

U Munich, CESifo and IZA

February 16, 2006

Abstract

Multinational labor demand responds to wage differentials at the extensive margin, when a multinational enterprise (MNE) expands into foreign locations, and at the intensive margin, when an MNE operates existing affiliates across locations. We derive conditions for parametric and nonparametric identification of an MNE model to infer elasticities of labor substitution at both margins, controlling for location selectivity. Prior studies have rarely found foreign wages or operations to affect employment. Our strategy detects salient adjustments at the extensive margin for German MNEs. With every percentage increase in German wages, German MNEs allocate 1,600 manufacturing jobs to Eastern Europe at the extensive margin and 3,900 jobs overall.

Keywords: Multinational enterprise; location choice; sample selectivity; labor demand; translog cost function; nonparametric estimation

JEL Classification: F21, F23, C14, C24, J23

*We thank Xiaohong Chen, Peter Egger, Gordon Hanson, Sebastian Kessing and Hal White as well as participants at various seminars and conferences for insightful suggestions. We thank Steve Redding for sharing code to compute market access statistics. Jennifer Poole, Robert Jäckle, Nadine Gröpl, and Daniel Klein provided excellent research assistance. Simone Hofer from UBS kindly shared the bank's international wage data. We gratefully acknowledge financial support from the VolkswagenStiftung under its grant initiative *Global Structures and Their Governance* and administrative and financial support from the Ifo Institute. Becker also gratefully acknowledges financial support from the Fritz-Thyssen-Stiftung.

[¶]muendler@ucsd.edu (*econ.ucsd.edu/muendler*), corresponding. Ph: +1 (858) 534-4799

1 Introduction

The expansion and operation of multinational enterprises (MNEs) is a driving force of international economic integration. The world's ten largest MNEs in 2000 produce almost one percent of world GDP, and the one hundred largest MNEs are responsible for more than four percent of world GDP (up from three-and-a-half percent in 1990).¹ MNEs with headquarters in the U.S., for instance, transact more than two in five exports and more than half of their imports through their affiliates (Slaughter 2000).

Surprisingly, however, the operation of MNEs has rarely been found to affect factor demands, although MNEs are important mediators of world trade, and demands for domestic manufactures and labor are strongly affected by trade. Feenstra and Hanson (1999) argue that fifteen or more percent of the relative U.S. wage changes over the past decades can be explained with outsourcing (within MNEs or across enterprises). Several studies on the operation of MNEs have concluded, however, that MNE production in low-wage locations has no detectable impact on their labor demand in the home market (e.g. Slaughter (2000) for U.S. MNEs). Studies on the expansion of MNEs have not found low wages or low per-capita incomes to be significant predictors of location choice (e.g. Devereux and Griffith (1998) for U.S., Head and Mayer (2004) for Japanese MNEs). Hanson, Mataloni, and Slaughter (2005), however, report that affiliates of U.S. MNEs process more intra-firm imports the lower are low-skilled wages. The result challenges the view that relative abundance in low-skilled labor fails to attract MNEs.

Like several earlier studies, we find labor costs to have little measurable impact on MNEs' observed choices of foreign expansions. An MNE's propensity to select a foreign location, however, is a weighty predictor of its labor demand so that even poor effects of wages on location choice can have a strong impact on multinational labor substitution. Findings in this paper for a representative sample of German manufacturing MNEs and their majority-owned foreign manufacturing affiliates show that a weak effect of wage differences on location selection, weighted with the salient effect of location selection on labor demand, results in a statistically significant and economically important wage effect on MNEs' labor demands. There are large sunk entry and exit costs so that expansions (or withdrawals) are infrequent but, when undertaken, they have a sizeable impact on labor demand.

We quantify the employment effect of permanent wage differentials at two important margins. An MNE's labor demand responds to international wage differentials at the *extensive margin*, when the MNE expands into a foreign market, and at the *intensive margin*, when the MNE operates existing affiliates and chooses employment. Our paper thus offers an integration of two strands of the empirical literature—one on MNEs' location choices (Devereux and Griffith 1998, Head and Mayer 2004) and one on MNE operations across existing locations (Slaughter 2000, Head and

¹UNCTAD press release TAD/INF/PR/47 (12/08/02).

Ries 2002, Hanson, Mataloni, and Slaughter 2005)—into a unified estimation framework. The MNE’s two-stage decision, to first expand (extensive margin) and then operate (intensive margin), has a well-defined econometric counterpart in sample selection. Aside from the economic interpretation of the extensive margin, labor demand or cost function estimates at the intensive margin are subject to selectivity bias unless corrected.

A methodological contribution of our paper is to extend the univariate sample selection case to one of multiple selections. We derive conditions under which the common Heckman (1979) selection correction can be applied location by location to a seemingly unrelated equation system of the MNE’s cost function. We also prove identification of a nonparametric selection model, which extends single-equation models (such as those in Das, Newey, and Vella (2003)) to the multivariate case. The nonparametric estimator is simple to implement in a two-stage approach and is applicable to the estimation of multivariate demand systems in general (for a recent parametric approach to multivariate demand, see e.g. Yen (2005)).

Similar to Slaughter (2000) for the operation of U.S. MNEs, Konings (2004) and Barba Navaretti and Castellani (2004) find no evidence for the hypothesis that the production of European MNEs in low-wage locations has a detectable impact on home-market labor demand. Braconier and Ekholm (2000) and Marin (forthcoming) respectively estimate wage elasticities of labor demand and intermediate imports from Central and Eastern Europe, and report no significant effect of foreign relative wages. Similarly, studies have not found relative abundance of low-skilled labor to explain sales patterns across locations (e.g. Brainard (1997) for aggregate U.S. MNE trade).

Neither have studies on the expansion of MNEs found low wages or low per-capita incomes to be significant predictors of location choices (e.g. Devereux and Griffith (1998) for U.S., Head and Mayer (2004) for Japanese, Buch, Kleinert, Lipponer, and Toubal (2005) for German MNEs). Multinomial logit estimation does turn wages into significant predictors of location selection in Disdier and Mayer (2004) for French MNEs, and in our exercise for Swedish MNEs and the same sample of German MNEs (Becker, Ekholm, Jäckle, and Muendler 2005). Multinomial logit estimation rests on the assumption, however, that independent decision makers within the MNE decide on distinct investment projects. This conceptual framework is incompatible with a model of MNE-wide profit maximization as in this paper. Devereux and Griffith (1998) estimate multinomial logit choice but, to be consistent, restrict their sample to MNEs who invest in only one location abroad, and do not find wages to be significant predictors of U.S. MNEs location choices in Europe.

Contrary to results from firm-level studies, Carr, Markusen, and Maskus (2001) find evidence in aggregate data that relatively abundant high-skilled labor is a significant predictor of foreign direct investment (FDI) of U.S. MNEs (and Blonigen, Davies, and Head (2003) find that low labor costs predict less MNE activity). Among

the firm-level studies, Brainard and Riker (2001) find that foreign affiliate employment substitutes modestly for U.S. parent employment but less so than for employment across foreign locations. Hanson, Mataloni, and Slaughter (2005) use firm-level data and, as an exception to much prior firm-level evidence, report that affiliates of U.S. MNEs process significantly more intra-firm imports the lower are low-skilled wages. The result challenges the view that relative abundance in low-skilled labor fails to attract MNEs. We revisit their result in the context of multinational labor substitution and extend the estimation framework to incorporate location choice.

We estimate elasticities of multinational labor substitution from a translog cost function, controlling for location selectivity. We find cross-wage elasticities at the extensive margin to be strictly positive. So, home and foreign employment are substitutes within MNEs not only at the intensive but also at the extensive margin. Elasticities at the extensive margin are smaller in magnitude than at the intensive margin especially in locations close to home. Bootstraps, however, reject equality between the intensive and the total elasticity of substitution (in all locations except for overseas developing countries), corroborating the importance of the extensive margin. Elasticity point estimates at both margins are robust across different samples, specifications, and parametric and nonparametric estimation techniques.

We evaluate the counterfactual question how many jobs MNEs would reallocate in response to shrinking wage differentials. A one-percent drop in German wages relative to the sample-mean level would reduce employment in Central and Eastern Europe (CEE) by around 3,900 jobs, for instance. Similarly, a one-percent increase in CEE wages would bring 650 jobs to Germany. These are sizeable figures. Wages in CEE are, on average, 12.8 percent of the German level in 2000. If the estimated elasticities of substitution were constant at all levels of wages, an increase in CEE wages of 780% ($=1/.128$) to remove the wage gap to Germany would bring 500,000 counterfactual manufacturing jobs to Germany—more than a third of the estimated home employment at German manufacturing MNEs. Of course, elasticities of substitution are not constant at all levels of wages so that the counterfactual prediction is crude. We nevertheless view the magnitude as indicative of the potential importance of multinational labor substitution.

This paper has five more sections. Section 2 presents a model of the expansion and operation of MNEs, and Section 3 derives identification conditions for its estimation under location selectivity. Section 4 presents the data and discusses descriptive statistics on location choice. Estimation results on multinational labor substitution are presented in Section 5, and interpreted in counterfactual evaluations. Section 6 concludes. Details on the MNE model and data construction are relegated to an Appendix.

2 Multinational Expansion and Operation

Let observed employment y_j^ℓ of MNE j at time t in location ℓ obey

$$y_{jt}^\ell = \mathbf{x}_{jt}^\ell \beta^\ell + \epsilon_{jt}^\ell$$

if MNE j is present at ℓ . Else, $y_{jt}^\ell = 0$. The vector \mathbf{x}_{jt}^ℓ of employment predictors includes output and factor prices (or additively separable transformations of output and factor prices such as in the translog case), and ϵ_{jt}^ℓ is a disturbance term. So, the conditional expectation of MNE j 's observed employment in location ℓ is

$$\bar{y}_{jt}^\ell \equiv \mathbb{E} [y_{jt}^\ell \mid \mathbf{x}_{jt}^\ell, \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}] = \mathbf{x}_{jt}^\ell \beta^\ell + \mathbb{E} [\epsilon_{jt}^\ell \mid \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}], \quad (1)$$

where the vector \mathbf{d}_{jt} of presence indicators d_{jt}^k reflects MNE j 's observed pattern of locations $k = 1, \dots, L$ at time t ($d_{jt}^k = 1$ if firm j is present in location k and $d_{jt}^k = 0$ otherwise) and contains $d_{jt}^\ell = 1$. The information set $\mathbf{z}_{j,t-\tau}$ at moment $t - \tau$ affects labor demand through the resulting choice of presence in location ℓ .

We define the *extensive* margin of labor demand to be the expected labor demand $\bar{y}_\ell^{\text{ext}}$ in location ℓ , predicted by a firm j 's current choices of presence around the world and its past information set $\mathbf{z}_{j,t-\tau}$

$$\bar{y}_{jt}^{\text{ext},\ell} \equiv \mathbb{E} [\epsilon_{jt}^\ell \mid d_{jt}^1, \dots, d_{jt}^\ell = 1, \dots, d_{jt}^L; \mathbf{z}_{j,t-\tau}], \quad (2)$$

where the optimal binary choices $(d_{jt}^1, \dots, d_{jt}^\ell, \dots, d_{jt}^L)$ are functions of MNE j 's information set at the moment of location choice $t - \tau$, and τ is the time it takes an MNE to implement location choices (two to four years, say). The information set $\mathbf{z}_{j,t-\tau}$ at moment $t - \tau$ predicts presence in location k with $d_{jt}^k = \mathbf{1}(H(\mathbf{z}_{j,t-\tau}) + \eta_{jt}^k > 0)$, where $H(\cdot)$ is an unknown function and η_{jt}^k is a disturbance that may affect an MNE's presence until moment t . Most important, $\mathbf{z}_{j,t-\tau}$ includes the then prevailing wage differentials between locations.

Labor demand at the *intensive* margin is then defined as

$$\bar{y}_{jt}^{\text{int},\ell} \equiv \bar{y}_{jt}^\ell - \bar{y}_{jt}^{\text{ext},\ell} = \mathbf{x}_{jt}^\ell \beta^\ell. \quad (3)$$

The labor demand effect at the extensive margin $\bar{y}_{jt}^{\text{ext},\ell} = \mathbb{E}[\epsilon_{jt}^\ell \mid \mathbf{d}_{jt}]$ is an additive component of conditional labor demand $\mathbb{E}[y_{jt}^\ell \mid \mathbf{x}_{jt}^\ell, \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}]$. Economically, an MNE's mere presence at a location typically raises the labor demand prediction for that location.² Statistically, the extensive margin needs to be included in the regression to correct for selectivity.

²To be precise, this is true if high home wages raise the probability of presence at a foreign location ℓ and the presence likelihood is positively correlated with labor demand at that foreign location ℓ . Both conditions are satisfied in our MNE sample.

Labor demand predictors \mathbf{x}_{jt}^ℓ at the intensive margin include the prevailing wage differentials between locations at time t .

MNE j produces a vector of location-specific outputs $\mathbf{q}_{jt} = (q_{jt}^1, \dots, q_{jt}^L)'$ at L locations. We consider MNEs to be price takers in input market, whereas they may have market power in output markets. (We estimate a cost function, so any pricing behavior in the sales market is consistent with our approach.) On the input side, we focus on employment. We view MNEs as wage takers in the local markets, competing with labor demand from non-tradeable goods sectors and foreign incumbent firms. Similarly, we consider demand for capital goods and intermediate inputs from non-MNEs as sufficiently large so that the remaining demand of MNEs for those goods has a negligible price impact.

Final goods prices are world-market prices that differentiated products *from* locations $\ell = 1, \dots, L$ can fetch, given product characteristics. Final goods are produced with labor and capital. After controlling for location choice in the formation of the MNE, we consider installed capital $\mathbf{k}_{jt} = (k_{jt}^1, \dots, k_{jt}^L)'$ to be a quasi-fixed factor in an MNE's short-run cost function C_{jt} (but put to use at locations $k = 1, \dots, L$ to different degrees). We consider labor at locations $k = 1, \dots, L$ to be immobile across national borders and its factor prices $\mathbf{w}_t = (w_t^1, \dots, w_t^L)'$ as specific to L locations.

2.1 Location choice

Define γ_N^ℓ as the fixed entry costs for establishing FDI presence at location ℓ and γ_X^ℓ as the fixed FDI exit costs from location ℓ .³ Then, gross fixed costs of presence or absence from location ℓ in t , anticipated at $t - \tau$, become

$$G_{j,t-\tau}^\ell(d_{jt}^\ell, d_{j,t-\tau}^\ell) = \gamma_N^\ell d_{jt}^\ell (1 - d_{j,t-\tau}^\ell) + \gamma_X^\ell (1 - d_{jt}^\ell) d_{j,t-\tau}^\ell,$$

where d_{jt}^ℓ is the indicator for MNE j 's current FDI presence at location ℓ , and $d_{j,t-\tau}^\ell$ for its past presence. Hence, the decision-relevant fixed cost difference $F_{j,t-\tau}^\ell \equiv G_{j,t-\tau}^\ell(1, d_{j,t-\tau}^\ell) - G_{j,t-\tau}^\ell(0, d_{j,t-\tau}^\ell)$ between presence at location ℓ and absence from ℓ at time t is

$$F_{j,t-\tau}^\ell = \gamma_N^\ell - (\gamma_X^\ell + \gamma_N^\ell) d_{j,t-\tau}^\ell, \quad (4)$$

where $(\gamma_X^\ell + \gamma_N^\ell)$ is sometimes called the *hysteresis band* and reflects the sunk cost effect that induces firms to continue operations in location ℓ (Dixit 1989).⁴

³For lack of empirical identification, the fixed costs of reentry into a given location after a period of absence are assumed to be equal to the presence costs at first entry γ_N^ℓ .

⁴Probit estimation with firm-fixed effects is known for problematic performance in panel data with a short time horizon (Heckman 1981). We therefore do not attempt to estimate MNE-specific sunk costs of presence $F_{j,t-\tau}^\ell$ at location ℓ . We can distinguish between entry and exit sunk cost components to account for MNE-specific differences in $F_{j,t-\tau}^\ell$ similar to Roberts and Tybout's (1997) model of sunk costs in exporting status.

To select locations (τ years prior to production and sales), MNE j considers maximal expected profits $\mathbb{E}_{j,t-\tau}[\mathbf{p}(\mathbf{q}_{i \neq j,t}, \mathbf{q}_{jt})' \cdot \mathbf{q}_{jt} - C_{jt}(\mathbf{q}_{jt}; \mathbf{k}_{jt}, \mathbf{w})]$. A fully specified model of multinational location choice in Appendix A shows that MNE j 's profit-maximizing rule for entry with FDI in location ℓ can be written as

$$\begin{aligned} d_{jt}^\ell &= \mathbf{1}\left(\mathbb{E}_{j,t-\tau}[p^\ell q_{jt}^{\ell,*}] + \mathbb{E}_{j,t-\tau}[C_{jt}(q_{jt}^\ell=0; \cdot) - C_{jt}(q_{jt}^{\ell,*}; \cdot)] - F_{j,t-\tau}^\ell + \eta_{jt}^\ell > 0\right) \\ &= \mathbf{1}\left(h(\mathbf{z}_{j,t-\tau}^0) - \gamma_N^\ell + (\gamma_X^\ell + \gamma_N^\ell) d_{j,t-\tau}^\ell + \eta_{jt}^\ell > 0\right) \\ &= \mathbf{1}\left(H(\mathbf{z}_{j,t-\tau}) + \eta_{jt}^\ell > 0\right), \end{aligned} \tag{5}$$

where the unknown function $h(\mathbf{z}_{j,t-\tau}^0)$ captures both revenues from producing the expected profit-maximizing quantity $q_{jt}^{\ell,*}$ at location ℓ and cost savings from producing at ℓ (see first line). Sunk costs of presence at location ℓ have an observable component $F_{j,t-\tau}^\ell$ by (4) and a disturbance η_{jt}^ℓ . This stochastic component η_{jt}^ℓ is known to the MNE but not to the researcher. To simplify notation, we write $H(\mathbf{z}_{j,t-\tau}) \equiv h(\mathbf{z}_{j,t-\tau}^0) - \gamma_N^\ell + (\gamma_X^\ell + \gamma_N^\ell) d_{j,t-\tau}^\ell$ and include past presence in any location in the information set $\mathbf{z}_{j,t-\tau}$.

Equation (5) is the selection equation: the empirical rule of presence in locations $\ell = 1, \dots, L$. We estimate the rule both parametrically (with a probit regression and $h(\mathbf{a}) = \mathbf{a}\gamma^\ell$) and nonparametrically.

2.2 Multiproduct cost function

To obtain theoretically well-defined estimates of elasticities of labor substitution across locations, we opt for a flexible parametric specification of the MNE's multiproduct cost function. We first augment the cost function with parametric corrections for location selectivity. We then proceed to a model with a parametric cost function part and a nonparametric correction for selectivity.

We use a short-run multiproduct translog cost function to estimate labor demand, and extend it to control for location selectivity.⁵ A short-run cost function, given MNE j 's location choice, treats MNE j 's vector of capital stocks \mathbf{k}_{jt} as quasi-fixed factors. We prefer a short-run over a long-run cost function because we already control for the installation of foreign affiliates through location selectivity (5) and because the inclusion of capital stock variables captures otherwise unobservable (firm-specific) user costs of capital across locations. Note that a translog cost function is not separable in individual products so that $C_{jt}(\mathbf{q}_{jt}; \mathbf{k}_{jt}, \mathbf{w}) \neq \sum_\ell c_{jt}^\ell(q_{jt}^\ell; \mathbf{w}_\ell)$

⁵We follow Brown and Christensen's (1981, eq. 10.21) short-run version of Christensen, Jorgenson, and Lau (1973) and extend the framework to multiple products. A main alternative would be Hall's (1973) generalization of Diewert's (1971) Leontief cost function to the multiproduct case. We favor the translog cost function because its dimensionality requirements are considerably leaner and permit higher-order approximations to the nonparametric correction for selectivity. The translog specification is common in the empirical trade literature since Kohli (1978).

for product-level cost functions $c_{jt}^\ell(\cdot)$. Otherwise, elasticities of substitution would be zero.

Applying Shepard's (1953) lemma to the short-run multiproduct translog cost function yields location-specific cost shares $s_{jt}^\ell \equiv w_t^\ell y_{jt}^\ell / C_{jt}$ (the wage bill at location ℓ in the MNE's total wage bill) as functions of $(\mathbf{q}_{jt}; \mathbf{k}_{jt}, \mathbf{w})$. We multiply the cost shares s_{jt}^ℓ with observation-specific scalars C_{jt}/w_t^ℓ to arrive at our outcome equation (labor demand at ℓ)

$$y_{jt}^\ell = \mathbf{x}_{jt}^\ell \beta^\ell + \epsilon_{jt}^\ell \quad (6)$$

with

$$\mathbf{x}_{jt}^\ell \beta^\ell = \alpha_\ell \frac{C_{jt}}{w_t^\ell} + \sum_{m=1}^L \left(\mu_{\ell m} \ln(q_{jt}^m)^{C_{jt}/w_t^\ell} + \kappa_{\ell m} \ln(k_{jt}^m)^{C_{jt}/w_t^\ell} + \delta_{\ell m} \ln(w_t^m)^{C_{jt}/w_t^\ell} \right)$$

(see Appendix B), where ϵ_{jt}^ℓ is a disturbance.

Compared to translog regression equations in wage bill shares s_{jt}^ℓ , the transformation with observation-specific scalars C_{jt}/w_t^ℓ to an equivalent regression of y_{jt}^ℓ on \mathbf{x}_{jt}^ℓ has three important advantages. First, there is no constant term among the regressors \mathbf{x}_{jt}^ℓ so that lacking identification of the constant in a nonparametric selection correction is no concern. Second, wages are regressors only and do not enter the dependent variable. Third, labor demand is not bounded above so that, conditional on \mathbf{x}_{jt}^ℓ , the labor demand disturbance satisfies the assumption of a one-sided truncation for (Heckman (1979) and nonparametric) selectivity correction.

Stacking locations with zero output and factor use. Most MNEs produce in some but not in all locations. For cases of zero output or input, however, equation (6) is not well defined. Especially zero turnover and zero capital stocks require attention because they are MNE-specific, but absence from a location also suggests dropping wage regressors when no employment occurs.

One possible treatment is estimation of separate equation systems for every single presence pattern in the data. The estimators are hard to interpret, however, and plagued by dimensionality: potential presence in up to $L - 1$ locations outside the home location implies that there are up to $2^{L-1} - 1$ regional presence patterns for an MNE.⁶ In the German sample in 2000, for instance, only 57 out of 1,770 MNEs are omnipresent in all four world locations while every single one of the 15 possible regional presence pattern occurs.

We choose to stack the observations of all MNEs in the sample. In general, stacking observations improves efficiency, collapses the up to $2^{L-1} - 1$ sets of estimates into one consistently estimated $(L-1)$ -equation system, and provides a single

⁶MNEs are present in their home location by sample definition, so only 2^{L-1} patterns are observable in principle. Firms that only operate domestically without any foreign affiliate are not MNEs by definition so that the single presence pattern with the only presence at the home location must be subtracted.

$L \times L$ matrix of estimates for wage elasticities of regional labor demands. Stacking is permissible under three conditions: (i) all MNEs face identical sunk cost $F_{j,t-\tau}^\ell$ for presence at location ℓ conditional on their prior presence and information set (so that presence is not correlated with inputs); (ii) MNEs face an identical short-run cost function $C_{jt}(\cdot) = C(\cdot)$ in all locations of presence, conditional on their characteristics (so that one common parameter vector is justified); and (iii) the disturbances ϵ_{jt}^ℓ are uncorrelated across observations.

There are at least two methods of stacking in our context. Because production theory offers little guidance, we pursue both methods and settle with the more conservative estimates for the elasticities of substitution. First, set all missing location variables for an absent MNE j to zero—that is log employment, turnover, capital stock and wages are zero at location m from where MNE j is absent. This is equivalent to interacting the translog cost function coefficients with presence indicators: $\mu_{\ell m} = 0$ when no output is produced at location m , and $\kappa_{\ell m} = \delta_{\ell m} = 0$ when MNE j employs no factors at location m . We make this our baseline specification and call it *lean stacking* because we do not restrict μ_{mk} , κ_{mk} , or δ_{mk} at the location of absence so that production and factor employment at other locations can interact with the location of absence. We include the set of absence indicators $(\mathbf{1} - \mathbf{d}_{jt})$ to offset the zero output prediction at the sample mean.

Under *strict stacking*, not only the translog cost function coefficients are interacted with presence indicators but all variables that enter equation m are set to zero in addition. This is equivalent to additionally interacting the translog coefficients μ_{mk} , κ_{mk} , or δ_{mk} with presence indicators for all k . Either method of stacking can induce correlations between the transformed regressors and the error ϵ_{jt}^ℓ in (6). To remove this source of potential bias, we include the set of absence indicators $(\mathbf{1} - \mathbf{d}_{jt})$ (with nuisance parameters β_d^ℓ) among the regressors in the outcome equation: $y_{jt}^\ell = \mathbf{x}_{jt}^\ell \beta^\ell = \mathbf{x}_{jt}^{0\ell} \beta^\ell + (\mathbf{1} - \mathbf{d}_{jt}) \beta_d^\ell$.

3 Estimation under Location Selectivity

The selection equation (5) for location ℓ is

$$d_{jt}^\ell = \mathbf{1} (H(\mathbf{z}_{j,t-\tau}) + \eta_{jt}^\ell > 0)$$

and, conditional on MNE j 's selection of location ℓ , expectations of the outcome (6) are

$$\mathbb{E} [y_{jt}^\ell | \mathbf{x}_{jt}^\ell, \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}] = \mathbf{x}_{jt}^\ell \beta^\ell + \mathbb{E} [\epsilon_{jt}^\ell | d_{jt}^1, \dots, d_{jt}^\ell = 1, \dots, d_{jt}^L; \mathbf{z}_{j,t-\tau}],$$

where disturbances ϵ_{jt}^ℓ and η_{jt}^ℓ are uncorrelated across observations (of MNEs i and j , and between periods t and $t+1$).

In this section, we discuss cross-regional distributional assumptions on $(\epsilon_{jt}^\ell, \eta_{jt}^\ell)$ and permissible estimation techniques under those conditions. For a parametric cost function specification (with well-defined elasticities of substitution), a parametric approach to selectivity appears natural to start with. We present sets of necessary and sufficient distributional assumptions for univariate Heckman (1979) corrections. Empirical evidence on the necessary assumptions, however, is mixed in our sample. For multivariate selectivity, an extension of the Heckman (1979) estimator has a complicated form (conditional moments of multivariate normal distributions have no known closed form for multiple truncations, see Kotz, Balakrishnan, and Johnson (2000)). Simulated maximum-likelihood would be a viable technique but requires joint multivariate normality.

To be free of distributional restrictions, we resort to a nonparametric multivariate selection model (similar to one in Das, Newey, and Vella (2003)) to account for cross-location correlations between labor demand choices at the extensive and intensive margins. We derive identification from common sufficient assumptions. The nonparametric procedure allows for unknown disturbance distributions, and for unknown functional forms of $\mathbb{E}[\epsilon_{jt}^\ell | \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}]$ and $\mathbf{1}(H(\mathbf{z}_{j,t-\tau}) + \eta_{jt}^k > 0)$.

3.1 Univariate selection corrections

Consider Heckman (1979) selectivity corrections location by location. There are two alternative sets of assumptions that allow for univariate Heckman (1979) corrections, whereby labor demand (6) in ℓ only requires correction for selectivity (5) into ℓ but not into any other locations $k \neq \ell$. We are interested in $\mathbb{E}[y_{jt}^\ell | \mathbf{x}_{jt}^\ell, \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}]$ and $H(\mathbf{z}_{j,t-\tau}) = \mathbf{z}_{j,t-\tau}\gamma^\ell - \gamma_N^\ell + (\gamma_X^\ell + \gamma_N^\ell)d_{j,t-\tau}^\ell$.

Assumption 1 *The disturbances $(\epsilon_{jt}^k, \eta_{jt}^\ell)$ are multivariate normally distributed and independent of \mathbf{x}_{jt}^m and $\mathbf{z}_{j,t-\tau}$ for all k, ℓ, m (and $\text{Var}(\eta_{jt}^\ell) = 1$). In addition, either*

- (a) *the part of the selection shock that correlates with labor demand shocks is an MNE-specific disturbance and does not vary by location (ϵ_{jt}^k and η_{jt}^ℓ correlate across locations $k \neq \ell$ but in the same way as ϵ_{jt}^ℓ and η_{jt}^ℓ), or*
- (b) *the labor-demand related part of the selection shock varies by location but is independent of labor demand shocks in other locations (ϵ_{jt}^k and η_{jt}^ℓ are independent for $k \neq \ell$),*

for $\ell, k = 1, \dots, L$.

Intuitively, in case (a) when the part of the selection shock that correlates with labor demand shocks is an MNE-specific disturbance and does not vary by location, all relevant information for labor demand at any location ℓ is fully contained in the single indicator d_{jt}^ℓ (which is as informative about η_{jt}^ℓ as any other location indicator).

Lemma 1 *Independent univariate Heckman (1979) selection corrections for L locations identify $\mathbf{x}_{jt}^\ell \beta^\ell$ and $\text{Cov}(\epsilon_{jt}^\ell, \eta_{jt}^\ell)$ if and only if Assumption 1 holds.*

Proof. Because any normally distributed variable can be linearly decomposed into a sum of independent standard normal variables, consider without loss of generality

$$\eta_{jt}^\ell = \sqrt{1-\omega} e_{jt}^\ell + \sqrt{\omega} \sum_{k \leq \ell} \frac{\pi_\eta^{k\ell}}{\sqrt{\sum_{k \leq \ell} (\pi_\eta^{k\ell})^2}} u_{jt}^k, \quad (7)$$

$$\epsilon_{jt}^\ell = \sum_k \lambda^{k\ell} e_{jt}^k + \sum_{k \leq \ell} \pi_\epsilon^{k\ell} v_{jt}^k \quad (8)$$

for independent standard normal variables $e_{jt}^k, u_{jt}^k, v_{jt}^k$ ($k = 1, \dots, L$), where $\omega \in [0, 1]$ is a weight to satisfy $(\sigma_\eta^\ell)^2 = \sigma_\eta^{\ell\ell} = 1$, and $\pi_\eta^{k\ell}, \pi_\epsilon^{k\ell}, \lambda^{k\ell}$ are parameters. To prove sufficiency, let $\pi_\eta^{k\ell} = \pi_\epsilon^{k\ell} = 0$ for $k \neq \ell$.

First consider (a) MNE-specific selection shocks η_{jt}^ℓ whose labor demand related part does not vary over locations. Concretely, set $e_{jt}^k = e_{jt}$ for all locations k , and denote $\lambda^\ell \equiv \sum_k \lambda^{k\ell}$. Then the variances and covariances of the selection shocks (7) are $\sigma_\eta^{\ell\ell} = 1$ and $\sigma_\eta^{k\ell} = 1 - \omega$. The variances and covariances of the labor demand shocks (8) are $\sigma_\epsilon^{\ell\ell} = (\lambda^\ell)^2 + (\pi_\epsilon^{\ell\ell})^2$ and $\sigma_\epsilon^{k\ell} = (\lambda^\ell)^2$. And the covariances between the selection shock in location k and the demand shock in location ℓ are $\sigma_{\eta\epsilon}^{k\ell} = \lambda^\ell$.

Second, consider (b) location-varying selection shocks η_{jt}^ℓ that are independent of labor demand shocks in other locations. Concretely, set $\lambda^{k\ell} = 0$ for $k \neq \ell$, and denote $\lambda^\ell \equiv \lambda^{\ell\ell}$ for comparability. Then the selection shock variances and covariances are $\sigma_\eta^{\ell\ell} = 1$ and $\sigma_\eta^{k\ell} = 0$. The variances and covariances of the labor demand shocks are $\sigma_\epsilon^{\ell\ell} = (\lambda^\ell)^2 + (\pi_\epsilon^{\ell\ell})^2$ and $\sigma_\epsilon^{k\ell} = 0$. The covariances between the selection shock in location k and the demand shock in location ℓ are $\sigma_{\eta\epsilon}^{\ell\ell} = \sqrt{1-\omega} \lambda^\ell$ and $\sigma_{\eta\epsilon}^{k\ell} = 0$ for $k \neq \ell$.

In both cases, the marginal likelihood function becomes

$$g(y_{jt}^\ell | \mathbf{x}_{jt}^\ell, \mathbf{z}_{j,t-\tau}) = \frac{\phi((y_{jt}^\ell - \mathbf{x}_{jt}^\ell \beta^\ell) / \sigma_\epsilon^\ell)}{\sigma_\epsilon^\ell \Phi(\mathbf{z}_{j,t-\tau} \gamma^\ell)} \cdot \Phi\left(\frac{\rho_{\eta\epsilon}^{\ell\ell} (y_{jt}^\ell - \mathbf{x}_{jt}^\ell \beta^\ell) + \mathbf{z}_{j,t-\tau} \gamma^\ell}{\sigma_\epsilon^\ell (1 - \rho_{\eta\epsilon}^{\ell\ell})^{1/2}}\right), \quad (9)$$

after integrating out u_{jt}^ℓ and v_{jt}^ℓ , where $\sigma_\epsilon^\ell = \sqrt{\sigma_\epsilon^{\ell\ell}}$ and $\rho_{\eta\epsilon}^{\ell\ell} = \sigma_{\eta\epsilon}^{\ell\ell} / \sigma_\eta^\ell \sigma_\epsilon^\ell$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and distribution functions. This is precisely the likelihood function for independent Heckman (1979) correction location by location.

For necessity, observe that parameters $\pi_\eta^{k\ell} \neq 0$ or $\pi_\epsilon^{k\ell} \neq 0$ for any $k \neq \ell$ cause cross-equation correlations and do not permit integrating out u_{jt}^ℓ and v_{jt}^ℓ to arrive at (9). Similarly, $\lambda^{k\ell} \neq 0$ for any $k \neq \ell$ precludes integrating out e_{jt}^ℓ to arrive at (9). ■

Note that cross-location correlations of labor demand shocks are not necessarily evidence against the applicability of univariate Heckman (1979) selection corrections. Case (a) of MNE-specific selection shocks induces a correlation between the labor demand shocks across locations (but no independent correlation is permissible for univariate selectivity corrections).

Estimation. Extending the univariate Heckman (1979) two-stage procedure to L locations, we first estimate equations (5) with probit regressions by location. Second, we estimate outcome (6) at location ℓ by including the predicted selectivity hazard (inverse of the Mills ratio) $\hat{\Lambda}_{jt}^\ell$ from the first stage among the regressors (we also include absence indicators $(\mathbf{1} - \mathbf{d}_{jt})$ among the regressors to prevent stacking bias). The coefficient on the predicted selectivity hazard equals $\beta_\lambda^\ell \equiv \rho_{\epsilon\eta}^\ell \sigma_\epsilon^\ell$. We implement the second-stage estimation of (6) for $L-1$ locations (excluding home) by iterating Zellner’s (1962) seemingly unrelated regression (SUR) over the estimated disturbance covariance matrix until the estimates converge. This is equivalent to maximum-likelihood estimation (Dhrymes 1971) and makes estimation invariant to the deleted location equation L (Barten 1969).⁷ Through constraints, we impose linear homogeneity in factor prices and symmetry of wage coefficients (see appendix B). We treat induced heteroskedasticity following Heckman (1979) (resulting in differing standard errors on symmetric coefficients). After estimation, we test whether either of the two possible sets of distributional assumptions are satisfied. We will find implications of both sets violated.

Tests. Implications of Assumption 1 are testable. In case (a) of MNE-specific selection shocks and for any $\omega < 1$, Assumption 1 implies that $\sigma_\eta^{k\ell}$ is the same for any pair of locations $k \neq \ell$. Note that we have no evidence on $\sigma_{\eta\epsilon}^{k\ell}$ for $k \neq \ell$ from location-by-location estimation. We obtain estimates of $\sigma_\eta^{k\ell}$ from multivariate probit estimation instead and use a χ^2 -test for their equality.

Under the additional assumption that $\omega = 0$, there is a further test to query case (a), whether selection shocks are purely MNE-specific. Probit (maximum likelihood) estimation of selection in the Heckman procedure does not predict the disturbances η_{jt} . A testable implication of an MNE-specific selection shock, however, is that, if an MNE is neither present in all locations nor absent from all locations, the choices of presence and absence must be consistent with a location-independent MNE-specific selection shock for all locations. Concretely, an MNE observation contradicts the assumption of a location-independent selection shock if $\mathbf{z}_{j,t-\tau}\gamma^k - F_{j,t-\tau}^k > \mathbf{z}_{j,t-\tau}\gamma^\ell - F_{j,t-\tau}^\ell$ for locations k of absence and locations ℓ of presence because η_{jt} can be subtracted from both sides of the inequalities. This implication is testable for the predicted values, which are normally distributed conditional on $\mathbf{z}_{j,t-\tau}$ and $\mathbf{d}_{j,t-\tau}$ by normality of η_{jt} .

For (b) location-variant selection shocks, the set of assumptions implies that $\sigma_\epsilon^{k\ell} = 0$. So, a regression of ϵ_{jt}^ℓ on $\epsilon_{jt}^1, \dots, \epsilon_{jt}^{\ell-1}, \epsilon_{jt}^{\ell+1}, \dots, \epsilon_{jt}^L$ must have zero coefficients. We test this implication.

Both sets (a) and (b) of assumptions imply that ϵ_{jt}^k is independent of d_{jt}^k for all k because ϵ_{jt}^k and η_{jt}^ℓ are independent. We include absence indicators $(\mathbf{1} - \mathbf{d}_{jt})$ among

⁷We also estimate the full system of $2(L-1)$ probit and labor demand equations with maximum likelihood, restricting $\sigma_\eta^{k\ell} = 0$. Results are available at <http://econ.ucsd.edu/muendler/research>.

the regressors in the outcome equation, however, so this is not a testable implication in our context.

3.2 Multivariate selection correction

In the nonparametric version of the multivariate binary choice model (5) and (6),

$$d_{jt}^\ell = \mathbf{1}(H(\mathbf{z}_{j,t-\tau}) + \eta_{jt}^\ell > 0), \quad (\ell = 1, \dots, L)$$

$$\mathbb{E}[y_{jt}^\ell | \mathbf{x}_{jt}^\ell, \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}] = \mathbf{x}_{jt}^\ell \beta^\ell + \mathbb{E}\left[\epsilon_{jt}^\ell \mid d_{jt}^\ell = 1, \mathbf{d}_{jt}^{k \neq \ell}; \mathbf{z}_{j,t-\tau}\right],$$

no distributional assumptions are placed on η_{jt}^ℓ or ϵ_{jt}^ℓ and $H(\cdot)$ is an unknown function.

We augment the nonparametric sample selection model in Das, Newey, and Vella (2003) to remain identified under multivariate binary selection (similar in spirit to a selection model with endogeneity in Das, Newey, and Vella (2003)). Suppose η_{jt}^k and ϵ_{jt}^ℓ are correlated. Suppose also that $\mathbf{z}_{j,t-\tau}$ and \mathbf{x}_{jt}^ℓ are correlated (e.g. wages in the past and present, as our data show). Because d_{jt}^k is a function of η_{jt}^k , it correlates with ϵ_{jt}^ℓ ; because d_{jt}^k is a function of $\mathbf{z}_{j,t-\tau}$, it correlates with \mathbf{x}_{jt}^ℓ . So, if the labor demand equation does not condition on d_{jt}^k , the identifying restriction that \mathbf{x}_{jt}^ℓ and y_{jt}^ℓ are uncorrelated will be violated.

Define the *propensity score* (the expected probability of selection conditional on $\mathbf{z}_{j,t-\tau}$) as $p_{jt}^\ell \equiv \mathbb{E}[d_{jt}^\ell | \mathbf{z}_{j,t-\tau}] = 1 - G(-H(\mathbf{z}_{j,t-\tau}))$, where $G(\cdot)$ is the cumulative distribution function of η_{jt}^ℓ . Then, assuming $G(\cdot)$ is one-to-one and changing variables with $u_{jt}^\ell = 1 - G(\eta_{jt}^\ell)$, labor demand at the extensive margin becomes

$$\begin{aligned} \mathbb{E}[\epsilon_{jt}^\ell | d_{jt}^\ell = 1, \mathbf{d}_{jt}^{k \neq \ell}, \mathbf{z}_{j,t-\tau}] &= \mathbb{E}[\epsilon_{jt}^\ell | \eta_{jt}^\ell > -H(\cdot), \mathbf{d}_{jt}^{k \neq \ell}, \mathbf{z}_{j,t-\tau}] \\ &= \mathbb{E}[\epsilon_{jt}^\ell | u_{jt}^\ell < p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}] \\ &= \int \int_0^{p_{jt}^\ell} \epsilon_{jt}^\ell f(\epsilon_{jt}^\ell, u_{jt}^\ell | \mathbf{d}_{jt}^{k \neq \ell}) d\epsilon_{jt}^\ell du_{jt}^\ell / p_{jt}^\ell \\ &= m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}). \end{aligned}$$

So, the conditional labor demand disturbance for location ℓ depends only on the propensity score for that location and the pattern of presence elsewhere. Observed labor demand then satisfies

$$\mathbb{E}[y_{jt}^\ell | \mathbf{x}_{jt}^\ell, \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}] = \mathbf{x}_{jt}^\ell \beta^\ell + m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}).$$

To establish identification, consider deviations from the truth $\Delta \xi^\ell(\mathbf{x}_{jt}^\ell) \equiv \mathbf{x}_{jt}^\ell (\hat{\beta}^\ell - \beta^\ell)$ and $\Delta m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) \equiv \hat{m}^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) - m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$, where hats denote estimates of the true (not hatted) functions. Assumption 2 states sufficient conditions for identification.

Assumption 2

- (i) $\mathbb{E}[\epsilon_{jt}^\ell | d_{jt}^\ell = 1, \mathbf{d}_{jt}^{k \neq \ell}, \mathbf{z}_{j,t-\tau}] = m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}),$
 - (ii) $\Pr(\Delta \xi^\ell(\mathbf{x}_{jt}^\ell) + \Delta m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) = 0 | d_{jt}^\ell = 1) = 1$ implies that $\Delta \xi^\ell(\mathbf{x}_{jt}^\ell)$ is constant,
 - (iii) $\nabla_{\mathbf{z}_{j,t-\tau}} p_{jt}^\ell \neq \mathbf{0}$ with probability one,
- for $\ell = 1, \dots, L$.

Part (i) states that the conditional expectation of the labor demand disturbance at location ℓ is only a function of the propensity score of presence at ℓ and observed presence elsewhere. So, in the regression of observed labor demand y_{jt}^ℓ on $\mathbf{x}_{jt}^\ell \beta^\ell$ and $m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$, $\mathbf{x}_{jt}^\ell \beta^\ell$ is a separate additive component. This specification extends nonparametric selectivity correction in Das, Newey, and Vella (2003) to the multivariate case.

Part (ii) is the same identification condition as in Das, Newey, and Vella (2003) and implies that p_{jt}^ℓ (which enters $m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$) depends on variables in $\mathbf{z}_{j,t-\tau}$ that are not in $\mathbf{x}_{jt}^\ell \beta^\ell$. Otherwise, a regression of y_{jt}^ℓ on $\mathbf{x}_{jt}^\ell \beta^\ell$ leaves $\Delta \xi^\ell(\mathbf{x}_{jt}^\ell) = m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$ and $\Delta m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) = -m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$ indeterminate—a violation of (ii). In our context, the exclusion restriction arises naturally because the MNE chooses \mathbf{x}_{jt}^ℓ in response to information after $t - \tau$, whereas the decision of presence is based on $\mathbf{z}_{j,t-\tau}$. In addition, parent-firm characteristics and competitor-level host-country characteristics are predictors of presence but not related to the labor-cost specific part of the cost function other than through wages themselves. The rank condition (iii) requires that the information set $\mathbf{z}_{j,t-\tau}$ predicts the propensity score.

Assumption 2 allows us to relax the earlier identifying assumption that $(\epsilon_{jt}^k, \eta_{jt}^\ell)$ is independent of \mathbf{x}_{jt}^m and $\mathbf{z}_{j,t-\tau}$ for all k, ℓ, m . Assumption 2 only requires that, conditional on the propensity score p_{jt}^ℓ , ϵ_{jt}^ℓ is uncorrelated with all functions of \mathbf{x}_{jt}^ℓ and $\mathbf{z}_{j,t-\tau}$. Moreover, the nonparametric estimator \mathbf{x}_{jt}^m allows for conditional heteroskedasticity of unknown form (and thus presents a nonparametric alternative to Chen and Khan's (2003) three-step estimator). Also note that we need no assumption on the cross-equation correlation of η_{jt}^ℓ if we include $\mathbf{d}_{jt}^{k \neq \ell}$. This makes nonparametric analysis a powerful tool for multivariate binary selection estimation.

Lemma 2 *If Assumption 2 holds and if $m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$ and $p_{jt}^\ell(\mathbf{z}_{j,t-\tau})$ are continuously differentiable and have continuous distribution functions almost everywhere, then $\mathbf{x}_{jt}^\ell \beta^\ell$ and $m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$ are identified up to additive constants.*

Proof. In any observationally equivalent model it must be the case that the observed outcome satisfies $\mathbb{E}[y_{jt}^\ell | \mathbf{x}_{jt}^\ell, \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}] = \mathbf{x}_{jt}^\ell \hat{\beta}^\ell + \hat{m}^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$ for some $\mathbf{x}_{jt}^\ell \hat{\beta}^\ell$ and $\hat{m}^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$. Equivalently, deviations from the truth $\Delta \xi^\ell(\mathbf{x}_{jt}^\ell) + \Delta m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) = 0$.

This identity must be differentiable with respect to \mathbf{x}_{jt}^ℓ and $\mathbf{z}_{j,t-\tau}$ by continuous differentiability of $m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$ and $p_{jt}^\ell(\mathbf{z}_{j,t-\tau})$. So,

$$\begin{aligned}\nabla_{\mathbf{x}_{jt}^\ell} \Delta \xi^\ell(\mathbf{x}_{jt}^\ell) &= \mathbf{0}, \\ (\partial \Delta m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) / \partial p_{jt}^\ell) \cdot \nabla_{\mathbf{z}_{j,t-\tau}} p_{jt}^\ell &= \mathbf{0}.\end{aligned}$$

The first equation implies that $\Delta \xi^\ell(\mathbf{x}_{jt}^\ell) = \mathbf{x}_{jt}^\ell (\hat{\beta}^\ell - \beta^\ell) = c_1$ for a constant c_1 and $\mathbf{x}_{jt}^\ell \beta^\ell$ is identified up to this constant. By $\nabla_{\mathbf{z}_{j,t-\tau}} p_{jt}^\ell \neq \mathbf{0}$, the second equation implies that $\Delta m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) = \hat{m}^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) - m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) = c_2$ for a constant c_2 and $m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell})$ is identified up to that constant. ■

Note that lacking identification of additive constants is not a problem in our context. The transformed cost function regressors $\mathbf{x}_{jt}^\ell \beta^\ell$ in equation (6) do not include a constant term. To assess the labor demand effect of permanent wage differentials at the extensive margin, we will evaluate $\nabla_{p_{jt}^\ell} m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) \cdot \nabla_{\mathbf{z}_{j,t-\tau}} p_{jt}^\ell$ (a scalar), for which the constant does not matter.

Conversely, if we want to include the propensity scores $\mathbf{p}_{jt}^{k \neq \ell}$ in the second-stage regression, instead of the presence indicators $\mathbf{d}_{jt}^{k \neq \ell}$, we can only do so if η_{jt}^ℓ and ϵ_{jt}^k are uncorrelated across locations ($k \neq \ell$). This is a drawback of identification under Assumption 2.

Suppose we are interested in a broader definition of the extensive margin,

$$\bar{y}_{jt}^{\text{ext},\ell} \equiv \mathbb{E} [\epsilon_{jt}^\ell \mid d_{jt}^\ell = 1; \mathbf{z}_{j,t-\tau}],$$

which does not condition on the observed location pattern outside ℓ . This definition allows us to investigate the impact of a permanent wage differential (in $\mathbf{z}_{j,t-\tau}$) through its effect on the entire grid of an MNE's potential locations. Formally, we can now evaluate $\nabla_{\mathbf{p}_{jt}} m^\ell(\mathbf{p}_{jt}) \cdot \nabla_{\mathbf{z}_{j,t-\tau}} \mathbf{p}_{jt}$ (a matrix), where \mathbf{p}_{jt} is the vector of propensity scores. Under the restriction that η_{jt}^ℓ and ϵ_{jt}^k are not correlated across locations ($k \neq \ell$), d_{jt}^k is not correlated with ϵ_{jt}^k because ϵ_{jt}^ℓ must be uncorrelated with all functions of $\mathbf{z}_{j,t-\tau}$. Then we can relax item (i) in Assumption 2 to $\mathbb{E}[\epsilon_{jt}^\ell \mid d_{jt}^\ell = 1, \mathbf{z}_{j,t-\tau}] = m^\ell(\mathbf{p}_{jt})$.

Assumption 3

- (i) $\mathbb{E}[\epsilon_{jt}^\ell \mid d_{jt}^\ell = 1, \mathbf{z}_{j,t-\tau}] = m^\ell(\mathbf{p}_{jt})$ and $\text{Cov}(\epsilon_{jt}^\ell, \eta_{jt}^k) = 0$ for $k \neq \ell$,
- (ii) $\Pr(\Delta \xi^\ell(\mathbf{x}_{jt}^\ell) + \Delta m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) = 0 \mid d_{jt}^\ell = 1) = 1$ implies that $\Delta \xi^\ell(\mathbf{x}_{jt}^\ell)$ is constant,
- (iii) $\nabla_{\mathbf{z}_{j,t-\tau}} p_{jt}^\ell \neq \mathbf{0}$ with probability one,

for $\ell = 1, \dots, L$.

Lemma 3 follows as a corollary to Lemma 2 (replacing the scalar derivative $\partial\Delta m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k\neq\ell})/\partial p_{jt}^\ell$ with the vector $\nabla_{\mathbf{p}_{jt}}\Delta m^\ell(\mathbf{p}_{jt})$, and $\nabla_{\mathbf{z}_{j,t-\tau}}p_{jt}^\ell$ with $\nabla_{\mathbf{z}_{j,t-\tau}}\mathbf{p}_{jt}$).

Lemma 3 *If Assumption 3 holds and if $m^\ell(\mathbf{p}_{jt})$ and $p_{jt}^\ell(\mathbf{z}_{j,t-\tau})$ are continuously differentiable and have continuous distribution functions almost everywhere, then $\mathbf{x}_{jt}^\ell\beta^\ell$ and $m^\ell(\mathbf{p}_{jt})$ are identified up to additive constants.*

Das, Newey, and Vella (2003) establish convergence rates and asymptotic normality of similar estimators on the basis of smoothness properties of $p_{jt}^\ell(\mathbf{z}_{j,t-\tau})$ and $m^\ell(\mathbf{p}_{jt})$ (and a generalization of $\mathbf{x}_{jt}^\ell\beta^\ell$ to a function of \mathbf{x}_{jt}^ℓ) for splines and power series. We use power series to approximate $p_{jt}^\ell(\mathbf{z}_{j,t-\tau})$ and $m^\ell(\mathbf{p}_{jt})$. Power series are root- n asymptotic normal and can estimate smooth functionals of unknown parameters (Newey 1997). Most important for our application, the first derivative of the power series estimator is a smooth functional and hence also root- n asymptotic normal.

Estimation. We first estimate equations (5) with individual linear regressions by location. We use a third-order polynomial in wages and two additional predictors, alongside otherwise linear predictors (to break the curse of dimensionality). Second, we include the predicted propensity scores \hat{p}_{jt}^ℓ from the first stage on the second stage (6). Under Assumption 2 we approximate $m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k\neq\ell})$ with a third-order polynomial in p_{jt}^ℓ , interacted with $\mathbf{d}_{jt}^{k\neq\ell}$ (we continue to include absence indicators $(\mathbf{1}-\mathbf{d}_{jt})$ without interactions to both approximate $m^\ell(\cdot)$ and remove potential stacking bias). Under Assumption 3 we approximate $m^\ell(\mathbf{p}_{jt})$ with a third-order polynomial in \mathbf{p}_{jt} (and include absence indicators $(\mathbf{1}-\mathbf{d}_{jt})$ among the regressors to remove potential stacking bias). We implement the second-stage estimation of (6) for $L-1$ locations (excluding home) by iterating SUR over the estimated disturbance covariance matrix until the estimates converge. Through constraints, we impose linear homogeneity in factor prices and symmetry of wage coefficients (see appendix B).

3.3 Wage Elasticities of Labor Demand

We use elasticities of substitution to quantify the responses of multinational labor demand y_{jt}^ℓ to permanent wage changes. The (constant-output) *cross-price elasticity of substitution* between factors ℓ and k is defined as $\varepsilon_{\ell k} \equiv \partial \ln y_{jt}^\ell / \partial \ln w^k$ and becomes

$$\varepsilon_{\ell k}^T = \frac{\psi_{\ell k} + s^\ell s^k}{s^\ell} \quad (k \neq \ell) \quad \text{and} \quad \varepsilon_{\ell \ell}^T = \frac{\psi_{\ell \ell} + s^\ell(s^\ell - 1)}{s^\ell} \quad (10)$$

for a short-run translog cost function function, where $s^\ell = w^\ell y^\ell / C$ is the cost share of the workforce at ℓ (the wage bill at location ℓ in the MNE's total wage bill) and

$\psi_{\ell k} \equiv \partial s_{jt}^\ell / \partial \ln w^k$ is the marginal change of the cost share at ℓ in response to a log wage change at k . These elasticities can be calculated both for each individual MNE- j observation and in the aggregate using sample means. Reported elasticity estimates are sample means from cost function coefficients and observed wage bill shares.

A permanent change of the wage level w^k in location k is reflected in both vectors of regressors \mathbf{x}_{jt}^ℓ (with w_t^k) and $\mathbf{z}_{j,t-\tau}$ (with $w_{t-\tau}^k$). So, the response of the cost share s_{jt}^ℓ to a permanent change in $\ln w_t^k$ is

$$\psi_{\ell k} = \delta_{\ell k} + \mathbb{E}[\epsilon_{jt}^\ell | \cdot, w_{t-\tau}^k] / \partial w_{t-\tau}^k \equiv \psi_{\ell k}^{\text{int}} + \psi_{\ell k}^{\text{ext}}. \quad (11)$$

The first term in (11) captures the labor demand response at the intensive margin $\psi_{\ell k}^{\text{int}} \equiv \partial s_{jt}^\ell / \partial w_t^k$. The second term in (11) is a measure of the labor demand response to a permanent change in w^k at the extensive margin $\psi_{\ell k}^{\text{ext}} \equiv \partial s_{jt}^\ell / \partial w_{t-\tau}^k$.

By (6), the labor demand response at the intensive margin is $\psi_{\ell k}^{\text{int}} = \delta_{\ell k}$ under any of the Assumptions 1 through 3. The labor demand response at the extensive margin, however, depends on the identifying assumption:

$$\psi_{\ell k}^{\text{ext}} = \begin{cases} \gamma_{w^k}^\ell \beta_\Lambda^\ell \Delta_{jt}^\ell \cdot w_t^\ell w_t^k / C_{jt} & \text{Assumption 1,} \\ (\partial m^\ell(p_{jt}^\ell, \mathbf{d}_{jt}^{k \neq \ell}) / \partial p_{jt}^\ell) \cdot (\partial p_{jt}^\ell / \partial w_{t-\tau}^k) \cdot w_t^\ell w_t^k / C_{jt} & \text{Assumption 2,} \\ \nabla_{\mathbf{p}_{jt}} m^\ell(\mathbf{p}_{jt}) \cdot \nabla_{w_{t-\tau}^k} \mathbf{p}_{jt} \cdot w_t^\ell w_t^k / C_{jt} & \text{Assumption 3.} \end{cases} \quad (12)$$

We multiply by present wages w_t^k because estimation on the first stage uses w_t^k as regressors, not their logs. We divide by C_{jt}/w_t^ℓ to convert estimates from labor demand equation (6) back into their cost share equivalents because we also use $\psi_{\ell k}^{\text{int}} = \delta_{\ell k}$ at the intensive margin. Under Heckman (1979) correction (Assumption 1), $\gamma_{w^k}^\ell$ is the wage coefficient in the selection equation, $\beta_\Lambda^\ell \equiv \rho_{\epsilon\eta}^{\ell\ell} \sigma_\epsilon^\ell$ is the coefficient on the selectivity hazard in the outcome equation, and Δ_{jt}^ℓ is the first derivative of the selectivity hazard Λ_{jt}^ℓ (the inverse of the Mills ratio) with respect to its scalar argument, $\Delta_{jt}^\ell(\mathbf{z}_{j,t-\tau}\gamma^\ell) \equiv \Lambda_{jt}^\ell(\mathbf{z}_{j,t-\tau}\gamma^\ell) [\Lambda_{jt}^\ell(\mathbf{z}_{j,t-\tau}\gamma^\ell) - \mathbf{z}_{j,t-\tau}\gamma^\ell]$. Because $\Delta_{jt}^\ell(\cdot) \in (0, 1)$, the sign of the log wage effect on the wage bill at the extensive margin is the sign of the product $\gamma_{w^k}^\ell \beta_\Lambda^\ell$ (the coefficients on the two stages of estimation). Under polynomial series estimation, the derivatives of $m^\ell(\cdot)$ and p_{jt}^ℓ are the marginal effects on the third-order polynomials, evaluated at the sample mean.⁸

We run 200 bootstraps on the two-stage procedure to find standard errors for our elasticity estimates. Bootstrapping does not require any treatment of insignificant wage coefficients from the first-stage regressions in our quantification of the extensive margin. Moreover, Eakin, McMillen, and Buono (1990) show in simulations that analytic confidence intervals for elasticity estimates can widely differ from bootstrapped confidence interval estimates.

⁸If w_t^ℓ is a strictly location-specific variable, equation (12) does not apply to $k = \ell$ since w_t^ℓ drops from a binary probit likelihood function. By our variable construction, w_t^ℓ is MNE j 's competitors' mean factor price exposure. It is thus also MNE-specific.

4 Data and Descriptive Statistics

Our main data source is a confidential three-dimensional panel (parent-affiliate-year observations) of German MNEs at Deutsche Bundesbank (BuBa). We retain manufacturing parents and majority-owned manufacturing affiliates only. We transform this panel to parent-location-year observations and combine the data with complementary information on wages and host-country characteristics from various sources.

Firm-level data. Information on turnover, employment and fixed assets of foreign affiliates stems from BuBa’s MIDI database (MIcro database Direct Investment, formerly DIREK), which derives its outward FDI information from legally mandated annual questionnaires on all foreign corporate holdings by German firms and households above minimum ownership shares and capital stock thresholds (Lipponer 2003). Individually identified outward FDI data are available for the years 1996-2001 and provide two-digit NACE 1.1 sector classifications for the parent (and affiliate) firms. We restrict our sample to majority-owned foreign affiliates because estimation of a multilocation cost function suggests the use of observations of parent firms with full managerial control and because majority ownership is insensitive to a change in the notification threshold in MIDI 1999. Assets and capital structure of *every* majority-owned foreign firm are reported in MIDI, including in years with zero turnover. Turnover does not distinguish within-MNE shipments from final sales but is nevertheless a proxy to affiliate production for cost function estimation.

Balance sheet and income statement information for German parent firms comes from BuBa’s USTAN database, which records this information for German firms that draw a bill of exchange (for a documentation in English see Deutsche Bundesbank (1998)). Whereas MIDI covers by design the full sample of German MNEs and their majority-owned affiliates, USTAN includes only firms that draw a bill of exchange in a given year. Because the bill of exchange is a common form of payment among firms of all sizes throughout the sample period 1996-2001 (though losing some popularity thereafter), USTAN is considered the most comprehensive source of balance sheet data for companies of all sizes outside the financial sector in Germany. The MIDI and USTAN data were linked by parent name and address in our previous work (Becker, Ekholm, Jäckle, and Muendler 2005).⁹

To obtain interpretable results, we lump host countries into four *aggregate locations*: CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), and WEU (Western Europe); see table 15 in the Appendix for definitions. As Table 1 shows, the four aggregate foreign locations

⁹Our conservative string matching routine filtered out potential duplicates from time-varying firm identifiers in USTAN. In manual treatments, only doubtlessly identifiable parent pairs from MIDI and USTAN were kept. At the expense of reduced sample size, this caution guarantees the formation of time-consistent parent pairs.

Table 1: EMPLOYMENT AT GERMAN MNEs IN 2000

	HOM	CEE	DEV	OIN	WEU
	(1)	(2)	(3)	(4)	(5)
Employment	1,387,817 ^a	394,579	319,221	245,721	332,622
Estimation sample employment	938,867	191,854	139,240	125,199	184,560
Mean employment per sample MNE	1,588.6	282.6	736.7	387.6	407.4

Sources: MIDI and USTAN 1996 to 2001, manufacturing MNEs and their majority-owned foreign manufacturing affiliates. Locations: HOM (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

^aPredicted German employment at in- and out-of-sample MNEs, based on linear employment regressions to account for incomplete MIDI-USTAN matches.

host similarly large manufacturing workforces for German manufacturing MNEs, between 250,000 and 400,000 employees. MIDI and USTAN matches are incomplete so that we do not observe parent employment for every German MNE. For comparisons, we predict total parent employment for the full sample of German manufacturing MNEs from a linear regression of parent employment on foreign employments and estimate that German manufacturing MNEs with majority-owned foreign manufacturing affiliates employ about 1.4 million German workers. Conditional on MNE presence, the largest employment per MNE in the estimation sample occurs at affiliates in DEV and the smallest employment at affiliates in CEE.

Table 2 shows changes to the presence patterns of German MNEs between 1996 and 2000. Adjustments are infrequent. More than half of all MNEs who are present in only one foreign location have a presence only in one foreign location in 2000 (small numbers in row 2). Among firms who remain MNEs in both years, more than four in five with a presence in only one location abroad in 1996 keep exactly one foreign location (large numbers in row 2). In general, entries along the diagonal exhibit the highest frequency in every row and every column. Regional expansions are gradual: the frequencies above the diagonal decrease monotonically in every row. Regional exits, however, are not gradual: MNEs who exit most frequently abandon all foreign locations at once; frequencies in the first column dominate frequencies below the diagonal in every row. There is a large number of complete withdrawals between 1996 and 2000 (477 out of 2,247 MNEs). Note that our FDI data are designed to cover the population of German firms with FDI, and sample attrition is mitigated by the legal obligation to report and Deutsche Bundesbank's commitment to follow up on missing questionnaires. MNEs who are present in several foreign locations in 1996 expand their operations to additional locations more frequently than do MNEs with presence in a smaller number of foreign locations: the frequency mass above the diagonal increases row by row among the German parents with MNE status in both years.

Table 2: MNE COUNTS OF FOREIGN LOCATIONS

<i>L</i> in 1996	<i>L</i> in 2000					<i>Total</i> (100%)
	1	2	3	4	5	
1	0.0%	83.5%	12.2%	2.6%	1.6%	794
2		83.7%	12.5%	3.2%	0.6%	687
	34.7%	54.7%	8.2%	2.1%	0.4%	1,052
3		23.7%	55.8%	15.8%	4.7%	190
	28.0%	17.1%	40.2%	11.4%	3.4%	264
4		11.1%	25.0%	45.8%	18.1%	72
	24.2%	8.4%	19.0%	34.7%	13.7%	95
5		7.4%	3.7%	22.2%	66.7%	27
	35.7%	4.8%	2.4%	14.3%	42.9%	42
<i>Total</i>		630	211	91	44	976
	477	1,293	308	112	57	2,247

Source: MIDI population 1996 and 2000 (not matched to USTAN), manufacturing MNEs and their majority-owned foreign manufacturing affiliates. Locations: Home (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe); see table 15 for definitions.

Among the 2,247 MIDI MNEs with foreign presence either in 1996 or 2000, CEE was the region where MNEs opened most new affiliates, 18.2 percent more in 2000 than in 1996, followed by DEV with 12.6 percent, OIN with 3.2 percent and WEU with 2.0 percent. German MNEs typically pursue a single-affiliate strategy of foreign expansions: the median number of affiliates of a German MNE per location is one. Table 3 shows that, once an MNE has established its presence in a given location with at least one affiliate, the number of affiliates hardly changes: 859 out of 1,259 observations of MNEs in given locations exhibit no change to the number of affiliates between 1996 and 2000; 247 out of 1,259 observations of MNEs in their locations increase or decrease the number of affiliates by one. The small remainder of 153 manufacturing parents chooses to change the number of affiliates by more. (The MNE total in Table 3 is smaller than that in Table 2 because we condition on presence in a location.) Together, Tables 2 and 3 suggest that MNEs face potentially large sunk costs of foreign presence. Changes to the number of host countries within locations are even more infrequent than changes to the number of affiliates: an analysis of host country changes similar to Table 3 shows that 947 out of 1,259 observations of MNEs in given locations exhibit no change in the number of selected host countries within the location.

Table 3: MNE COUNTS OF CHANGING AFFILIATE NUMBERS

$N_{2000} - N_{1996}$	CEE	DEV	OIN	WEU	<i>MNE Total</i>
	(1)	(2)	(3)	(4)	(5)
≤ -3	2	3	2	15	22
-2	3	11	3	14	31
-1	6	17	11	64	98
0	186	131	145	397	859
+1	25	32	20	72	149
+2	11	11	4	16	42
+3	2	6	4	10	22
$\geq +4$	7	11	4	14	36
<i>MNE Total</i>	242	222	193	602	1,259
\bar{N}_{2000}	1.49	2.38	1.56	1.96	
\bar{N}_{1996}	1.41	2.28	1.50	2.01	

Sources: MIDI population 1996 and 2000 (not matched to USTAN). MNEs with regional presence of at least one affiliate in 1996; manufacturing MNEs and their majority-owned foreign manufacturing affiliates. Locations: Home (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe). Median number of affiliates by MNE, location and year: 1.

Infrequent net changes to the number of affiliates and countries could, in principle, conceal gross changes such as changes to the country composition within a location or exit and reentry with a different affiliate. In fact, only small fractions of MNEs who maintain a constant number of affiliates within a location change countries within the aggregate location. In both CEE and WEU 4.2 percent of MNEs with constant affiliate numbers between 1996 and 2000 change country, and 7.2 percent of the MNEs with constant affiliate numbers in DEV change country, but none do so in OIN. Similarly small fractions are associated with changing affiliate IDs. Motivated by these findings, we define the extensive margin (selection into a location) as the presence of an MNE in an aggregate location with at least one affiliate. We do not distinguish the few country changes within aggregate locations for selection estimation, but our labor demand (outcome) estimation accounts for varying country-level exposure.

We deflate parent variables with the German CPI and deflate affiliate variables with country-level CPIs (from the IMF's International Financial Statistics). CPI deflation factors are re-based to unity at year end 1998. We transform foreign currency values to their EUR equivalents in December 1998 in order to remove nominal exchange rate fluctuations. December 1998 is the mid point in time for our

1996-2001 sample. Introduction of the euro in early 1999 makes December 1998 a natural reference date. See Appendix C for details on currency conversion.

Complementary data. Wage information is not reported in MIDI. For foreign wages, we use OWW data by occupation (Occupational Wages around the World, Freeman and Oostendorp 2001) between 1983 and 1999 and conduct robustness checks with UBS wage data for 1994, 1997, 2000 and 2003. We deflate and currency-convert wages in accordance with all other variables, and transform them into annual wages. We reuse OWW wages from 1999 in 2000 and 2001, and we linearly interpolate UBS wages between survey years. Though domestic wage information is available from USTAN, we take (NACE 1.1 two-digit) sector wages at the occupation level from the German annual earnings survey (table 62321 from the federal statistical online portal *destatis.de/genesis*). Occupational wage information from the German annual earnings survey enters the ILO database, on which OWW wages are based, so that our foreign and domestic wage measures are compatible. To mitigate possible compositional effects in our labor demand regression on wages, we take sectoral medians over seven occupation groups for Germany and take country medians over 161 OWW occupation groups. In the robustness checks with UBS data, we use UBS wages also for Germany (and lose sectoral variation). Appendix D provides further details on wage variable construction.

National accounts information for host-country regressors comes from the World Bank’s World Development Indicators and the IMF’s International Financial Statistics series. We use CEPII bilateral trade and geographic data (*www.cepii.fr*) to compute market access to a host country as in Redding and Venables (2004), see Appendix E. To condition selection estimation on skill endowments beyond labor costs, we include the host country’s percentage of high-school or higher educated residents in 1999 from Barro and Lee (2001) and interact the variable with an indicator whether the percentage exceeds that in Germany (19.5%).¹⁰

Table 4 shows means of variables by location in the censored panel for labor-demand estimation. In our main specification, we consider multinational labor demand during the years 1998-2001 (called t) for a sample of 1,345 MNEs and their location selection two years prior to production ($t - \tau$). For robustness checks, we also use a single cross-section of 266 MNEs in 2000 and their location selection in 1996. The frequency of MNE presence abroad increased by two to four percentage points between 1996-99 and 1998-2001 in all locations but WEU (Western European countries) where it slightly fell. German MNEs spend the bulk of their wage bill

¹⁰For estimation of location selection, we also experimented with German import and export data from 2000 as controls for trade in the MNE’s home sector. The import and export data were at the two-digit product level (matching NACE 1.1 two-digit sector codes) and by country of destination or origin (Fachserie 7, Reihe 7 from the federal statistical online portal *destatis.de/genesis*) but did not prove to be significant predictors of location selection.

Table 4: SAMPLE MEANS OF VARIABLES

(t : 1998-2001, $t - \tau$: 1996-99)	HOM	CEE	DEV	OIN	WEU
	(1)	(2)	(3)	(4)	(5)
Indic.: Presence in t	1	.380	.367	.349	.717
Indic.: Presence in $t - \tau$	1	.344	.334	.327	.718
MNE-wide regressors (Labor demand estimation)					
Wage bill share (t)	.810	.075	.029	.176	.172
ln Fixed assets (t)	17.433	14.860	15.216	15.840	15.251
ln Turnover (t)	18.575	15.987	16.309	17.311	16.967
ln Wage (t)	10.189	7.983	7.583	9.938	9.812
Competitor-average regressors (Selection estimation)					
ln sample-mean Wage ($t - \tau$)	10.175	7.988	7.799	9.914	9.801
Hosts' ln Market access ($t - \tau$)	11.456	10.745	12.816	12.742	11.701
Hosts' skill share < Home ($t - \tau$)	20.564	19.244	22.649	22.467	20.927
Hosts' skill share \geq Home ($t - \tau$)	42.783	40.834	48.838	48.945	44.217
Hosts' distance ($t - \tau$)	32.343	30.168	36.494	36.313	33.086
Hosts' ln Consumption p.c. ($t - \tau$)	28.478	26.669	31.867	31.770	29.092
Parent-level regressors (Selection estimation)					
Indic.: West German parent ($t - \tau$)	.975	.962	.978	.967	.977
ln Count of host countries ($t - \tau$)	1.204	1.416	1.697	1.515	1.344
ln Employment ($t - \tau$)	5.924	6.074	6.691	6.421	6.130
ln Equity ($t - \tau$)	16.813	17.070	17.955	17.632	17.180
ln Liability ($t - \tau$)	17.864	18.150	18.742	18.402	18.113
ln Capital-labor ratio ($t - \tau$)	11.412	11.650	11.696	11.608	11.515
Parent observations	1,345	394	418	455	829

Sources: MIDI and USTAN 1996 to 2001, censored (second-stage) estimation sample of 1,345 MNEs. Averages of MNE variables are conditional on presence. Locations: HOM (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

(81 percent) at home. From German MNEs, CEE receives labor expenditures of a magnitude similar to the remaining developing world combined. (Note that shares do not add to unity across columns because averages are conditional on presence, omitting absent MNEs). A similar cross-location pattern arises for turnover and capital stocks.

Substantial wage disparities persist across locations. Between Germany and CEE, for instance, MNE wages differ by a factor of around 900 percent ($\exp(10.189 - 7.983) = 9.08$ for 1998-2001). This MNE-level difference is larger than the unconditional wage gap of 780 percent ($1/.128$) in the raw OWW wage data in 2000. The larger conditional differential reflects MNE selection into CEE: if parent firms in sectors with higher home wages are more likely to expand to CEE, or if MNEs are

more likely to invest in CEE countries with lower wages than regional average, then the conditional MNE-level wage differential exceeds the unconditional differential.

Choice-specific variables (host country attributes) are not identified in binomial choice models such as probit for univariate normal selection correction. We estimate our model also in an MNE cross-section where we have no time-varying host country attributes. We therefore transform host country attributes to competitor-averages by MNE, and use competitor-average transformations in all procedures for comparability. We group MNEs into eight manufacturing sectors¹¹ and calculate mean host-country attributes over all competitor observations by location and sector. We take the total of competitors' foreign employments as host-country weights within the location. The wage at $t - \tau$ in CEE, for example, is the average wage paid at competitor's affiliates in CEE. In Table 4, we only take means over MNEs with presence in a given location so that the table reports CEE wages of the competitors of a German MNE with FDI in CEE.¹² German MNEs in CEE, compared to any other location, face competitors in host countries that offer the least market access, that have the smallest skill endowments, that are geographically the closest and that exhibit the smallest per-capita consumption. The CEE wages paid by competitors of MNEs in CEE are slightly above those paid by competitors in DEV. MNEs in OIN face competitors with the strongest host-country market access and host-country skill endowments.

Parent-level covariates are suggestive of selectivity effects at their means. Parents with headquarters in East Germany (including West Berlin) are slightly more likely to expand to CEE and OIN than the average German MNE. For all other parent-level regressors, regional conditional means (columns 2 to 5) exceed the unconditional mean (column 1), and regional means tend to be the lower the higher the frequency of MNE presence. Conditional on their presence abroad, MNEs exhibit larger home workforces, larger parent-level equity or debt, and higher parent-level capital-labor ratios.

5 Estimation

A permanent wage differential between an MNE's home and a foreign location directly affects employment at the intensive margin through labor reallocation across existing affiliates. A permanent wage differential indirectly affects labor demand at the extensive margin by altering the likelihood of presence, which in turn changes conditional expectations of labor demand. We estimate both margins.

¹¹The sectors are: food; textiles and leather; wood, pulp and paper; chemicals, rubber, plastic and energy producing materials; mineral and metal products; machinery and equipment; transport equipment; manufactures not elsewhere classified.

¹²We use the wage level at $t - \tau$ as a regressor in selection estimation, not its log. For comparisons to the the log wage at t , we report the log of the sample-mean wage at $t - \tau$ in Table 4.

The effect of home wages on employment is identifiable at both margins from sector variation in a cross-section of German MNEs because individual wage-taking firms face bargained earnings schedules from sectoral agreements between unions and employers’ associations (with one-year to two-year terms).¹³ Time variation of home wages provides additional identification. Similarly, both time variation and variation across locations identify employment effects of foreign wages at the intensive margin. Identification of foreign wages at the extensive margin is more limited, however. Because binomial choice models (of presence or absence) cannot identify coefficients of choice-specific variables (host country attributes), foreign wage changes at the extensive margin are mainly identified over time. We obtain additional variation by considering competitor-average foreign wages which vary by MNE. To clear wage variables of workforce composition effects, we use country-wide skill-group medians for foreign wages and sector-wide skill medians for home wages. Estimation at the intensive margin conditions on a firm’s MNE status.

5.1 Location choice

We estimate binomial choices of presence in up to four foreign locations—CEE, DEV, OIN and WEU—with probit regressions for univariate normal selectivity corrections (Assumption 1) and with series estimators of the selection propensity for nonparametric selectivity corrections (Assumptions 2 or 3).

Probit estimation. To have a first idea of sunk costs in location choice, Table 5 shows probit probability estimates from a short regression of MNE presence (with at least one affiliate in a location) on past presence indicators across locations and a constant. In our sample of MNEs, past presence between 1996 and 1999 at a given location is a highly significant predictor of presence two years later in that location (and will continue to be highly significant in the long regression). Presence indicators elsewhere serve as rudimentary controls.

Coefficient estimates from Table 5 decompose the sunk cost components. Recall that the sunk cost part of location selection (5) can be represented with

$$F_{j,t-\tau}^\ell = \gamma_N^\ell - (\gamma_X^\ell + \gamma_N^\ell) d_{j,t-\tau}^\ell,$$

where γ_N are sunk entry costs, γ_X^ℓ sunk exit costs, and $(\gamma_X^\ell + \gamma_N^\ell)$ is also called the hysteresis band. Table 6 shows the decomposition result. Because the entry and exit cost decomposition involves the estimate of the constant, entry and exit costs cannot be expressed in marginal probability terms of their own. A marginal probability measure can be inferred for their sum, the hysteresis band.

¹³The use of median sector home wages and location selectivity controls removes potential firm-level bargaining effects behind labor demand coefficients on home wages. Foreign affiliates of German MNEs are few and small, with no impact on foreign wage levels.

Table 5: SUNK-COST COEFFICIENTS IN SHORT PROBIT REGRESSION

	CEE	DEV	OIN	WEU
	(1)	(2)	(3)	(4)
Indic.: FDI in CEE ($t - \tau$)	2.112 (.060)***	-.181 (.067)***	-.131 (.071)*	-.290 (.058)***
Indic.: FDI in DEV ($t - \tau$)	-.169 (.069)**	2.200 (.063)***	.124 (.070)*	-.156 (.061)**
Indic.: FDI in OIN ($t - \tau$)	-.149 (.071)**	.146 (.069)**	2.274 (.066)***	-.140 (.063)**
Indic.: FDI in WEU ($t - \tau$)	-.461 (.056)***	-.220 (.059)***	-.310 (.062)***	1.760 (.051)***
Const.	-.872 (.044)***	-1.241 (.049)***	-1.319 (.050)***	-.707 (.042)***
Obs.	3,392	3,392	3,392	3,392
Pseudo R^2	.380	.395	.419	.304

Sources: MIDI 1996 to 2001, pooled sample of manufacturing MNEs and their majority-owned foreign manufacturing affiliates with two-year selection lags ($\tau = 2$). Standard errors in parentheses: * significance at ten, ** five, *** one percent. Locations: Home (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

Past presence increases the likelihood of future presence in a given location by more than seventy percent in all but WEU, where the marginal effect predicts a more than sixty percent increase. The long probit regression will confirm these magnitudes. The total, however, hides the differential impact of entry and exit costs. Entry cost are the largest in the distant low-income and high-income locations DEV and OIN, and dominate exit costs there. Conversely, entry costs are the lowest in the nearby low-income and high-income locations CEE and WEU, and significantly smaller than exit costs. Among the exit costs are the opportunity costs of absence. German MNEs are considerably less reluctant to leave distant locations DEV and OIN than they abandon the neighboring locations CEE or WEU.

Indicators for past FDI presence may not exclusively capture sunk costs but also firm heterogeneity. In the long regression, we include a set of firm-level predictors alongside competitor-average host country attributes. Table 7 presents the marginal effects for the full list of covariates.¹⁴ In the long regression, only past presence indicators for the region itself are significantly different from zero, whereas past presence in other regions has no predictive power (except for past presence in CEE

¹⁴Sample size drops from 3,392 to 2,414 mainly because of missing information from parent balance sheets.

Table 6: SUNK ENTRY AND EXIT COSTS IN PROBABILITY TERMS

	CEE	DEV	OIN	WEU
	(1)	(2)	(3)	(4)
Sunk entry cost: γ_N	.872***	1.241***	1.319***	.707***
Sunk exit cost: γ_X	1.240***	.959***	.954***	1.053***
Hysteresis band: $(\gamma_N + \gamma_X)$	2.112***	2.200***	2.274***	1.760***
Marginal effect of hysteresis band	.704***	.710***	.714***	.621***

Sources: MIDI 1996 to 2001, 3,392 pooled observations of manufacturing MNEs and their majority-owned foreign manufacturing affiliates with two-year selection lags. Estimates are probit coefficients from Table 5. Significance levels from χ^2 tests. Standard errors in parentheses: * significance at ten, ** five, *** one percent. Locations: Home (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

significantly reducing the frequency of future presence in DEV). The probability effects of past presence at the same location confirm the salience of the hysteresis band. Beyond a binary indicator of past presence, the number of host countries in the past significantly raises the likelihood of presence in the future. German MNEs with large home employment, low parent debt, and a high capital-labor ratio at the parent firm are significantly more likely to be present at most or all foreign locations within two years.

Wages are weak predictors of binary location choices, where only their time and, by construction, their competitor variation identify effects. At the ten (five) percent significance level, five (three) out of twenty wage variable coefficients are statistically different from zero. The home wage coefficient has a counterintuitive negative sign in selection estimation for CEE but the expected positive sign in all other locations. Insignificant effects of host country wages on location choice are common in the literature on location choice (e.g. Devereux and Griffith (1998) for U.S., Head and Mayer (2004) for Japanese, and Buch, Kleinert, Lipponer, and Toubal (2005) for German MNEs). Bootstraps over both estimation stages will show nevertheless that statistically weak wage predictions of location selection, combined with strong labor demand effects of selection, significantly matter for elasticities of labor substitution.

Further covariates (not reported) include competitors' hosts' market access, competitors' hosts' skill shares, competitors' hosts' distance, competitors' hosts' per-capita consumption, an indicator of parents' headquarters in West Germany, equity, and parent profits/equity. No covariate is significant at the five-percent level in any location. To tentatively control for an outside margin of arm's length trade between independent firms, we also included a set of sector and location specific import and export measures but found the trade variables not to be significant pre-

Table 7: MARGINAL EFFECTS IN LONG PROBIT REGRESSION

Predictors ($t - \tau$)	CEE	DEV	OIN	WEU
	(1)	(2)	(3)	(4)
Indic.: FDI in CEE	.779 (.021)***	-.065 (.025)***	-.009 (.023)	-.022 (.041)
Indic.: FDI in DEV	-.040 (.033)	.669 (.028)***	.011 (.025)	-.050 (.043)
Indic.: FDI in OIN	-.018 (.032)	-.003 (.027)	.711 (.026)***	-.049 (.041)
Indic.: FDI in WEU	-.002 (.034)	-.046 (.031)	-.035 (.026)	.743 (.020)***
ln Count of host countries	.040 (.040)	.085 (.035)**	.017 (.029)	.121 (.052)**
ln Employment	.119 (.026)***	.058 (.024)**	.070 (.022)***	.142 (.031)***
ln Liability	-.089 (.022)***	-.046 (.019)**	-.055 (.017)***	-.158 (.026)***
ln Capital-labor ratio	.085 (.022)***	.026 (.020)	.039 (.018)**	.072 (.026)***
Home sector wage	-.007 (.010)	.002 (.009)	.0005 (.008)	.016 (.012)
Competitors' wages CEE	-.292 (.139)**	-.312 (.136)**	-.115 (.120)	-.005 (.172)
Competitors' wages DEV	-.005 (.011)	-.021 (.012)*	-.024 (.014)*	.009 (.014)
Competitors' wages OIN	.044 (.033)	.018 (.031)	.006 (.027)	-.094 (.040)**
Competitors' wages WEU	-.0003 (.009)	.003 (.009)	.008 (.008)	-.006 (.011)
Obs.	2,414	2,414	2,414	2,414
Pseudo R^2	.559	.523	.547	.450

Sources: MIDI and USTAN 1996 to 2001, pooled sample of manufacturing MNEs and their majority-owned foreign manufacturing affiliates with two-year selection lags ($\tau = 2$). Standard errors in parentheses: * significance at ten, ** five, *** one percent. Further regressors (not significantly different from zero at five percent level in any location): Competitors' hosts ln Market access, Indic. of Headquarters West Germany, ln Equity, Parent profits/equity, Competitors' hosts skill shares, Competitors' hosts distance, Competitors' hosts ln Consumption per capita. Locations: CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

dictors of location choice. Compared to those potentially important but statistically insignificant predictors of FDI, wages perform better.

Nonparametric propensity score approximations. We estimate the propensity score of location choice with a third-order polynomial in wages, market access, and the count of an MNE’s host countries, alongside the same linear predictors as for probit estimation. The predicted propensity scores are .335 for CEE, .323 for DEV, .300 for OIN and .631 for WEU—slightly under-predicting the actual frequencies of presence in Table 4 but properly reflecting the relative frequencies across locations.

Table 8 reports coefficient estimates by location. Marginal effects are very close to those in the probit regressions. Estimates of the hysteresis band along the diagonal of past presence indicators continue to have the same magnitude as in probit estimation but past presence in WEU, in addition to that in CEE, now significantly predicts absence from locations off the diagonal. All significant parent-level covariates from probit remain significant predictors. Similarly, insignificant parent-level covariates remain insignificant except for the coefficient on market access which becomes a significant predictor in OIN after including its higher-order terms and wage interactions in the series expansion. We present F -tests of joint significance of individual wages for p values at or below the .1 threshold. Now ten (five) out of twenty wage coefficients are significant at the ten (five) percent level.

5.2 Translog estimation with selectivity correction

We proceed to estimation of the short-run translog cost function and include predicted selection hazards from probit estimation as regressors in the equation system (univariate normal selectivity correction, Assumption 1). Alternatively, we include predicted propensity scores from nonparametric selection estimation (Assumption 2 or 3).

Translog cost function estimation. Table 9 presents estimates of translog cost function equations for 1,345 German MNE observations between 1998 and 2001 under lean stacking. (We lose observations mainly because of missing wage information at affiliate locations.) Coefficient estimates are from iterated seemingly unrelated regressions of wage bill shares on their translog predictors for four out of five locations, excluding home. For the regression, wage bill shares and covariates are scaled by observation-specific cost-wage ratios to remove truncation at unity. Beyond the reported wage coefficients, the equations include the full sets of turnover and fixed asset regressors, the scaled equivalent of the constant, and indicators of absence from all other locations. All but one wage coefficient in Table 9 is significantly different from zero at the one percent level. Most coefficients on output and fixed assets (not reported) are similarly highly significant.

Table 8: MARGINAL EFFECTS IN NONPARAMETRIC PROBABILITY MODEL

Predictors ($t - \tau$)	CEE (1)	DEV (2)	OIN (3)	WEU (4)
Indic.: FDI in CEE	.758 (.018)***	-.048 (.018)***	-.018 (.017)	-.018 (.022)
Indic.: FDI in DEV	-.037 (.019)*	.669 (.020)***	-.008 (.018)	-.028 (.024)
Indic.: FDI in OIN	-.022 (.018)	-.015 (.019)	.712 (.017)***	-.036 (.022)
Indic.: FDI in WEU	-.024 (.019)	-.044 (.020)**	-.040 (.019)**	.707 (.024)***
ln Count of host countries	-41.729 (24.664)*	-26.072 (25.317)	-27.678 (23.695)	-48.850 (30.293)
Competitors' hosts ln Market access	2.552 (9.313)	2.412 (9.560)	28.153 (8.947)***	2.760 (11.438)
ln Employment	.061 (.014)***	.042 (.014)***	.044 (.013)***	.084 (.017)***
ln Liability	-.046 (.011)***	-.030 (.012)***	-.034 (.011)***	-.093 (.014)***
ln Capital-labor ratio	.042 (.011)***	.023 (.012)*	.026 (.011)**	.042 (.014)***
Competitors' hosts ln Cons. p.c.	.0008 (.013)	-.018 (.013)	-.006 (.012)	-.036 (.016)**
Series terms: p -values from F tests				
Home sector wage terms	.074	.030	.009	.032
Competitors' CEE wage terms	.097	.033	.027	
Competitors' DEV wage terms			.058	.072
Competitors' OIN wage terms	.098			
Competitors' WEU wage terms				
Obs.	2,414	2,414	2,414	2,414
R^2	.664	.619	.634	.560

Sources: MIDI and USTAN 1996 to 2001, pooled sample of manufacturing MNEs and their majority-owned foreign manufacturing affiliates with two-year selection lags ($\tau = 2$). Standard errors in parentheses: * significance at ten, ** five, *** one percent. Further regressors (not significantly different from zero at five percent level in any location): Indic. of Headquarters West Germany, ln Equity, Parent profits/equity, Competitors' hosts skill shares, Competitors' hosts distance. Locations: CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

Table 9: TRANSLOG COST PARAMETER ESTIMATES

Labor cost shares in: (transformed)	CEE (1)	DEV (2)	OIN (3)	WEU (4)
Univariate Normal Selectivity Correction (Assumption 1)				
<i>ln Wages</i>				
HOM	.020 (.001)***	-.015 (.0007)***	.123 (.006)***	.111 (.007)***
CEE	-.011 (.001)***	.002 (.0001)***	-.003 (.0006)***	-.008 (.0006)***
DEV	.002 (.0002)***	.007 (.0006)***	-.0005 (.0004)	.007 (.0005)***
OIN	-.003 (.0002)***	-.0005 (.00008)***	-.131 (.005)***	.011 (.002)***
WEU	-.008 (.0002)***	.007 (.0001)***	.011 (.002)***	-.121 (.007)***
Selectivity hazard	23.183 (8.425)***	1.685 (16.609)	78.205 (18.861)***	78.764 (22.380)***
R^2	.726	.813	.938	.853
Nonparametric Selectivity Correction (Assumption 2)				
<i>ln Wages</i>				
HOM	.019 (.002)***	-.011 (.001)***	.114 (.009)***	.087 (.008)***
CEE	-.008 (.002)***	.001 (.0002)***	-.003 (.0008)***	-.009 (.0008)***
DEV	.001 (.0002)***	.006 (.0008)***	-.001 (.0005)***	.005 (.0006)***
OIN	-.003 (.0008)***	-.001 (.0005)***	-.123 (.008)***	.012 (.003)***
WEU	-.009 (.0008)***	.005 (.0006)***	.012 (.003)***	-.095 (.008)***
Series terms				
χ^2 tests (p -value)	320.2 (.000)	385.1 (.000)	237.1 (.000)	125.6 (.000)
R^2	.741	.823	.941	.846

Sources: MIDI and USTAN 1996 to 2001. Stacked observations of 1,345 MNEs. Further regressors: \ln Turnover, \ln Fixed assets, Absence indicators, Transformed constant (in univariate normal selectivity regression). Standard errors in parentheses: * significance at ten, ** five, *** one percent. Standard errors corrected for first-stage estimation of selectivity hazards (thus not symmetric on restricted coefficients). Locations: HOM (omitted), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

Equation estimates in the upper panel of Table 9 include the predicted selectivity hazards (inverses of Mills ratios) by location (Assumption 1). Selectivity hazards are statistically different from zero at the one percent level in all equations except DEV, where significance lacks at common levels. The lower panel presents estimates from nonparametric selectivity correction (Assumption 2), using third-order polynomials in the location's propensity score interacted with indicators for presence at all other locations. χ^2 tests on the series terms overwhelmingly reject their joint equality to zero. The regression fits (as measured by R^2) are slightly higher under nonparametric selectivity correction (Assumption 2) in three equations but not in WEU. Overall, we consider the significance of selectivity correction terms strong evidence for the importance of the extensive margin.

Tests for univariate selectivity correction. We test whether Assumption 1 of univariate normal selection correction is satisfied in our context. There are two cases: (a) the part of the selection shock that correlates with labor demand shocks is an MNE-specific disturbance and does not vary by location, and (b) the labor-demand related part of the selection shock varies by location but is independent of labor demand shocks in other locations. We test the two cases in turn. The evidence is mixed. Tests are not inconsistent with case (a) but they reject case (b).

First consider (a) MNE-specific selection shocks whose labor demand related part does not vary by location. This case implies that the covariance between selection disturbances is the same for any pair of locations $k \neq \ell$. We obtain estimates of those covariances from multivariate probit estimation of simultaneous selection into the four foreign locations. In the cross section of MNEs in 2000 with multivariate probit regressors from 1996, we fail to reject the joint equality of six correlation coefficients between the four equations with a χ^2 test statistic of 3.545 (p value .738). Under a restriction on the selection disturbance, another implication of case (a) is that, if an MNE is neither present in all locations nor absent from all locations, the choices of presence and absence must be consistent with a location-independent MNE-specific selection shock for all locations. We calculate the regression prediction for all MNEs that are not omnipresent, pick the maximum probit prediction among all locations of absence and the minimum probit prediction among all locations of presence thus stacking the cards of the test against case (a). Although 202 out of 1,942 observations show the wrong pattern, a one-sided t -test of the null hypothesis that absence and presence predictions are reversed is rejected overwhelmingly with a statistic of 77.67 (zero p value).

Second, consider (b) location-varying selection shocks that are independent of labor demand shocks in other locations. This assumption implies that a regression of labor demand residuals from one location on other residuals from other locations must have zero correlation coefficients. We overwhelmingly reject this hypothesis for all six pairs of the four location residuals with zero p values in five cases and a

p value of .01 in one case (OIN-WEU).

While case (b) is soundly rejected, there is no clear evidence against case (a) where selection shocks correlate with labor demand shocks only through an MNE-specific error but not through location-specific errors. Note that cross-location correlations of labor demand shocks are not evidence against case (a) because MNE-specific selection shocks themselves induce a correlation between the labor demand shocks across locations. Case (a) is plausible in economic terms. Suppose selection disturbances include both host country-specific parts such as, for example, surprising changes to profit repatriation policies and include MNE-specific parts such as shocks to sunk entry costs. Changes to host country repatriation policies affect the entry decision. But once the MNE operates in the host country, it maximizes profits irrespective of entry-relevant host-country shocks so that cost function disturbances are unrelated to the entry-relevant policy shocks. We therefore regard estimation under univariate normal selectivity correction (Assumption 1) a relevant benchmark.

Elasticities of multinational labor substitution. Table 10 shows own-wage and cross-wage substitution elasticities for permanent wage changes by one percent in different locations, separately for the extensive and the intensive margins. There is no well-defined extensive margin for selection into the home location (Germany) in a sample of MNEs, which are observable to the statistician only when active in the home location. The standard errors are from 200 bootstraps over the two estimation steps of univariate normal selectivity corrected translog estimates (Assumption 1). One margin at a time is set to zero to isolate the effect at the other margin. Cross-price elasticities are affine transformations of translog coefficients (equation (10)). While log wage effects on wage bill shares are additive in translog estimation at the intensive and the extensive margin (equation (6)), cross-wage substitution elasticities are not.

Own-wage elasticities along the diagonal—for both the intensive and the extensive margins—are uniformly negative as production theory requires. While this is expected for estimates at the intensive margin, it is a reassuring finding for estimates at the extensive margin. Note that we impose linear homogeneity in factor prices and symmetry of wage coefficients at the intensive margin through constraints on the translog regression, but we do not restrict estimates at the extensive margin because those constraints are not well defined at the observation level—neither under univariate normal nor under nonparametric selectivity correction. The own-wage elasticity of substitution is considerably larger in most foreign locations than at home, suggesting that MNE employment abroad responds more sensitively to labor costs there than home employment responds to home wages.

Cross-wage elasticities in the first row (foreign wage effects on home employment) and in the first column (home wage effects on foreign employment) are significantly positive for ten out of thirteen estimates at the intensive and the extensive margins.

Table 10: CROSS-WAGE ELASTICITIES UNDER UNIVARIATE NORMAL SELECTIVITY

Employment change (%) in	Wage change (by 1%) in				
	HOM (1)	CEE (2)	DEV (3)	OIN (4)	WEU (5)
HOM <i>intensive</i>	-.493***	.047***	-.010	.213***	.243***
CEE <i>extensive only</i>	.663***	-1.651***	-.0007	.752	.102
CEE <i>intensive only</i>	1.711***	-1.466***	.090***	-.095	-.240
DEV <i>extensive only</i>	.808***	-.049	-.995***	.087	.111
DEV <i>intensive only</i>	-.886	.223***	-.188	.00006	.851***
OIN <i>extensive only</i>	.881	-1.828	-.275	-.243	.920
OIN <i>intensive only</i>	2.877***	-.035	.00001	-3.139***	.297
WEU <i>extensive only</i>	1.372***	.002	.037	-2.484	-1.027***
WEU <i>intensive only</i>	1.849***	-.050	.071***	.167	-2.037***

Sources: MIDI and USTAN 1996 to 2001. Elasticities at the extensive and intensive margins from 1,345 stacked MNE observations. Underlying labor demand estimates $Y_j^\ell \beta^\ell$ from univariate normal selectivity-corrected ISUR estimates (Assumption 1, Table 9). Standard errors from 200 bootstraps: ** significance at five, *** one percent. Locations: HOM (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

A one-percent reduction in the wage in CEE, for instance, is associated with a .05 percent drop in home employment at German MNE parents. In contrast, a one-percent increase in the German home wage is associated with a 1.7 percent boost to MNE employment in CEE. So, home and CEE employment are substitutes within MNEs. The large difference in cross-wage effects is consistent with two main facts. First, employment at German MNE parents is larger in levels than at their CEE affiliates so that a smaller percentage wage drop in Germany means a larger reduction in employment in absolute terms. Second, CEE workers tend to be less productive than German workers, which is reflected in translog cost function coefficients.

The extensive margin is a noticeable component of adjustment, beyond its crucial role in correcting cost function estimates for location selectivity. Elasticities at the extensive margin are strictly positive. So, home and foreign employment are substitutes within MNEs not only at the intensive but also at the extensive margin. Although home wage effects on selection were not statistically different from zero on the first stage (Table 7), the strong significance of the selection effect on labor demand on the second stage (selectivity hazard coefficients in Table 9) turns home wage effects into significant predictors of employment substitution at the extensive margin. Beyond the marginal log wage coefficients from two-step estimation, ob-

served wage bill shares provide information for elasticity estimation and contribute to the significance of elasticity estimates.

Elasticities at the extensive margin are smaller in magnitude than at the intensive margin in the close locations CEE and WEU. In DEV, the extensive margin counteracts the insignificant (negative) elasticity at the intensive margin and we find a .8 percent increase in DEV employment in response to a one-percent home wage increase—similar in magnitude to that in CEE. In CEE, a one-percent increase in the German home wage is associated with a .7 percent increase in MNE employment at the extensive margin, if no adjustment occurs at the intensive margin.

We also add the intensive and extensive wage effects on wage bills and compute the total home wage elasticities of foreign labor demand. We find highly significant estimates for the total elasticities at three locations: 1.57 in CEE, 2.95 in OIN and 2.42 in WEU (significantly different from zero at the one-percent level). Our 200 bootstraps allow us to test whether the elasticities at the intensive margin are significantly different from the total elasticities. We reject their equality in all three locations (with t statistics between 2.7 and 16.7), corroborating the importance of the extensive margin. Note, however, that the effect of the home wage on selection into CEE (Table 7) has an unexpected negative sign so that the total cross-elasticity of labor demand with respect to a permanent home wage change (rather than margin-specific wage changes) is smaller than the elasticity at the intensive margin only. In the other selection regressions, the home wage has the expected positive sign so that the total elasticity exceeds the intensive elasticity.

Cross-wage estimates beyond the first row and column are for the most part not statistically different from zero. Notable exceptions at the intensive margin are two pairs of highly significant positive cross-wage effects: on the one hand of CEE on DEV (.22) and vice versa (.09), and on the other hand of DEV on WEU (.07) and vice versa (.85). The significantly positive and mutually consistent effects suggest that MNE employment is a substitute at the intensive margin between DEV and CEE and between DEV and WEU. There is no detectable substitution effect between home and DEV employment at German MNEs.

5.3 Specification comparisons

To assess the robustness of our estimates, we compare several further specifications and report the first rows of the cross-wage elasticity matrices (foreign wage effects on home employment) in Table 11, and the first columns separately by intensive and extensive margin in Tables 12 and 13 (home wage effects on foreign employment).

Foreign-wage elasticities of home employment are robust across specifications (Table 11). Estimates on our benchmark sample (first row) with OWW wages and MNEs between 1998 and 2001 under Assumption 1 conform closely to several other specifications. The striking similarity between the 1998-2001 MNE sample and the

Table 11: FOREIGN-WAGE ELASTICITIES OF HOME EMPLOYMENT

Home employment change (%)	Wage change (1%) in					Obs.
	HOM	CEE	DEV	OIN	WEU	
	(1)	(2)	(3)	(4)	(5)	(6)
Lean Stacking						
Ass. 1, OWW 98-01	-.493 (.061)***	.047 (.008)***	-.010 (.005)	.213 (.038)***	.243 (.045)***	1,345
Ass. 1, OWW 00	-.442 (.113)***	.043 (.019)**	.002 (.011)	.216 (.092)**	.181 (.041)***	266
Ass. 1, UBS 98-01	-.541 (.089)***	.026 (.009)***	-.029 (.019)	.136 (.036)***	.408 (.058)***	1,607
Ass. 2, OWW 98-01	-.457 (.060)***	.046 (.010)***	-.005 (.005)	.202 (.039)***	.214 (.039)***	1,345
Ass. 3, OWW 98-01	-.475 (.063)***	.047 (.009)***	-.006 (.005)	.225 (.039)***	.209 (.040)***	1,345
Omnipresent MNEs						
Ass. 1, OWW 98-01	-.711 (.589)	-.037 (.069)	-.010 (.042)	.661 (.433)	.098 (.263)	62
Stacking						
Ass. 1, OWW 98-01	-1.005 (.130)***	.086 (.033)***	-.012 (.008)	.442 (.078)***	.489 (.107)***	1,345

Sources: MIDI and USTAN 1996 to 2001. Elasticities of wage effects on home employment (first row of elasticity matrix) at the intensive margin. Standard errors from 200 bootstraps: ** significance at five, *** one percent. Locations: HOM (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

single cross section of MNEs in 2000 (with location choice in 1996) in the second row is consistent with the view that cross sectional and not time series variation is the main source of identification at the intensive margin. UBS wage data in the third row result in a larger estimation sample mainly because of their comprehensiveness for DEV and WEU, but sketchy UBS wage information in CEE (only wages for Budapest are continuously available from 1994 through 2003) perhaps introduces an attenuation bias in the CEE coefficient. Nonparametric estimation does not yield statistically different estimates, neither under Assumption 2 nor 3. The sample of 62 omnipresent MNEs between 1996 and 2001 is small and does not result in significant estimates on the second stage (we predict selectivity hazards from first-stage regressions on the full sample). Strict stacking of MNE observations considerably raises the point estimates. We choose to base our further analysis on the more conservative estimates from lean stacking. Estimates for DEV are not significant in any specification.

Table 12: HOME-WAGE ELASTICITIES AT THE INTENSIVE MARGIN

Employment change (%)	Home wage change (1%), by regression specification						
	Lean Stacking					Omnipr.	Stacking
	OWW 98-01	OWW 00	UBS 98-01	OWW 98-01	OWW 98-01	OWW 98-01	OWW 98-01
	Ass. 1	Ass. 1	Ass. 1	Ass. 2	Ass. 3	Ass. 1	Ass. 1
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
CEE	1.711 (.310)***	1.764 (.866)**	.947 (.303)***	1.677 (.371)***	1.719 (.374)***	-1.535 (2.777)	3.152 (1.196)***
DEV	-.886 (.534)	.195 (1.094)	-1.054 (.694)	-.482 (.483)	-.560 (.484)	-.454 (1.868)	-1.105 (.806)
OIN	2.877 (.516)***	2.925 (1.302)**	1.451 (.375)***	2.726 (.495)***	3.042 (.531)***	2.021 (1.330)	5.973 (1.108)***
WEU	1.849 (.328)***	1.628 (.339)***	1.844 (.255)***	1.629 (.284)***	1.589 (.290)***	.406 (1.070)	3.712 (.791)***
Obs.	1,345	266	1,607	1,345	1,345	62	1,345

Sources: MIDI and USTAN 1996 to 2001. Elasticities of home wage effects on foreign employment (first column of elasticity matrix) at the intensive margin. Standard errors from 200 bootstraps: ** significance at five, *** one percent. Locations: HOM (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

Home-wage elasticities of foreign employment at the intensive margin (Table 12) are robust too. Estimates on our benchmark sample (now in the first column) conform closely to several other specifications. In fact, the comments on the rows of Table 11 above apply also to the columns of Table 12.

At the extensive margin, home-wage elasticities of foreign employment (Table 13) are only significant in certain specifications. Not surprisingly, the single cross section of MNEs in 2000 (with location choice in 1996) does not result in significant estimates (second column). Competitor-level variation in wages for selection estimation does not suffice (when time series variation lacks) but, where standard errors are small (in CEE and WEU), point estimates are similar to other specifications. Estimates on our benchmark sample with MNEs between 1998 and 2001 under Assumption 1 are highly significant and largely insensitive to using OWW wages (first column) or UBS wages (third column).

Nonparametric estimates of elasticities at the extensive margin are sample means of the first derivatives of our third-order polynomial series expansions. We compute the elasticities after dropping outlier predictions, for which the first-stage probability model results in propensity scores outside the zero-one range. Contrary to our findings for the intensive margin, nonparametric estimates for the extensive mar-

Table 13: HOME-WAGE ELASTICITIES AT THE EXTENSIVE MARGIN

		Lean Stacking					Omnipr.	Stacking
		OWW	OWW	UBS	OWW	OWW	OWW	OWW
		98-01	00	98-01	98-01	98-01	98-01	98-01
Employmt. change (%)		Ass. 1	Ass. 1	Ass. 1	Ass. 2	Ass. 3	Ass. 1	Ass. 1
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
CEE		.663 (.190)***	.702 (.449)	.527 (.205)**	-1.627 (3.303)	9.753 (9.800)	.538 (.134)***	.644 (.480)
DEV		.808 (.187)***	-.930 (2.360)	.754 (.364)**	1.216 (8.555)	23.185 (22.484)	.543 (.398)	.753 (.538)
OIN		.881 (1.088)	-.421 (1.725)	1.586 (.858)	6.294 (4.179)	-5.559 (6.739)	.604 (.163)***	.762 (.685)
WEU		1.372 (.417)***	1.171 (.621)	1.682 (.587)***	-.827 (2.187)	4.679 (4.056)	.545 (.234)**	.860 (.511)
Obs.		1,345	266	1,607	1,345	1,345	62	1,345

Sources: MIDI and USTAN 1996 to 2001. Elasticities of home wage effects on foreign employment (first column of elasticity matrix) at the extensive margin. Standard errors from 200 bootstraps: ** significance at five, *** one percent. Locations: CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

gin (columns 4 and 5 of Table 13) are not statistically different from zero. Point estimates, however, are considerably larger in absolute size than in the benchmark specification, especially under Assumption 3. We view this as an indication that our parametric benchmark estimates of the extensive margin are conservative figures. Point estimates for omnipresent MNEs (column 6) are slightly smaller than in the benchmark specification, arguably because this highly selected sample expands to foreign location more frequently. Strict stacking yields (insignificant) point estimates of about the same magnitude as in the benchmark specification.

Taken together, our results confirm the statistical plausibility of the benchmark estimates from univariate normal selectivity correction (Assumption 1). Several tests for the validity of Assumption 1 fail to reject the identifying hypothesis that selection shocks correlate with labor demand shocks only through an MNE-specific error but not through location-specific errors. Nonparametric estimation yields very similar and highly significant elasticity estimates at the intensive margin. At the extensive margin, the benchmark estimates from univariate normal selectivity correction are highly significant and smaller in magnitude than nonparametric estimates. In short, the benchmark estimates from univariate normal selectivity correction are statistically robust. We now turn to the economic implications of our estimates for multinational labor substitution.

Table 14: COUNTERFACTUAL EMPLOYMENT EFFECTS OF A ONE-PERCENT REDUCTION IN THE HOME-FOREIGN WAGE GAP

Employment effect at the intensive margin on	Permanent wage gap reduction by one percent between Home and			
	CEE (1)	DEV (2)	OIN (3)	WEU (4)
Home ^a	651 (115)***	-135 (74)	2,955 (527)***	3,379 (622)***
Foreign ^b <i>extensive margin</i>	-1,628 (466)***	-2,689 (620)***	-2,813 (3474)	-5,413 (1644)***
Foreign ^b <i>total</i>	-3,858 (853)***	2,932 (1778)	-9,432 (3845)**	-9,537 (2314)***

Sources: Own calculations based on selectivity corrected translog estimates for 1,345 German manufacturing MNEs and their majority-owned foreign manufacturing affiliates in MIDI and USTAN between 1996 and 2001. Point estimates from univariate normal selectivity correction (Assumption 1, Table 10) multiplied by employment in 2000 (Table 1). Standard errors from 200 bootstraps: ** significance at five, *** one percent. Home (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

^aGap reducing foreign wage increases (by one percent).

^bGap reducing home wage reduction (by one percent).

5.4 Counterfactual Evaluation

Our hypothetical experiment is a permanent change in the wage differential between home and foreign locations. How much larger would parent employment be if the wage gap to foreign locations narrowed? How much smaller would affiliate employment be? Counterfactual predictions in Table 14 give answers to these questions. We use the home-wage elasticities of foreign labor demand and the foreign-wage elasticities of home labor demand from our selectivity corrected translog benchmark estimates for the 1998-2001 MNE sample (Table 10). These estimate reflect the employment responses of the mean MNE (the mean MNE in the stacked sample has propensities of presence abroad as in the first row of Table 4). We multiply the elasticity estimates with the workforce totals in Table 1 and obtain the implied employment changes from one-percent increases in wages by margin.

A one percent smaller wage gap between Germany and locations in CEE, for instance, is associated with 650 more jobs at German parents and 3,900 less jobs at affiliates in CEE. CEE workers tend to have lower labor productivity than German workers so that employment in CEE is more sensitive to home wage changes than home employment responds to foreign wages. The labor substitution effects between home locations and CEE are smaller than the effects relative to OIN or

WEU. In absolute magnitude, however, a complete closing of the HOM-CEE wage gap *at constant elasticities* results in larger employment effects than elimination of the HOM-OIN or HOM-WEU wage gaps. Wages in CEE are, on average, 12.8 percent of the German level in 2000. If the estimated elasticities of substitution are constant at all levels of wages, an increase in CEE wages of 780% (=1/.128) to remove the wage gap to Germany would bring 500,000 counterfactual manufacturing jobs (with a standard error of 90,000 jobs) to Germany—more than a third of the estimated home employment at German manufacturing MNEs (Table 1). The wage level in WEU (OIN) is 70.3 (81.6) percent of that in Germany so that an increase in WEU wages of 140% (=1/.703) (120% (=1/.816)) to close the gap would attract 480,000 (360,000) counterfactual manufacturing jobs to German plants of German manufacturing MNEs.

Elasticities of labor substitution are local properties of the MNE’s cost function, however, and the assumption of a constant elasticity of substitution at all wage levels is coarse. The rough calculations are merely intended to put an economic meaning to the abstract elasticity figures. In our view, the magnitude of our calculations for constant elasticities nevertheless underscores the potential importance of job substitution within MNEs for labor market outcomes.

6 Conclusion

While the public debate over outsourcing seems to have settled on the idea that multinational enterprises (MNEs) substitute jobs at home for foreign employment, economic studies on MNE labor demand across locations have found weak or no evidence of job substitution. We integrate two distinct approaches of the literature—one on predictions of MNEs’ location choices, and one on labor substitutability across established MNE locations—into a single econometric model that corrects cost function estimation for location selectivity. In our framework, multinational labor demand responds to wage differentials across locations both at the extensive margin, when a multinational enterprise expands into foreign locations, and at the intensive margin, when an MNE reallocates jobs across existing foreign affiliates. We derive conditions for common Heckman (1979) selectivity corrections, location by location, and for nonparametric identification. Cost function estimation conditions on MNE output. The empirical exercise thus leaves aside the counterfactual question how the market share and size of an MNE would differ if its access to foreign locations were limited in spite of global product market competition. This matter is part of our ongoing research.

Empirical evidence from a German sample of manufacturing MNEs shows that firms change multinational presence only infrequently and hardly alter their number of affiliates within regions. These infrequent changes to multinational presence at

the extensive margin give rise to rare but salient labor demand effects in response to permanent wage differentials across locations. Like much of the prior literature, we find wages to be poor predictors of binomial location choices. Statistically strong labor demand effects of location selection, however, magnify the weak wage effects and turn labor adjustment at the extensive margin into a statistically and economically significant predictor of MNE employment. With every percentage increase in Central and Eastern European wages, for instance, German manufacturing MNEs are found to allocate 650 jobs to Germany. With every percentage increase in German wages, German MNEs allocate 1,600 jobs to Central and Eastern Europe at the extensive margin and 3,900 jobs in total. Given the sizeable wage differential between Germany and Central and Eastern Europe (requiring a 750 percent increase in Eastern European wages in 2000 to close the gap), we conclude that international wage differentials have a salient impact on multinational labor substitution.

Appendix

A A Model of the MNE

We think of an MNE's choice of multinational production and sales as a two-stage decision problem. At moment $t - \tau$ (i.e. τ periods prior to production and sales), MNE j determines at which locations to produce and faces uncertainty over other MNEs' future output $\mathbf{q}_{i \neq j, t}$, input prices \mathbf{w}_t , and its own realized output \mathbf{q}_{jt} . With its location choice, the MNE also chooses its optimal capital stock vector \mathbf{k}_{jt} across L locations.

On the second stage at time t , all uncertainty is removed and MNE j chooses output \mathbf{q}_{jt} given its cost function (or, by duality, optimal factor employment given its production function). The optimal quantity choice \mathbf{q}_{jt}^* at time t can be characterized with the first-order conditions

$$p^\ell(\mathbf{q}_{i \neq j}, \mathbf{q}_{jt}^*) \left(1 - 1/\varepsilon_{jt}^{q^\ell}\right) \leq \frac{\partial C_{jt}(\mathbf{q}_{jt}^*; \mathbf{k}_{jt}, \mathbf{w}_t)}{\partial q_{jt}^\ell} \quad (\ell = 1, \dots, L), \quad (\text{A1})$$

where $p^\ell(\cdot)$ is the price of a good from location ℓ as a function of competitors' and own worldwide output, and $\varepsilon_{jt}^{q^\ell}$ is the elasticity of demand for q_{jt}^ℓ with respect to price p^ℓ . By the Kuhn-Tucker theorem, $\mathbf{q}_{jt}^* = 0$ if inequality holds. So, even if MNE j is present at location ℓ , it may find it optimal to produce $q_{jt}^\ell = 0$ once factor price and competitors' output are revealed.

On the first stage, MNE j 's linear programming problem can be characterized by the rules for FDI at locations $\ell = 1, \dots, L$

$$d_{jt}^\ell = \mathbf{1} \left(\mathbb{E}_{j, t-\tau} \left[\Pi_{jt}(q_{jt}^{\ell, *}) - \Pi_{jt}(q_{jt}^\ell = 0) \mid \mathbf{z}_{j, t-\tau} \right] - F_{j, t-\tau}^\ell + \eta_{jt}^\ell > 0 \right), \quad (\text{A2})$$

where $F_{j, t-\tau}^\ell$ denotes MNE j 's relevant fixed costs for presence at location ℓ and η_{jt}^ℓ is an MNE-specific disturbance. Expectations depend on MNE j 's information set $\mathbf{z}_{j, t-\tau}$. MNE j 's linear programming problem on the first stage involves the simultaneous evaluation of (A2) for each location ℓ given the 2^{L-1} possible combinations of outputs at all remaining locations $L-1$.

For its location choice on the first stage, an MNE j maximizes its expected profits $\mathbb{E}_{j, t-\tau}[\Pi_{jt}]$ where expectations are conditional on the MNE's information set in period $t - \tau$. The MNE can produce the vector of outputs $\mathbf{q}_{jt} = (q_{jt}^1, \dots, q_{jt}^L)'$ at L locations ($\ell = 1, \dots, L$). So, future expected profits are

$$\mathbb{E}_{j, t-\tau} [\mathbf{p}(\mathbf{q}_{i \neq j, t}, \mathbf{q}_{jt})' \cdot \mathbf{q}_{jt} - C_{jt}(\mathbf{q}_{jt}; \mathbf{k}_{jt}, \mathbf{w}_t)]. \quad (\text{A3})$$

The estimated presence rule (5) in the text follows using expected profits (A3) in criterion (A2).

B Multiproduct translog cost function

Consider the short-run multiproduct translog function with quasi-fixed capital:¹⁵

$$\begin{aligned}
\ln C_{jt} &= \varphi + \sum_{m=1}^L \varphi_m^0 \ln q_{jt}^m + \sum_{\ell=1}^L \alpha_\ell \ln w_t^\ell + \sum_{m=1}^L \sum_{\ell=1}^L \mu_{\ell m} \ln q_{jt}^m \ln w_t^\ell \\
&+ \frac{1}{2} \sum_{m=1}^L \sum_{\ell=1}^L \varphi_{\ell m}^1 \ln q_{jt}^m \ln q_{jt}^\ell + \frac{1}{2} \sum_{m=1}^L \sum_{\ell=1}^L \delta_{\ell m} \ln w_t^m \ln w_t^\ell \\
&+ \sum_{m=1}^L \zeta_m^0 \ln k_{jt}^m + \sum_{m=1}^L \sum_{\ell=1}^L \zeta_{\ell m}^{11} \ln k_{jt}^m \ln q_{jt}^\ell \\
&+ \sum_{m=1}^L \sum_{\ell=1}^L \kappa_{\ell m} \ln k_{jt}^m \ln w_t^\ell + \frac{1}{2} \sum_{m=1}^L \sum_{\ell=1}^L \zeta_{\ell m}^1 \ln k_{jt}^m \ln k_{jt}^\ell.
\end{aligned} \tag{B1}$$

By Shepard's (1953) lemma, MNE j 's demand for employment y_{jt}^ℓ is equal to $\partial C_{jt}/\partial w_t^\ell$ so that the share $s_{jt}^\ell \equiv w_t^\ell y_{jt}^\ell / C_{jt}$ of the wage bill at location ℓ in the MNE's total wage bill becomes

$$s_{jt}^\ell = \frac{\partial C_{jt}/\partial w_t^\ell}{C_{jt}/w_t^\ell} = \alpha_\ell + \sum_{m=1}^L (\mu_{\ell m} \ln q_{jt}^m + \kappa_{\ell m} \ln k_{jt}^m + \delta_{\ell m} \ln w_t^m)$$

for $\ell = 1, \dots, L$. We transform these L equations into L simultaneous labor demand functions by multiplying the dependent variable and all regressors with the observation-specific scalars C_{jt}/w_t^ℓ and obtain $y_{jt}^\ell = \partial C_{jt}/\partial w_t^\ell = s_{jt}^\ell C_{jt}/w_t^\ell$ as in equation (6).

With L locations, there are $L(L-1)/2$ symmetry restrictions $\delta_{k\ell} = \delta_{\ell k}$ for any k, ℓ . Linear homogeneity in factor prices requires that $\sum_{\ell=1}^L \alpha_\ell = 1$ and that $\sum_{\ell=1}^L \mu_{\ell m} = \sum_{\ell=1}^L \kappa_{\ell m} = \sum_{\ell=1}^L \delta_{\ell m} = \sum_{\ell=1}^L \delta_{m\ell} = 0$ for all m . We impose those restrictions on estimation but do not constrain estimates of factor price coefficients at the extensive margin. We do not impose any returns-to-scale restrictions.

C Currency conversion and deflation

We convert all economic data of foreign affiliates into euro (EUR) and deflate them. In BuBa's original MIDI data, all information on foreign affiliates is reported in German currency using the exchange rate at the closing date of the foreign affiliate's

¹⁵Slaughter (2000) adds $\ln(k/q)$ terms to a version of (B1). Given the additive logarithmic structure, this is equivalent to an affine transformation of the parameter pairs (α_k, ζ_k) and $(\mu_{k,\ell}, \kappa_{k,\ell})$ because $\ln(k/q) = \ln k - \ln q$.

balance sheet. We apply the following deflation and currency conversion method to all financial variables. Deutschmark (DEM) figures are converted into euro figures at the rate 1/1.95583 (the conversion rate at inception of the euro in 1999). (i) We use the market exchange rate on the end-of-month day closest to an affiliate’s balance sheet closing date to convert the DEM or EUR figures into local currency for every affiliate. This reverses the conversion applied to the questionnaires at the date of reporting. (ii) A CPI factor for every country deflates the foreign-currency financial figures to the December-1998 real value in local currency. (iii) For each country, the average of all end-of-month exchange rates vis-à-vis the DEM or EUR between January 1996 and December 2001 is used as a proxy for purchasing power parity of foreign consumption baskets relative to the DEM or EUR. All deflated local-currency figures are converted back to DEM or EUR using this purchasing-power proxy.

We use the foreign countries’ CPIs (Consumer Price Indices from the IMF’s International Financial Statistics) to deflate the figures. Whenever a country’s CPI is not available from IFS but the main currency used in that country is issued in some other country, we use the CPI of the currency-issuing country. The CPI deflation factors for all countries are rebased to unity at year-end 1998. For the UBS wage data, we first translate U.S. dollars into Euros and then proceeded as detailed above. Parent-level and sector-level domestic variables are transformed into December 1998 Euros using the German CPI.

D Wages

The German annual earnings survey provides representative monthly wages for employees, broken down into three blue-collar and four white-collar occupation groups by sector (two-digit NACE 1.1). We compute median wages over these seven occupation groups by sector, deflate them with the German CPI (standardized to unity in December 1998), and multiply them by twelve to arrive at annual earnings for cost function estimation.

The OWW data contain monthly average wage rates of male workers at the country level for 161 occupations in over 150 countries between 1983 and 1999. We follow Freeman and Oostendorp’s (2001) recommendation and pick the base calibration with lexicographic weighting for the aggregate wages by country. We deflate the wages with the country-level CPI (standardized to unity in December 1998) and convert the foreign currency to EUR at the December 1998 exchange rate. We fill missing values, by country and occupation group, with information from the latest preceding year that has wage information available. We compute the median wage across all reported occupation groups by country and multiply the monthly wages by twelve to approximate annual earnings for cost function estimation. We reuse

OWW wages from 1999 in 2000 and 2001.

For robustness checks, we use UBS wage data collected by the Swiss commercial bank for metropolitan areas around the world in 1994, 1997, 2000 and 2003 (UBS 2003). UBS carried out surveys in approximately 70 cities during the second quarter of 1994, 1997 and 2000, and during the first quarter of 2003. Questionnaires request detailed information on wage components, wage deductions and working hours across thirteen occupations. UBS converts wage figures into U.S. dollars and smoothes the effect of day-to-day currency fluctuations by using the average daily spot rate during the quarter of the UBS survey. We convert UBS wages into EUR at the average USD/EUR exchange rate during the survey quarter and deflate figures with the German CPI (standardized to unity in December 1998). We linearly interpolate UBS wages between survey years to cover our sample period 1996-2001. We use the machinist wage as the most closely comparable wage to median OWW and German wages.

E Market access

We construct market access measures following Redding and Venables (2004). We obtain bilateral trade data for 1996 through 2001 and geographic information on country pairs from CEPII (www.cepii.fr). After filling in missing imports to B from A with exports information from country A to B , we drop all exports information and set exports from A to B equal to B 's imports from A . We adopt this procedure because we consider imports, whenever available, more reliably measured than exports.

Our regression specification for an unbalanced panel of country pairs by year is

$$\ln X_{ij} = \alpha_i x_i + \beta_j m_j + \delta \ln d_{ij} + \mu b_{ij} + \epsilon_{ij},$$

where X_{ij} denotes country i 's aggregate exports in USD (+1) to country j , x_i an exporter country dummy, m_j an importer country dummy, d_{ij} the geographical distance between country i and j , and b_{ij} a dummy variable indicating a common border. We compute market access A_i to country i as

$$A_i = \exp\{\beta_i m_i\} \left(.67 \sqrt{\text{area}_i / \pi} \right)^{\delta/2} + \sum_{j \neq i} \exp\{\beta_j m_j\} (d_{ij})^\delta \exp\{\mu b_{ij}\}.$$

This is measure MA(3) in Redding and Venables (2004).

Table 15: LOCATION DEFINITIONS

Locations	Countries
WEU	Western European countries (EU 15 plus Norway and Switzerland)
OIN	Overseas Industrialized countries including Australia, Canada, Japan, New Zealand, USA as well as Iceland and Greenland
CEE	Central and Eastern European countries including accession countries and candidates for EU membership
DEV	Developing countries including Turkey, Russia and Central Asian economies as well as dominions of Western European countries and the USA

References

- BARBA NAVARETTI, G., AND D. CASTELLANI (2004): “Investments Abroad and Performance at Home: Evidence from Italian Multinationals,” *CEPR Discussion Paper*, 4284.
- BARRO, R. J., AND J. W. LEE (2001): “International Data on Educational Attainment: Updates and Implications,” *Oxford Economic Papers*, 53(3), 541–63.
- BARTEN, A. P. (1969): “Maximum Likelihood Estimation of a Complete System of Demand Equations,” *European Economic Review*, 1(1), 7–73.
- BECKER, S. O., K. EKHOLM, R. JÄCKLE, AND M.-A. MUENDLER (2005): “Location Choice and Employment Decisions: A Comparison of German and Swedish Multinationals,” *Review of World Economics*, 141(4), 693–731.
- BLONIGEN, B. A., R. B. DAVIES, AND K. HEAD (2003): “Estimating the Knowledge-Capital Model of the Multinational Enterprise: Comment,” *American Economic Review*, 93(3), 980–94.
- BRACONIER, H., AND K. EKHOLM (2000): “Swedish Multinationals and Competition from High- and Low-Wage Locations,” *Review of International Economics*, 8(3), 448–61.
- BRAINARD, S. L. (1997): “An Empirical Assessment of the Proximity-Concentration Trade-off between Multinational Sales and Trade,” *American Economic Review*, 87(4), 520–44.
- BRAINARD, S. L., AND D. A. RIKER (2001): “Are US Multinationals Exporting US Jobs?,” in *Globalization and labour markets*, ed. by D. Greenaway, and D. R. Nelson, vol. 2, pp. 410–26. Elgar, Cheltenham, UK and Northampton, MA.
- BROWN, R. S., AND L. R. CHRISTENSEN (1981): “Estimates of Elasticities of Substitution in a Model of Partial Static Equilibrium: An Application to U.S. Agriculture, 1947-1974,” in *Modeling and measuring natural resource substitution*, ed. by E. R. Berndt, and B. C. Field, chap. 10, pp. 209–29. MIT Press, Cambridge, Massachusetts.
- BUCH, C. M., J. KLEINERT, A. LIPPONER, AND F. TOUBAL (2005): “Determinants and Effects of Foreign Direct Investment: Evidence from German Firm-Level Data,” *Economic Policy*, 41, 51–98.
- CARR, D. L., J. R. MARKUSEN, AND K. E. MASKUS (2001): “Estimating the Knowledge-Capital Model of the Multinational Enterprise,” *American Economic Review*, 91(3), 693–708.

- CHEN, S., AND S. KHAN (2003): “Semiparametric Estimation of a Heteroskedastic Sample Selection Model,” *Econometric Theory*, 19(6), 1040–64.
- CHRISTENSEN, L. R., D. W. JORGENSON, AND L. J. LAU (1973): “Transcendental Logarithmic Production Frontiers,” *Review of Economics and Statistics*, 55(1), 28–45.
- DAS, M., W. K. NEWEY, AND F. VELLA (2003): “Nonparametric Estimation of Sample Selection Models,” *Review of Economic Studies*, 70(1), 33–58.
- DEUTSCHE BUNDESBANK (1998): “The Methodological Basis of the Deutsche Bundesbank’s Corporate Balance Sheet Statistics,” *Monthly Report*, 1998(10), 49–64.
- DEVEREUX, M. P., AND R. GRIFFITH (1998): “Taxes and the Location of Production: Evidence from a Panel of US Multinationals,” *Journal of Public Economics*, 68(3), 335–67.
- DHRYMES, P. J. (1971): “Equivalence of Iterative Aitken and Maximum Likelihood Estimators for a System of Regression Equations,” *Australian Economic Papers*, 10(16), 20–24.
- DI EWERT, W. E. (1971): “An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function,” *Journal of Political Economy*, 79(3), 481–507.
- DISDIER, A. C., AND T. MAYER (2004): “How Different Is Eastern Europe? Structure and Determinants of Location Choices by French Firms in Eastern and Western Europe,” *Journal of Comparative Economics*, 32(2), 280–96.
- DIXIT, A. K. (1989): “Entry and Exit Decisions under Uncertainty,” *Journal of Political Economy*, 97(3), 620–38.
- EAKIN, B. K., D. P. MCMILLEN, AND M. J. BUONO (1990): “Constructing Confidence Intervals Using the Bootstrap: An Application to a Multi-Product Cost Function,” *Review of Economics and Statistics*, 72(2), 339–44.
- FEENSTRA, R. C., AND G. H. HANSON (1999): “The Impact of Outsourcing and High-Technology Capital on Wages: Estimates for the United States, 1979–1990,” *Quarterly Journal of Economics*, 114(3), 907–40.
- FREEMAN, R. B., AND R. H. OOSTENDORP (2001): “The Occupational Wages around the World Data File,” *International Labour Review*, 140(4), 379–401.

- HALL, R. E. (1973): “The Specification of Technology with Several Kinds of Output,” *Journal of Political Economy*, 81(4), 878–92.
- HANSON, G. H., R. J. MATALONI, AND M. J. SLAUGHTER (2005): “Vertical Production Networks in Multinational Firms,” *Review of Economics and Statistics*, 87(4), 664–78.
- HEAD, K., AND T. MAYER (2004): “Market Potential and the Location of Japanese Investment in the European Union,” *Review of Economics and Statistics*, 86(4), 959–72.
- HEAD, K., AND J. RIES (2002): “Offshore Production and Skill Upgrading by Japanese Manufacturing Firms,” *Journal of International Economics*, 58(1), 81–105.
- HECKMAN, J. J. (1979): “Sample Selection Bias as a Specification Error,” *Econometrica*, 47(1), 153–61.
- (1981): “The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process,” in *The structural analysis of discrete data*, ed. by C. Manski, and D. F. McFadden, pp. 179–95. MIT Press, Cambridge, MA.
- KOHLI, U. R. (1978): “A Gross National Product Function and the Derived Demand for Imports and Supply of Exports,” *Canadian Journal of Economics*, 11(2), 167–82.
- KONINGS, J. (2004): “The Employment Effects of Foreign Direct Investment,” *EIB Papers*, 9(1), 86–108.
- KOTZ, S., N. BALAKRISHNAN, AND N. L. JOHNSON (2000): *Continuous Multivariate Distributions*, vol. 1 of *Wiley Series in Probability and Statistics*. Wiley, New York, 2nd edn.
- LIPPONER, A. (2003): “A “New” Micro Database for German FDI,” in *Foreign Direct Investment in the Real and Financial Sector of Industrial Countries*, ed. by H. Herrmann, and R. Lipsey, pp. 215–44. Springer, Berlin.
- MARIN, D. (forthcoming): “A New Division of Labor in Europe: Offshoring and Outsourcing into Eastern Europe,” *Journal of the European Economic Association*.
- NEWKEY, W. K. (1997): “Convergence Rates and Asymptotic Normality for Series Estimators,” *Journal of Econometrics*, 79(1), 147–68.

- REDDING, S., AND A. J. VENABLES (2004): “Economic Geography and International Inequality,” *Journal of International Economics*, 62(1), 53–82.
- ROBERTS, M. J., AND J. R. TYBOUT (1997): “The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs,” *American Economic Review*, 87(4), 545–64.
- SHEPARD, R. W. (1953): *Cost and production functions*. Princeton University Press, Princeton.
- SLAUGHTER, M. J. (2000): “Production Transfer within Multinational Enterprises and American Wages,” *Journal of International Economics*, 50(2), 449–72.
- UBS (2003): “Prices and Earnings: A comparison of purchasing power around the world,” *ubs.com* Zürich.
- YEN, S. T. (2005): “A Multivariate Sample-Selection Model: Estimating Cigarette and Alcohol Demands with Zero Observations,” *American Journal of Agricultural Economics*, 87(2), 453–66.
- ZELLNER, A. (1962): “An Efficient Method for Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias,” *Journal of the American Statistical Association*, 57(298), 348–68.