

# Margins of Multinational Labor Substitution

By MARC-ANDREAS MUENDLER AND SASCHA O. BECKER

## Online Appendix

### I. Global Production

#### A. Multiproduct translog cost function

Consider the short-run multiproduct translog function with quasi-fixed capital:<sup>30</sup>

$$\begin{aligned}
 \ln C_{jt} &= \varphi + \sum_{n=1}^L \varphi_n^0 \ln q_{jt}^n + \sum_{\ell=1}^L \alpha_\ell \ln w_t^\ell + \sum_{n=1}^L \sum_{\ell=1}^L \mu_{\ell n} \ln q_{jt}^n \ln w_t^\ell \\
 &\quad + \frac{1}{2} \sum_{n=1}^L \sum_{\ell=1}^L \varphi_{\ell n}^1 \ln q_{jt}^n \ln q_{jt}^\ell + \frac{1}{2} \sum_{n=1}^L \sum_{\ell=1}^L \delta_{\ell n} \ln w_t^n \ln w_t^\ell \\
 (A1) \quad &\quad + \sum_{n=1}^L \zeta_n^0 \ln k_{jt}^n + \sum_{n=1}^L \sum_{\ell=1}^L \zeta_{\ell n}^{11} \ln k_{jt}^n \ln q_{jt}^\ell \\
 &\quad + \sum_{n=1}^L \sum_{\ell=1}^L \kappa_{\ell n} \ln k_{jt}^n \ln w_t^\ell + \frac{1}{2} \sum_{n=1}^L \sum_{\ell=1}^L \zeta_{\ell n}^1 \ln k_{jt}^n \ln k_{jt}^\ell.
 \end{aligned}$$

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

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

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  $C_{jt} / w_t^\ell$  and obtain  $y_{jt}^\ell = \partial C_{jt} / \partial w_t^\ell = s_{jt}^\ell C_{jt} / w_t^\ell$  as in eq. (1).

With  $L$  locations, there are  $L(L-1)/2$  symmetry restrictions  $\delta_{k\ell} = \delta_{\ell k}$  for any  $k, \ell$ . Linear homogeneity in factor prices requires that  $\sum_{\ell=1}^L \alpha_\ell = 1$  and that  $\sum_{\ell=1}^L \mu_{\ell n} = \sum_{\ell=1}^L \kappa_{\ell n} = \sum_{\ell=1}^L \delta_{\ell n} = \sum_{\ell=1}^L \delta_{n\ell} = 0$  for all  $n$ . We impose these restrictions on intensive-margin estimation but do not constrain extensive-margin coefficients.

#### B. Stacking

Eq. (1) requires treatment for locations of absence because outputs and capital inputs are missing where MNEs do not operate. Our maintained assumptions imply that stacking of observations is a viable and attractive procedure.<sup>31</sup> Stacking means that we set regressors for locations of absence to zero. Stacking is easily implemented, improves efficiency, collapses the up to  $2^{L-1} - 1$

<sup>30</sup>Slaughter (2000) adds  $\ln(k/q)$  terms to a version of (A1). Given the additive logarithmic structure, this is equivalent to an affine transformation of the parameter pairs  $(\alpha_k, \zeta_k)$  and  $(\mu_{k,\ell}, \kappa_{k,\ell})$  because  $\ln(k/q) = \ln k - \ln q$ .

<sup>31</sup>Estimation of separate equation systems for all possible presence patterns is plagued by dimensionality: potential presence in up to  $L-1$  locations outside home means that there are up to  $2^{L-1} - 1$  regional presence patterns. Lung Fei

sets of estimates into one consistently estimated  $(L-1)$ -equation system, and provides a single  $L \times L$  matrix of estimates for wage elasticities of regional labor demands.

More formally, stacking interacts the parameters in (1) with presence indicators:  $\mu_{\ell n} = 0$  when no output is produced at location  $n$ , and  $\kappa_{\ell n} = \delta_{\ell n} = 0$  when MNE  $j$  employs no factors at location  $n$ . Stacking is permissible under three natural assumptions in our framework: (i) all MNEs face the same sunk cost function  $F_{j,t-\tau}^{\ell}$  conditional on prior presence (so that presence is mean independent of inputs); (ii) MNEs face an identical short-run cost function  $C(\cdot)$  in all locations of presence (but not necessarily where absent) conditional on characteristics (so that a common parameter vector is justified); and (iii) the disturbances  $\epsilon_{jt}^{\ell}$  are uncorrelated across observations of MNEs  $i$  and  $j$ . To prevent any bias from stacking, we include a set of absence indicators  $(\mathbf{1} - \mathbf{d}_{jt}^{n \neq \ell})$  in the outcome equation. Absence indicators control for shadow inputs. To check robustness of the stacking procedure, we repeat estimation for the subsample of omnipresent MNEs that operate affiliates in all locations.

## II. Selection Correction

### A. Parametric selection correction

Given our parametric cost function, a parametric approach to selectivity is a natural benchmark. Plausible distributional assumptions permit individual Heckman (1979) corrections location by location.<sup>32</sup> Consider linear selection predictions  $H(\mathbf{z}_{j,t-\tau}) = \mathbf{z}_{j,t-\tau} \gamma^{\ell}$  and jointly normally distributed disturbances  $(\epsilon_{jt}^k, \eta_{j,t-\tau}^{\ell})$  so that a probit model describes the choice of presence (6).

The correlation between  $\epsilon_{jt}^n$  and  $\eta_{j,t-\tau}^{\ell}$  across separate locations  $n \neq \ell$  is crucial for estimation of outcomes (2). Our data reject independence of  $\epsilon_{jt}^n$  and  $\eta_{j,t-\tau}^{\ell}$ .<sup>33</sup> To specify the correlation structure, we depart from the idea that selection disturbances include both location-specific parts such as, for example, surprising changes to profit repatriation policies in the host country and include MNE-specific parts such as idiosyncratic shocks to a firm's sunk entry costs. Changes to host-country repatriation policies affect the entry decision. But once the MNE operates in the host country, it minimizes costs irrespective of entry-related host-country shocks. So, we consider it plausible to assume that there is an MNE-specific, location-independent component  $e_{jt}$  to the selection shock  $\eta_{j,t-\tau}^n$  and that the labor-demand shock  $\epsilon_{jt}^{\ell}$  correlates with the selection shock  $\eta_{j,t-\tau}^n$  elsewhere only through the MNE-specific component  $e_{jt}$ . The assumption is not rejected in our data. Note that, under this assumption, cost function disturbances do covary with entry shocks across locations, but only through an MNE-specific component.

**ASSUMPTION A:** *The disturbances  $\epsilon_{jt}^n$  and  $\eta_{j,t-\tau}^{\ell}$  are multivariate normally distributed with  $\epsilon_{jt}^n = \lambda e_{jt} + \pi_{\epsilon}^n v_{jt}^n$  and  $\eta_{j,t-\tau}^{\ell} = \sqrt{1-\nu} e_{jt} + \sqrt{\nu} u_{jt}^{\ell}$ , where  $\nu \in [0, 1]$  and the standard normal variables  $e_{jt}, u_{jt}^{\ell}, v_{jt}^n$  are independent of  $\mathbf{x}_{jt}^m$  and  $\mathbf{z}_{j,t-\tau}$  for all  $\ell, m, n$ .*

Any normally distributed random variable can be decomposed into an affine function of standard normal variables. Assumption A does this. Under Assumption A, the variances and covariances of the selection shocks are  $\sigma_{\eta}^{\ell\ell} = 1$ , as is common for probit, and  $\sigma_{\eta}^{n\ell} = 1 - \nu$ . The variances

Lee and Mark M. Pitt (1986) propose an estimator related to J. Peter Neary and Kevin W. S. Roberts's (1980) shadow price approach. Bertrand M. Koebel (2006) conducts Box-Cox transformations on inputs.

<sup>32</sup>For multivariate selectivity, an extension of the univariate Heckman (1979) estimator has a complicated form (conditional moments of multivariate normal distributions have no known closed form for multiple truncations, see Samuel Kotz, N. Balakrishnan and Norman L. Johnson (2000)). Simulated maximum-likelihood would be a viable technique but requires joint multivariate normality, which we prefer to relax in nonparametric estimation.

<sup>33</sup>SUR estimation of the outcome equations shows that  $\epsilon_{jt}^n$  and  $\epsilon_{jt}^{\ell}$  correlate so that  $\epsilon_{jt}^n$  and  $\eta_{j,t-\tau}^{\ell}$  must be correlated because  $\epsilon_{jt}^{\ell}$  and  $\eta_{j,t-\tau}^{\ell}$  are correlated.

and covariances of the labor demand shocks are  $\sigma_\epsilon^{\ell\ell} = \lambda^2 + (\pi_\epsilon^{\ell\ell})^2$  and  $\sigma_\epsilon^{n\ell} = \lambda^2$ . And the covariances between the selection shock in location  $n$  and the demand shock in location  $\ell$  are  $\sigma_{\eta\epsilon}^{n\ell} = \lambda$ . So, cost function disturbances do correlate with entry-relevant policy shocks across locations, but only through an MNE-specific shock. The assumption accommodates potential serial correlation in location selection, defining  $u_{jt}^\ell \equiv \sum_\varsigma \alpha_\varsigma^u \tilde{u}_{j,t-\varsigma}^\ell$ . Assumption A is testable. We obtain estimates of  $\sigma_\eta^{n\ell} = 1 - \nu$  from multivariate probit estimation (on the same set of regressors as in Table 5) and use a  $\chi^2$ -test for their equality. We fail to reject equality.

Intuitively, all selection-related information that is relevant for labor demand at any location  $\ell$  is fully contained in the single presence indicator  $d_{jt}^\ell$ , which is as informative about  $\eta_{j,t-\tau}^\ell$  as any other location indicator. So, location-by-location correction for selectivity is permissible.

**LEMMA 1:** *If Assumption A holds, independent selection correction for  $L$  locations identifies  $\mathbf{x}_{jt}^\ell \beta^\ell$  and  $m^\ell(\Pr^\ell(\mathbf{z}_{j,t-\tau}))$ .*

**PROOF:**

Denote the standard normal density and distribution functions with  $\phi(\cdot)$  and  $\Phi(\cdot)$ . Under Assumption A, the marginal likelihood function is

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

after concentrating out  $u_{jt}^\ell$  and  $v_{jt}^\ell$ , where  $\sigma_\epsilon^\ell = \sqrt{\sigma_\epsilon^{\ell\ell}} = \sqrt{\lambda^2 + (\pi_\epsilon^{\ell\ell})^2}$  and  $\rho_{\eta\epsilon}^{\ell\ell} = \sigma_{\eta\epsilon}^{\ell\ell} / \sigma_\epsilon^\ell = \lambda / \sqrt{\lambda^2 + (\pi_\epsilon^{\ell\ell})^2}$ . This is the likelihood function for independent Heckman (1979) correction location by location, where  $m^\ell(\Pr^\ell(\mathbf{z}_{j,t-\tau})) = \beta_\Lambda^\ell \Lambda_{jt}^\ell(\mathbf{z}_{j,t-\tau} \gamma^\ell)$  and  $\beta_\Lambda^\ell = \rho_{\eta\epsilon}^{\ell\ell} \sigma_\epsilon^\ell$  is the coefficient on the selectivity hazard  $\Lambda_{jt}^\ell(\mathbf{z}_{j,t-\tau} \gamma^\ell)$  (the inverse of the Mills ratio) in the outcome equation. ■

Under Heckman (1979) correction (Assumption A), the extensive-margin term in (5) simplifies to  $\beta_\Lambda^\ell \Delta_{jt}^\ell \cdot \gamma_{w^n}^\ell \cdot w_t^\ell w_t^n / C$ , where  $\gamma_{w^n}^\ell$  is the wage coefficient in the selection equation,  $\beta_\Lambda^\ell$  is the coefficient on the selectivity hazard in the outcome equation, and  $\Delta_{jt}^\ell$  is the first derivative of the selectivity hazard  $\Lambda_{jt}^\ell$  (the inverse of the Mills ratio) with respect to its 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  $\gamma_{w^n}^\ell \beta_\Lambda^\ell$  (the coefficients on the two stages of estimation).

### B. Nonparametric selection correction

To establish identification, consider the following deviations from the truth:  $\Delta \xi^\ell(\mathbf{x}_{jt}^\ell) \equiv \mathbf{x}_{jt}^\ell (\hat{\beta}^\ell - \beta^\ell)$  and  $\Delta m^\ell(\mathbf{P}_{jt}) \equiv \hat{m}^\ell(\mathbf{P}_{jt}) - m^\ell(\mathbf{P}_{jt})$ , where hats denote estimates of the true (not hatted) functions.

Assumption B formally states one set of sufficient conditions for identification.

**ASSUMPTION B:**

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

(iii)  $\nabla_{\mathbf{z}_{j,t-\tau}} \mathbf{P}_{jt} \neq \mathbf{0}$  with probability one,

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

Part (i) posits that the conditional expectation of the labor demand disturbance at location  $\ell$  is a function of the propensity scores of presence at any location  $k = 1, \dots, L$ . So, in the regression of observed labor demand  $y_{jt}^\ell$  on  $\mathbf{x}_{jt}^\ell \beta^\ell$  and  $m^\ell(\mathbf{P}_{jt})$ ,  $\mathbf{x}_{jt}^\ell \beta^\ell$  is a separate additive component. This specification applies nonparametric selectivity correction with a single outcome equation (but multiple selection thresholds) in Das, Newey and Vella (2003) to the multivariate outcome case.<sup>34</sup> The generalization to simultaneous location selection (multivariate selectivity) comes at a price. To maintain identifying restrictions similar to Das, Newey and Vella (2003), we need to assume cross-equation independence in the selection disturbance conditional on observable variables.

Part (ii) is the same identification condition as in Das, Newey and Vella (2003) and implies that  $\mathbf{P}_{jt}$  (which enters  $m^\ell(\mathbf{P}_{jt})$ ) depends on variables in  $\mathbf{z}_{j,t-\tau}$  that are not in  $\mathbf{x}_{jt}^\ell \beta^\ell$ . Otherwise, a regression of  $y_{jt}^\ell$  on  $\mathbf{x}_{jt}^\ell \beta^\ell$  leaves  $\Delta \xi^\ell(\mathbf{x}_{jt}^\ell) = m^\ell(\mathbf{P}_{jt})$  and  $\Delta m^\ell(\mathbf{P}_{jt}) = -m^\ell(\mathbf{P}_{jt})$  indeterminate—a violation of (ii). In our context, parent-firm characteristics and competitor-level host-country characteristics are among the  $\mathbf{z}_{j,t-\tau}$  predictors of presence but not related to the labor-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.

**LEMMA 2:** *If Assumption B holds and if  $m^\ell(\mathbf{P}_{jt})$  and  $\mathbf{P}_{jt}(\mathbf{z}_{j,t-\tau})$  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.*

**PROOF:**

In any observationally equivalent model it must be the case that the observed outcome satisfies  $\mathbb{E}[y_{jt}^\ell | \mathbf{x}_{jt}^\ell, \mathbf{d}_{jt}, \mathbf{z}_{j,t-\tau}] = \mathbf{x}_{jt}^\ell \hat{\beta}^\ell + \hat{m}^\ell(\mathbf{P}_{jt})$  for some  $\mathbf{x}_{jt}^\ell \hat{\beta}^\ell$  and  $\hat{m}^\ell(\mathbf{P}_{jt})$ . Equivalently, deviations from the truth  $\Delta \xi^\ell(\mathbf{x}_{jt}^\ell) + \Delta m^\ell(\mathbf{P}_{jt}) = 0$ . This identity must be differentiable with respect to  $\mathbf{x}_{jt}^\ell$  and  $\mathbf{z}_{j,t-\tau}$  by continuous differentiability of  $m^\ell(\mathbf{P}_{jt})$  and  $\mathbf{P}_{jt}(\mathbf{z}_{j,t-\tau})$ . So,

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

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  $\nabla_{\mathbf{z}_{j,t-\tau}} \mathbf{P}_{jt} \neq \mathbf{0}$ , the second equation implies that  $\Delta m^\ell(\mathbf{P}_{jt}) = \hat{m}^\ell(\mathbf{P}_{jt}) - m^\ell(\mathbf{P}_{jt}) = c_2$  for a constant  $c_2$  and  $m^\ell(\mathbf{P}_{jt})$  is identified up to that constant. ■

Under nonparametric location selection (Assumption B) and polynomial series estimation, the derivatives of  $m^\ell(\cdot)$  and  $\Pr_{jt}^\ell$  at the extensive margin are the marginal effects on the polynomial terms  $\nabla_{\mathbf{P}_{jt}} m^\ell(\mathbf{P}_{jt}) \cdot \nabla_{w_{t-\tau}^n} \mathbf{P}_{jt} \cdot w_t^\ell w_t^n / C$ , which we evaluate at the sample mean.

<sup>34</sup>A semiparametric alternative would be the Lung Fei Lee (1995) estimator, a multivariate extension to Roger W. Klein and Richard H. Spady's (1993) semiparametric maximum-likelihood estimator. Lee (1995) partitions the covariates  $\mathbf{z}_{j,t-\tau}$  to appear in  $H(\mathbf{z}_{j,t-\tau})$  through multiple indexes. Note, however, that in our context the information set  $\mathbf{z}_{j,t-\tau}$  includes location selection predictors from every world region; so there is no natural subpartition. A nonparametric estimator for  $H(\mathbf{z}_{j,t-\tau})$  accommodates the multiple-index case and simultaneous selection into more than one location.

### III. Modelling Wage Endogeneity

Cross-wage elasticities (4) were derived in the context of competitive labor markets. MNEs, however, pay wage premia over local competitors. Suggested reasons include relatively skilled workforces and rent sharing through efficiency wages or bargaining.

Under wage bargaining, cross-wage elasticities (4) remain consistent with departures from competitive labor markets. Stole and Zwiebel (1996a; 1996b), for instance, consider bargaining between a firm and its individual workers, whose contracts cannot bind them to the firm. Their model relates bargaining outcomes to a firm's individual profitability and can explain within-industry wage differences between firms, such as mark-ups at MNEs relative to local competitors, if there are fixed hiring costs at wage-bargaining firms. A wage-bargaining firm's cost function does not necessarily exhibit first-degree homogeneity in paid wages. But, in line with our translog cost specification where we use location-wide median wages as outside wages, the wage-bargaining firm's cost function is homogeneous of degree one in location-specific reservation wages.

To establish homogeneity of degree in location wages, note that the first-order condition in Stole and Zwiebel (1996a; 1996b) for single-product firms requires that, at the optimal employment level  $\tilde{n}$ , realized profits are equal to average profits over all putative inframarginal workforce sizes

$$\pi(n, k) = p q(n, k) - \underline{w}n - \underline{r}k = (1/\tilde{n}) \int_0^{\tilde{n}} \pi(s, k) ds \equiv \tilde{\pi}(\tilde{n}, k),$$

where  $\underline{w}$  and  $\underline{r}$  are reservation factor prices. Since optimal profits  $\pi(\tilde{n}, k)$  are homogeneous of degree one in reservation prices by this first-order condition (an instance of the envelope theorem), the cost function is homogeneous of degree one in reservation wages. Similarly, Shepard's lemma holds for the reservation wage.

The consequences of wage bargaining for labor demand are theoretically ambiguous when contracts are non-binding (Stole and Zwiebel, 1996a; Asher Wolinsky, 2000; Catherine C. de Fontenay and Joshua S. Gans, 2003). To capture potential employment distortions, whatever their direction, we include MNE-specific wage residuals beyond the main location-specific median wages in our labor demand system.

We use the predicted log wage residual  $\psi_{jt}^{\ell}$  from a reduced form regression, mirroring the selection equation, to control for potential bias that could arise from omitting an MNE-specific wage premium or discount relative to the industry-wide median wage at the location:

$$\ln w_{jt}^{\ell} = W(\mathbf{z}_{j,t-\tau}) + \psi_{jt}^{\ell},$$

where  $w_{jt}^{\ell}$  is the MNE's paid wage at location  $j$ ,  $W(\cdot)$  mirrors the functional form of the location-selection equation (linear in the Heckman model), and  $\mathbf{z}_{j,t-\tau}$  is the vector of selection predictors. We include the set of predicted residual wages  $\psi_{jt}^{\ell}$  for all locations with an MNE's presence as additional regressors in outcome eq. (2) on the second stage. By construction, the log MNE-wage residuals are orthogonal to the propensity score. So any wage variation associated with the propensity score of presence is assigned to the extensive margin, as our selection model requires.

Including the estimated log wage residuals in labor-demand equations addresses concerns with profit-related pay and unobserved workforce heterogeneity. As suggested by firm-level wage bargaining described above, more productive MNEs may share rents with their workforces across locations and the MNEs' profitability may covary with industry-median wages at those locations. Alternatively, more productive MNEs may employ more skilled workforces, which can be associated with industry-median wages if the productivity dispersion is industry dependent. So, unobserved MNE productivity could bias our employment estimation unless firm-specific wage residuals are included alongside industry-wide wages. The firm-specific wage residuals serve as controls only. In line with the structural MNE model, estimation of cross-wage elasticities from

labor-demand outcomes exclusively rests on the industry-wide median wage coefficients ( $\delta_{\ell n}$ ).

## IV. Data

### A. Currency conversion and deflation

We deflate parent variables with the German consumer price index and deflate affiliate variables with country-level consumer price indices (from the IMF's International Financial Statistics).<sup>35</sup> CPI series are available for a broader set of countries than producer or wholesale price series. CPIs properly reflect the opportunity costs for investors who are the beneficiaries of firms' profit maximization. We re-base CPI deflation factors to unity at year end 1998 and 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.

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 balance sheet. Concretely, we apply the following conversion to all financial variables, including the physical capital stock (fixed assets). Deutschmark (DEM) figures are transformed into EUR at the rate 1/1.95583 (the conversion rate at euro inception 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.

### B. Wages

Our main estimation sample uses sectoral manufacturing wages by country between 1996 and 2001 from the UNIDO Industrial Statistics Database at the 3-digit ISIC level, Rev. 2 (UNIDO, 2005). The UNIDO measure of annual sectoral wage bills includes all payments to workers at establishments in the reference sector and year (wages and salaries, remuneration for time not worked, bonuses and gratuities, allowances, and payments in kind; but excludes contributions to social security, pensions, insurance, severance and termination pay). We divide the sectoral wage bill by the sectoral number of workers and employees. The UNIDO data cover 109 countries and result in the largest overlap with MIDI observations.

For robustness checks, we use wage data collected by the Swiss commercial bank UBS for metropolitan areas around the world in 1994, 1997, 2000 and 2003 (UBS, 2003). We linearly interpolate UBS wages between survey years to cover our sample period 1996-2001. UBS surveyed 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 use the machinist wage as the most closely comparable wage to German manufacturing wages (and to median OWW wages). We convert UBS wages into

<sup>35</sup>We use the CPI in the currency-issuing country whenever a country's CPI is not available from IFS but the main currency is issued elsewhere. We use current exchange rates and the German price deflator whenever foreign price deflators are missing or period-average exchange rate information is incomplete.

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 also use OWW monthly average wage rates of male workers at the country level for 161 occupations in 155 countries between 1983 and 1999. Missing observations, however, reduce the overlap with MIDI data below the overlap that UNIDO (or UBS) wage data provide. We follow Freeman and Oostendorp's (2001) recommendation and pick the base calibration with lexicographic weighting for the aggregate wages by country. We fill missing values, by country and occupation group, with information from the latest preceding year that has wage information available and reuse OWW wages from 1999 in 2000 and 2001. To mitigate workforce composition effects, we take country medians over 161 OWW occupation groups for foreign wages. We multiply the resulting monthly median occupation wage by twelve to approximate annual earnings for cost function estimation. Complementing foreign OWW wages, we use the German annual earnings survey (table 62321 from *destatis.de/genesis*) and obtain sectoral monthly wages, 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 and deflate figures with the German CPI (standardized to unity in December 1998). Occupational wage information from the German annual earnings survey enters the ILO database, on which OWW wages are based, so that these foreign and domestic wages are compatible.

### C. Complementary data

National accounts information for host-country regressors comes from the World Bank's World Development Indicators and the IMF's International Financial Statistics. To condition selection estimation on skill endowments beyond labor costs, we include the host country's percentage of highly educated residents in 1999 from Robert J. Barro and Jong Wha Lee (2001) and interact the variable with an indicator whether the percentage exceeds that in Germany (19.5%). We construct market access measures following Stephen Redding and Anthony J. Venables (2004), using their measure MA(3). To capture relevant cross-sectional variation, we compute competitor-level averages of the host-country characteristics MNE by MNE. Many host-country regressors are nevertheless statistically insignificant predictors in binary choice estimation, conditional on parent-level observable variables and host-country wages.

## V. Translog Estimates and Alternative Location Definitions

Table A.1 presents estimates of translog cost function equations for 1,654 stacked MNE observations between 1998 and 2001, as discussed in Section III-C. Beyond wages, specifications include turnover and fixed asset regressors, the scaled equivalent of the constant, and indicators of absence from all other locations. Estimates in the upper panel of Table A.1 include the predicted selectivity hazards (inverses of Mills ratios) by location (Assumption A). The lower panel presents estimates from nonparametric selectivity correction (Assumption B), using third-order polynomials in the predicted propensity scores for all locations.

Our definition of aggregate locations is motivated by geographical proximity and broad institutional similarity (Table A.2). As a robustness check, we split the world into the home country and four artificial regions defined by the quartiles of UNIDO manufacturing wages in the initial sample year 1996. Table A.3 reports estimated cross-wage elasticities for the wage-quartile groups of countries, as discussed in Section III-F.

TABLE A.1—TRANSLOG COST PARAMETER ESTIMATES

Employment in: <sup>a</sup>	CEE	DEV	OIN	WEU
	(1)	(2)	(3)	(4)
<b>Parametric Selectivity Correction (Assumption A)</b>				
ln Wages <sup>a</sup>				
HOM	.001 (.0006)	-.013 (.001)***	.027 (.008)***	.054 (.006)***
CEE	.001 (.0005)**	-.008 (.0001)***	.008 (.00004)***	-.002 (.00006)***
DEV	-.008 (.0003)***	.011 (.001)***	.009 (.0001)***	.0006 (.0001)***
OIN	.008 (.0004)***	.009 (.0008)***	-.086 (.008)***	.043 (.002)***
WEU	-.002 (.0005)***	.0006 (.0006)	.043 (.001)***	-.095 (.006)***
Selectivity hazard	12.058 (11.923)	24.432 (13.443)*	-19.821 (11.606)*	35.824 (14.625)**
R <sup>2</sup>	.977	.975	.969	.948
<b>Nonparametric Selectivity Correction (Assumption B)</b>				
ln Wages <sup>a</sup>				
HOM	.001 (.0006)**	-.008 (.001)***	.023 (.007)***	.059 (.006)***
CEE	-.0008 (.0005)	-.006 (.0003)***	.007 (.0004)***	-.002 (.0005)***
DEV	-.006 (.0003)***	.010 (.001)***	.007 (.0007)***	-.004 (.0007)***
OIN	.007 (.0004)***	.007 (.0007)***	-.079 (.008)***	.042 (.002)***
WEU	-.002 (.0005)***	-.004 (.0007)***	.042 (.002)***	-.096 (.005)***
Series terms				
χ <sup>2</sup> tests ( <i>p</i> -value)	495.52 (.000)	246.04 (.000)	151.17 (.000)	244.62 (.000)
R <sup>2</sup>	.979	.977	.974	.959

Sources: MIDI and USTAN 1996 to 2001 (UNIDO wages).

Notes: Stacked observations of 1,654 MNEs. Further regressors: ln Turnover, ln Fixed assets, ln MNE wage residuals, Absence indicators, Transformed constant (in parametric selectivity regression). Standard errors in parentheses: \* significance at ten, \*\* five, \*\*\* one percent. Standard errors corrected for first-stage estimation of selectivity hazards (hence not symmetric on restricted coefficients). Locations: HOM (Germany), CEE (Central and Eastern Europe), DEV (Developing countries), OIN (Overseas Industrialized countries), WEU (Western Europe).

<sup>a</sup>Transformed wage-bill shares and regressors.



TABLE A.2—AGGREGATE LOCATIONS

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 as well as Balkan countries, Belarus, Turkey, and Ukraine
DEV	Developing countries including Russia and Central Asian economies as well as dominions of Western European countries and of the USA

TABLE A.3—CROSS-WAGE ELASTICITIES BETWEEN WAGE QUARTILE GROUPS

Employment change (%) in	Wage change (by 1%) in				
	HOM (1)	Qrtl. 4 (2)	Qrtl. 3 (3)	Qrtl. 2 (4)	Qrtl. 1 (5)
HOM <i>intensive</i>	-.467**	.402**	.043*	.015*	-.001
Qrtl. 4 <i>intensive only</i>	1.193**	-1.339**	.104***	.025*	.009**
<i>extensive only</i>	.703***	-.763***	.030***	.019***	.004**
Qrtl. 3 <i>intensive only</i>	1.026*	.833***	-1.695***	-.190***	.018
<i>extensive only</i>	.703***	.237***	-.970***	.019***	.004**
Qrtl. 2 <i>intensive only</i>	.572*	.317*	-.297***	-.619**	.020
<i>extensive only</i>	.703***	.237***	.030***	-.981***	.004**
Qrtl. 1 <i>intensive only</i>	-.175	.561*	.134	.096	-.624
<i>extensive only</i>	.703***	.237***	.030***	.019***	-.996***

Sources: MIDI and USTAN 1996 to 2001 (UNIDO wages).

Notes: Elasticities at the extensive and intensive margins from 663 stacked MNE observations. Underlying labor demand estimates from parametric selectivity-corrected ISUR estimates (Assumption A). Standard errors from 200 bootstraps: \* significance at ten, \*\* five, \*\*\* one percent. Locations: HOM (Germany) and four foreign-country groups by manufacturing-wage quartiles, fourth quartile with top wages.

## REFERENCES

- Barro, Robert J., and Jong Wha Lee.** 2001. "International Data on Educational Attainment: Updates and Implications." *Oxford Economic Papers*, 53(3): 541–63.
- Das, Mitali, Whitney K. Newey, and Francis Vella.** 2003. "Nonparametric Estimation of Sample Selection Models." *Review of Economic Studies*, 70(1): 33–58.
- de Fontenay, Catherine C., and Joshua S. Gans.** 2003. "Organizational Design and Technology Choice under Intrafirm Bargaining: Comment." *American Economic Review*, 93(1): 448–55.
- Freeman, Richard B., and Remco H. Oostendorp.** 2001. "The Occupational Wages around the World Data File." *International Labour Review*, 140(4): 379–401.
- Heckman, James J.** 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1): 153–61.
- Klein, Roger W., and Richard H. Spady.** 1993. "An Efficient Semiparametric Estimator for Binary Response Models." *Econometrica*, 61(2): 387–421.
- Koebel, Bertrand M.** 2006. "Exports and Labor Demand: Searching for Functional Structure in Multi-Output Multi-Skill Technologies." *Journal of Business and Economic Statistics*, 24(1): 91–103.
- Kotz, Samuel, N. Balakrishnan, and Norman L. Johnson.** 2000. *Continuous Multivariate Distributions*. Vol. 1 of *Wiley Series in Probability and Statistics*. 2nd ed., New York: Wiley.
- Lee, Lung Fei.** 1995. "Semiparametric Maximum Likelihood Estimation of Polychotomous and Sequential Choice Models." *Journal of Econometrics*, 65(2): 381–428.
- Lee, Lung Fei, and Mark M. Pitt.** 1986. "Microeconomic Demand Systems with Binding Nonnegativity Constraints: The Dual Approach." *Econometrica*, 54(5): 1237–42.
- Neary, J. Peter, and Kevin W. S. Roberts.** 1980. "The Theory of Household Behaviour under Rationing." *European Economic Review*, 13(1): 25–42.
- Redding, Stephen, and Anthony J. Venables.** 2004. "Economic Geography and International Inequality." *Journal of International Economics*, 62(1): 53–82.
- Slaughter, Matthew J.** 2000. "Production Transfer within Multinational Enterprises and American Wages." *Journal of International Economics*, 50(2): 449–72.
- Stole, Lars A., and Jeffrey Zwiebel.** 1996a. "Intra-firm Bargaining under Non-binding Contracts." *Review of Economic Studies*, 63(3): 375–410.
- Stole, Lars A., and Jeffrey Zwiebel.** 1996b. "Organizational Design and Technology Choice under Intrafirm Bargaining." *American Economic Review*, 86(1): 195–222.
- UBS.** 2003. "Prices and Earnings: A Comparison of Purchasing Power Around the World." [ubs.com](http://ubs.com) Zürich.
- UNIDO.** 2005. "INDSTAT3 2005 ISIC Rev. 2 User's Guide." [unido.org/doc/3531](http://unido.org/doc/3531), Vienna.
- Wolinsky, Asher.** 2000. "A Theory of the Firm with Non-binding Employment Contracts." *Econometrica*, 68(4): 875–910.