Fixed Costs and FDI: The Conflicting Effects of Productivity Shocks*

Assaf Razin†, Yona Rubinstein‡and Efraim Sadka§

October, 2004

Abstract

The paper develops a model with lumpy setup costs of new investment, which govern the flows of FDI. Foreign investment decisions are two-fold: whether to export FDI and, if so, how much. The first decision is governed by total profitability considerations, whereas the second is governed by marginal profitability considerations. A positive productivity shock in the host country may, on the one hand, increases the volume of the desired FDI flows to the host country but, on the other hand, somewhat counter-intuitively, lowers the likelihood of the making new FDI flows by the source country, at all. Every country is potentially both a source for FDI flows to several host countries, and a host for FDI flows from several source

*We wish to thank Gita Gopinath, Pierre-Olivier Gourinchas, Galina Hale, Elhanan Helpman, Anil Kashyap, Michael Klein, Maury Obstfeld, Philip Lane, Jim Markusen, Helene Rey, John Romalis, Guido Tabellini and Stephen Ross Yeaple, for many useful comments and suggestions. We benefitted from seminars in the University of California - Berkeley, the University of Boconni, the University of Chicago, Trinity College (Dublin), Tel-Aviv University and the Hebrew University, and from comments of the NBER IFM program. We are indebted to Ashoka Mody, IMF, for developing the comprehensive data set that we use in this paper. We also thank Pavel Vacek for research assistance.

†Tel-Aviv University and Cornell University, CEPR, NBER and CESifo.

‡Tel-Aviv University and CEPR.

§Tel-Aviv University, CESifo and IZA.

countries. Thus, the model could generate two-way FDI flows, but not all source-host FDI flows get realized. We employ a sample of 24 OECD countries, over the period 1981-1998. We observe many pairs of countries with no FDI flows between them. Zero reported flows could indicate measurement errors, or true zeroes that are due to fixed costs (in situations where they dominate marginal productivity conditions). Empirical literature on the determinants of FDI flows which uses the Tobit procedure aims at a correction for measurement errors provides nevertheless biased estimates in the presence of fixed costs. By employing the Heckman selection procedure, we demonstrate how to get unbiased estimates of the fixed-costs effects on FDI flows. Indeed, the evidence suggests that the decisions on whether to export FDI at all, and on how much FDI cannot be regarded as a single decision on FDI flows, because the exogenous shocks are found to have conflicting effects on the likelihood of FDI flows and the magnitude of the flows.

1 Introduction

The paper develops an international capital flows' model, with lumpy set up costs of new investment which govern the flow of FDI.[1] The model works like this. First, a potential FDI investor decides how much she would like to invest. This decision is governed by marginal profitability considerations so as to equate marginal factor productivity to factor prices (that is, the standard first-order conditions). In the econometric terminology, this decision is described by a flow (gravity) equation. Second, because of fixed costs of new investments, the potential FDI investor must also decide wether to carry out at all new investments. This decision is governed by the total (rather than the marginal) profitability of the new investment. In the econometric terminology, the decision is described by a so-called participation equation. In the model, every country is potentially both a source

[1] The international trade literature appeals often to fixed costs. These costs play a very important role in determining the extent of trade-based foreign direct investment through the reallocation of capital across industries and the emergence of comparative advantages; see Zhang and Markusen (1999), Carr, Markusen and Maskum (2001), and Helpman, Melitz, and Yeaple (2004).

for FDI flows to several host countries, and a host for FDI flows from several source countries. But because of fixed costs, some of the source-host country pairs are inactive. In the presence of fixed costs, a productivity shock in the host country may also, on the one hand, increases the volume of the desired FDI flows to this country; as expected; but, on the other hand, and somewhat counter-intuitively, the shock lowers the likelihood of making new FDI flows at all, by the source country.

Our sample consists of 24 OECD countries over the period 1981-1998.[2] When one looks at data on international capital flows of FDI, one is immediately struck by the lack of flows from some source countries to many host countries. Only 17 countries are a source for FDI outflows, and each one of them exports FDI to only a few host countries. Thus there is a *prima facia* evidence for the existence of fixed setup costs of investment that shut off the potential of "small" capital flows, even though they may have been called for by marginal productivity conditions.

Previous empirical literature on the determinants of FDI flows frequently make use of the Tobit procedure. But this procedure, which is proper to handle measurement errors, reduces in essence, the flow and participation equations into just one equation. In contrast, by employing Heckman (1979) selection procedure, the two equations that are jointly estimated yield estimates that provide insight about the two equations separately. Indeed, the evidence suggests that the two equations cannot be combined, as in theory, exogenous shocks have conflicting effects on the likelihood of FDI flows and their magnitudes. Put it econometrically, the errors terms in the two equations are negatively correlated, and this implies that the Tobit procedure yields biased estimates. Controlling for the selection into source-host pairs of countries, and for time and country fixed effects, we find: (1) The magnitude of FDI flows respond positively to advances in the host country level of education (relative to the source country level of education). (2) The source-country level of education is a good predictor of the formation of source-host

[2]In Razin, Rubinstein and Sadka (2003) we employ a sample of 45 countries, both developed and developing countries. But the OECD data set is inacurate about the exports of FDI to non-OECD countries.

country pairs; (3) FDI flows respond positively to improvements in host country financial risk ratings, relative to the corresponding source country ratings; (3) The rich-rich FDI flow volumes in the OECD data set depend on the host-source income level differentials.

The organization of the paper is as follows. Section 2 presents our model of fixed setup costs of foreign direct investment. Section 3 presents the econometric approach. The data are described in Section 4. Estimation results of the determinants of FDI flows, and whether source-host flows are formed at all, are presented in Section 5. The results are interpreted in Section 6. Conclusions are drawn in Section 7.

2 Marginal Profitability Versus Total Profitability

We employ a "lumpy" adjustment cost for new investment, in the form of a fixed setup cost of investment.[3] This specification, which has been recently supported empirically by Caballero and Engel (1999, 2000), creates a situation in which FDI decisions become two-fold decisions: whether to export FDI at all, and, if so, how much. These decisions are pair-wise: that is, they are made by each source country with respect to each host country, as the "lumpy" adjustment cost is specific for each source-host pair. In our setup of exogenous shocks can affect these two decisions in opposite directions. That is, a shock can lower the likelihood of exporting FDI from a certain source country to a certain host country; but, if such an export is carried out, its magnitude is even higher.

Consider a representative industry in a given host country (H), in a world of free capital mobility which fixes the world rate of interest, denoted by r. There is a single good which serves both for consumption and investment. For simplicity, suppose that firms in this industry, all identical, last for two periods. The initial stock of capital is denoted by K_H^0. If the firm invests I in the first period, it augments its capital stock

[3] Evidently, this specification gives rise to economies of scale. Such economies either in the production or investment technologies are also a key contributor to the gains from trade and economic integration. For example, based on estimates taken from a partial equilibrium analysis, the Cecchini (1988) Report assessed that the gains from taking advantage of economies of scale will constitute about 30 percent of the total gains from the European market integration in 1992.

to $K_H^0 + I$ and its gross output in the second period will be $A_H F(K, L)$, where $F(\cdot)$ is a concave production function, L is the labor input, and A_H is a productivity factor.

We assume that there exists a fixed setup cost of investment. For simplicity, assume that this fixed cost is generated by a fixed input (L_H^C) of domestic labor. Thus, the fixed cost is equal to $W_H L_H^C$, where W_H is the wage rate in the host country. In order for the firm to be able to incur such a setup cost, we assume that, due to some (suppressed) fixed factor, $F(\cdot)$ exhibits diminishing returns to scale in K and L, that is F is strictly concave. Thus, the implied average cost curve is U-shaped, which is consistent with perfect competition that we assume.

Consider a representative firm which does invest in the first period an amount $I = K - K_H^0$ in order to augment its stock of capital to K. Its present value becomes

$$V^+(A_H, w_H, L_H^C) = \underset{\{K,L\}}{Max} \left\{ \frac{A_H F(K, L) - w_H L + K}{1 + r} - [(K - K_H^0) + w_H L_H^C] \right\} \qquad (1)$$

(For simplicity, it is assumed that capital does not depreciate).

The demand of such a firm for K and L are denoted by $K^+(A_H, W_H, L_H^C)$ and $L^+(A_H, W_H, L_H^C)$, respectively, they are defined by the marginal productivity conditions

$$A_H F_K(K, L0 = r \qquad (2)$$

and

$$A_H F_L(K, L) = w \qquad (3)$$

Note, however, that the firm may choose not to invest at all (that is, to stick to its existing stock of capital K_H^0) and avoid the lumpy setup cost $W_H L_H^C$. In this case its labor input, denoted by $L^-(A_H, w_H, L_H^C)$ is defined by:

$$A_H F_L(K_H^0, L) = w_H. \qquad (4)$$

In this case its present value is:

$$V^-(A_H, K, w_H) = \underset{L}{Max}\{\frac{A_H F(K_H^0, L) - w_H L + K_H^0}{(1 + r)}\}. \qquad (5)$$

The firm will make a new investment if, and only if,

$$\begin{aligned} & V^{+}\left(A_H, w_H, L_H^C\right) \\ \geqq \; & V^{-}\left(A_H, w_H\right). \end{aligned} \tag{6}$$

That is, the firm makes the amount of investment that is called for by the marginal productivity conditions (2) and (3), if and only if some global [condition (6)] is met. We assume that labor is confined within national borders. Denoting the country's endowment of labor by L_H^0, we have the following labor market clearance equation:

$$\left.\begin{array}{c} L_H^C + L^{+}(A_H, w_H, L_H^C) = L_H^0 \quad \text{if} \;\; V^{+}(A_H, w_H, L_H^C) \geqq V^{-}(A_H, w_H) \\ L^{-}(A_H, w_H) = L_H^0 \qquad \text{if} \;\; V^{+}(A_H, w_H, L_H^C) < V^{-}(A_H, w_H) \end{array}\right\} \tag{7}$$

This market clearance equation determines the wage rate in the host country as a function $W_H(A_H)$ of the productivity factor (and other exogenous factors, such as L_H^0, which are kept constant and are therefore suppressed).

Note that no similar market clearance condition is specified for capital, as we assume that capital is f reely mobile internationally and its return is fixed at r.

We now turn to discuss FDI flows from the source country S to the host country H. We treat as FDI the investment of source-country entrepreneurs in the acquisition of host country firms. Suppose that the source country entrepreneurs are endowed with some "intangible" capital, or know-how, stemming from their specialization or expertise in the industry at hand. We model this comparative advantage by assuming that the lumpy setup cost of investment in the host country, when investment is done by the source country entrepreneurs (FDI investors) is only L_H^{*C} which is below L_H^C, the lumpy setup cost of investment when carried out by the host country direct investors. This means that the foreign direct investors can bid up the direct investors of the host country in the purchase of the investing firms in the host country. The representative firm is purchased at its value which is $V^{+}[A_H, (w_H), L_H^{*C}]$. This essentially assumes that competition among the foreign direct investors pushes the price of the acquired firm to its maximized value. Thus, the FDI investors shift all the gains from their lower setup cost to the host-country original

owners of the firm. The new owners also invest an amount $K^{+}[A_H, w_H(A_H), L_H^{*C}, K_H^0\ ]$ to expand the capital stock of the acquired the firm. On the other hand, if condition (6) does not hold then there will be no FDI flows from S to H. Thus, aggregate foreign direct investment is equal to:

$$FDI = \begin{cases} V^{+}[A_H, w_H(A_H), L_H^{*C}] + K^{+}[A_H, w_H(A_H), L_H^{*C}] = K_H^0 + w_H L_H^{*C} \\ \qquad\qquad \text{if } \ V^{+}[A_H, w_H(A_H), L_H^{*C}] \geqq V^{-}[A_H, w_H(A_H)] \\ 0 \qquad\qquad \text{if } \ V^{+}[A_H, w_H(A_H) L_H^{*C}) < V^{-}[A_H, w_H(A_H)] \end{cases} \tag{8}$$

The model thus suggests that if the productivity factor (A_H) is sufficiently high, and/or the wage rate (w_H) is sufficiently low, and/or the setup cost ($w_H L_H^{*C}$) is sufficiently low, then FDI flows from country S to country H are positive. Otherwise, the flow of FDI from S to H is zero.

Recall the model's special feature is the two-fold mechanism of FDI decisions. First, one decides how much to invest abroad while ignoring the fixed setup cost. Second a decision is made whether to invest at all, while taking into account this cost. The hallmark of our empirical approach is based on the two equations (conditions) that govern these decisions. First, ignoring the setup cost, the FDI flows from Country S to country H (denoted by FDI_{NO}) is governed by a "flow" equation:

$$\begin{aligned} FDI_{NO}^{*} \ &= \ V^{+}[A_H, w_H(A_H), L_H^{*C}] + K^{+}[A_H, w_H(A_H), L_H^{*C}] \\ &\quad -K_H^0 + w_H(A_H) L_H^{*C} \end{aligned} \tag{9}$$

That is, the quantity of investment (K^{+}) and the acquisition price (V^{+}) are governed by the marginal productivity conditions (2) and (3). Second, the question whether FDI flows from S to H are at all positive, is governed by a "participation" equation (condition):

$$V^{+}[A_H, w_H, (A_H) L_H^{*C}] - V^{-}[A_H, w_H(A_H)] \geqq 0. \tag{10}$$

Consider now the effect of an increase in the host country's productivity factor A_H. Suppose initially that the wage rate in the host country (w_H) is fixed [that is, ignore the labor market clearance condition in equation (7)]. An increase in A_H raises the quantity of new investment (K^{+}), if investment is at all carried out, the acquisition price (V^{+})

that FDI investors pay, and the amount of FDI; see the appendix A. It also raises the demand for labor in the host country; see also appendix A. This will raise the wage rate w_H in the host country (and the fixed setup cost $w_H L_H^{*C}$), thereby countering the above effects on K^+, V^+, and FDI. With a unique equilibrium, the initial effects of the increase in A_H are likely to dominate the subsequent counter effects of the rise in w_H, so that FDI still rises.[4]

Thus, an increase in the host country's productivity factor A_H raises the flow volume of FDI from S to H that is governed by the flow equation. But at the same time, the rise in A_H increases also the value of the lumpy setup cost $w_H(A_H)L_H^{*C}$. Thus, it may weaken the advantage of carrying out positive FDI flows from S to H at all. In other words, the gap between V^+ and V^- in the participation equation narrows down. Thus, a positive productivity shock (unobserved in the data) raises the observed FDI flows in the flow equation and, at the same time, may lower the likelihood of observing positive FDI flows at all. In other words, the model may generate a negative correlation in the data between the residuals of the flow and participation equations.

3 The Econometric Approach

The preceding section presents a model of bilateral foreign direct investment flows distinguished by lumpy setup costs of investment.[5] Our empirical investigation is in the tradition of the often used gravity models,[6] but with adjustments for a selection bias of

[4]However, with fixed setup cost the equilibrium need not be unique, and an increase in A_H may, somewhat counter-intuitively, reduce FDI, possibly even to zero. For a similar phenomenon, see Razin, Sadka and Coury (2003).

[5]Helpman, Melitz and Yeaple (2004) pose the question of how a source country can simultaneously make both FDI outflows and exports goods to the same host country. Their answer rests on productivity heterogeneity within the source country, and differences in the setup costs associated with FDI and exports. Their explanation is thus geared toward firm-level decisions on exports and FDI in the source country.

[6]Gravity models postulate that bilateral international flows (goods, FDI, etc.) between any two economies are positively related to the size of the two economies (e.g., population, GDP), and negatively

all potential country pairs into *source* and *host* countries. With n countries in the sample, there are potentially $n(n-1)$ pairs of source-host $(s-h)$ countries. In fact, as we show in the data section below, the actual number of $s-h$ pairs is smaller. Therefore, the selection into $s-h$ pairs, which is naturally *endogenous*, cannot be ignored; that is, this selection cannot be taken as exogenous, which has been a standard practice in most gravity models.

Denote by $Y_{i,j,t}$ the flow of FDI from source country i to host country j in period t. The corresponding FDI flows from source country j to host country i are denoted by $Y_{j,i,t}$. Note that with this notation, $Y_{i,j,t}$ almost always non-negative.[7] But, it may well be zero, because typically, in a global economy, there are only a few countries which significantly export FDI to all, or most countries.

The existence of a setup cost of investment makes investment "lumpy". This means that the conventional determinants of FDI flows (such as standard marginal productivity conditions) have to generate a sufficiently large infra-marginal profits, so as to surpass a certain *unobserved* threshold. Otherwise, the observed FDI flows are practically zero. We argue that the sub-sample of FDI source countries is not a random sample of the countries in the global economy, if setup costs play a significant role in the determination of FDI flows. We now develop a simple econometric approach to study the effect of setup costs

related to the distance (physical or other such as tariff barriers, information asymmetries, etc.) between them. For instance, using population as the size variable, Loungani, Mody and Razin (2002) find that imports are less than proportionately related to the host country population, while they are close to proportionately related to the source country population. Correspondingly, FDI flows increase by more than proportionately with both the source and the host-country populations. For early works with gravity models of international trade in goods, see Eaton and Tamura (1994) and Eichengreen and Irwin (1998).

[7] This ignores rare cases of FDI and flows from country i to country j being negative, when investors from country i liquidate their investment in country j in the aggregate. For instance, outflows from the U.S. to Finland, Japan, New Zealand and Spain wee all negative in 1991. We take care of *negative outflows* in our empirical approach by allowing for two types of lumpy adjustment costs: one for setting up new investments (positive flows) and another one for liquidating existing investments (negative flows). We correct for liquidation in Table 4.

and correct for selection bias in the analysis of FDI flows.[8]

3.1 The Participation Equation

To estimate the gravity FDI flow model, and to identify the role of setup costs, the statistical model takes full advantage of the well-known Heckman's selection model [see Heckman (1979) and Kyriazidou (1996)].

To simplify, but without losing generality, let us assume that, in an imaginary world with *no setup costs,* potential FDI flows $(Y_{i,j,t})$ exhibit the following linear form:

$$Y_{i,j,t} = X_{F,i,j,t}\beta + U_{F,i,j,t}, \tag{11}$$

where $X_{F,i,j,t}$ stands for a vector of observed variables that potentially explain the pattern of FDI flows (hence the F subscript). This equation is the analogue of equation (9) in the preceding section. Such variables are, for example, per-capita income differentials between country i and country j (reflecting differences in the capital-labor ratio), as well as, language, geographical distance, legal system, and communication or transportation costs. The vector β represents the *ceteris paribus* effect of $X_{i,j,t}$ on $Y_{j,i,t}$.

The error term $U_{F,i,j,t}$ is a composite of (i) an unobserved time invariant heterogeneity $(\theta_{i,j})$, which reflects, persistent gaps between, for instance, the wage in the i source and the j host countries $(\varepsilon_{i,j})$; and (ii) a random shock term, which is $i-j$ pairwise-specific $\left(\eta_{i,j,t}\right)$, reflecting both deviations from the "long-run" wage gap $(\Delta\varepsilon_{i,j,t})$, as well as other fluctuations in macroeconomic policy, political events, etc., that are unique to the $i-j$ source-host pair.

[8] Correction for selection bias is rare in international economics literature. Notable exceptions are Broner, Lorenzoni and Schmukler (2003) and Smarzyska and Wei (2001). Broner, Lorenzoni and Schmukler applied the Heckman selection model in estimating the average maturity of sovereign debt. They take into account the incidental truncation of the data, since the average maturity is available only for countries which issue bonds to the world market. The missing observations cannot be treated as zero maturity. They show, as expected, that countries with weak macroeconomic stance are less likely to issue bonds. In this case the problem reduces to be the standard Tobin model. Smarzyska and Wei applied Heckman method to study the effects of corruption on FDI in transition economies.

Let $Z_{i,j,t}$ be a *latent* variable, which represents profits from the direct investment made in host country j, by the firms in the source country i, in period t. To simplify, we assume that profits are a linear function of the flow of FDI, which takes the form $\tilde{Z}_{i,j,t} \equiv Y_{i,j,t} - C_{i,j,t}$, where $C_{i,j,t}$ is the setup cost. Define $Z_{i,j,t} = \tilde{Z}_{i,j,t}/\sigma_{\tilde{Z}}$, where $\sigma_{\tilde{Z}}$ is the standard deviation of $\tilde{Z}$. We further assume that $Z_{i,j,t}$ exhibits the following linear form:

$$Z_{i,j,t} = X_{2,i,j,t}\gamma + V_{i,j,t}, \tag{12}$$

where $X_{2,i,j,t}$ and γ are a regressor row vector and a coefficient vector, which affect the normalized profits, respectively, and $V_{i,j,t}$ is the error term respectively. Note that the variables in the vector for X_1 are all included in the vector X_2. But vector X_2 includes also fixed-cost variables. In a random sample, we assume that the classical assumptions, regarding the error term, hold. We further assume that the error terms are normally distributed:

$$\begin{aligned} U_{i,j,t} &\sim N(0,\, \sigma_U^2), \\ V_{i,j,t} &\sim N(0, 1). \end{aligned} \tag{13}$$

We also assume that the error terms $U_{i,j,t}$ and $V_{i,j,t}$ follow a bivariate normal distribution:

$$(U_{i,j,t},\, V_{i,j,t}) \sim N(0, \Omega),$$

with variances σ_U^2 and σ_v^2, respectively.

$$\Omega = \begin{vmatrix} \sigma_U^2 & \rho \cdot \sigma_U \\ \rho \cdot \sigma_U & 1 \end{vmatrix}, \tag{14}$$

where ρ is the correlation coefficient between the cross-equation error terms.

3.1.1 Setup Costs and Selection Bias

The (statistical) population-regression function for equation (11) is:

$$E\left(Y_{i,j,t} \mid X_{F,i,j,t}\right) = X_{F,i,j,t}\beta. \tag{15}$$

Many previous studies aimed at estimating the effects of X on Y , in the context of international capital mobility, typically ignore the effect of the unobserved setup costs on the (observed) capital flows. According to our model, FDI flows ($Y_{i,j,t}$) are positive, if and only if $Z_{i,j,t} \geq 0$. Thus, we define a binary variable $D_{i,j,t}$, by

$$D_{i,j,t} = \left\{ \begin{array}{cc} 1 \quad if \quad Z_{i,j,t} = X_{2,i,j,t}\gamma + V_{i,j,t} \geq 0 \\ 0 \qquad\qquad\qquad otherwise \end{array} \right\}. \tag{16}$$

Note that whereas $Z_{i,j,t}$ *is not* observed, the binary variable $D_{i,j,t}$ *is* indeed observed. Assuming that the errors in the underlying latent equation are distributed normally then the probability setup for the probit equation exhibits the following form.

$$\Pr(D_{i,j,t} = 1 \mid \cdot) = \Pr(X_{2,i,j,t}\gamma \geq -V_{i,j,t}) = \Phi(X_{2,i,j,t}\gamma). \tag{17}$$

where Φ is the cdf of the unit normal distribution.

Therefore, the regression function for the *sub-sample* of countries for which we do indeed observe positive FDI flows is:

$$E(Y_{i,j,t} \mid X_{i,j,t},\, D_{i,j,t} = 1) = X_{i,j,t}\beta + E(U_{i,j,t} \mid X_{i,j,t},\, D_{i,j,t} = 1) \tag{18}$$

Note that the last term, the conditional expectation of $U_{i,j,t}$ does no longer equal to zero. Furthermore, it depends on $X_{i,j,t}$, thus upsetting the classical assumptions concerning regression functions when applied to random samples.

To see this, one can substitute equations (12) and (16) into equation (17) to get:

$$E(Y_{i,j,t} \,|X_{i,j,t},\, D_{i,j,t} = 1) = X_{i,j,t}\beta + E(U_{i,j,t}|\, V_{i,j,t} \geq -X_{i,j,t}\gamma). \tag{19}$$

Because $U_{i,j,t}$ and $V_{i,j,t}$ follow a bivariate normal distribution with correlation ρ and with variances σ_U^2 and σ_V^2, respectively, it follows that the expected volume of FDI flows from the source country i into the host country j in equation (18) is equal to:

$$E(Y_{i,j,t} \mid X_{i,j,t},\ D_{i,j,t} = 1) = X_{i,j,t}\beta + \rho \cdot \sigma_U \cdot \lambda_{i,j,t}, \tag{20}$$

where the inverse Mill's ratio, $\lambda_{i,j,t}$, is defined by:

$$\lambda_{i,j,t} \equiv E(V_{i,j,t} \mid V_{i,j,t} \geq -X_{i,j,t}\gamma) = \frac{\phi(-X_{i,j,t}\gamma)}{1 - \Phi(-X_{i,j,t}\gamma)} = \frac{\phi(X_{i,j,t}\gamma)}{\Phi(x_{i,j,t}\gamma)}, \tag{21}$$

and where ϕ and Φ are the unit normal density and the cumulative distribution functions, respectively. The bias (in the population) term is equal to the partial derivative of the *conditional* expectations of U with respect to X. That is:

$$bias = -\gamma \cdot \rho \cdot \sigma_U \cdot \delta_{i,j,t}, \tag{22}$$

where $\delta_{i,j,t}$ is a positive number.[9]

Figure 1 provides the intuition for the case where $\rho > 0$. Suppose, for instance, that $X_{i,j,t}$ measures the per-capita income differential between the ith source country and the jth potential host country, holding all other variables constant, namely per-capita income differentials between the ith source country and all the rest of the countries. Our theory predicts that parameter β is positive in this case. This is shown by the upward sloping line AB. Note that this slope is an estimate of the "true" underlying effect of $X_{i,j,t}$ on $Y_{i,j,t}$. But, recall that flows could be equal to zero if the set up cost are sufficiently high. The capital-flow threshold derived from the setup costs is shown as line TT' in Figure 1.

However, recall that the data include only those country pairs for which $Y_{i,j,t}$ is positive. This sub-sample is, therefore, no longer random . Moreover, as equation (12) makes clear the *selection* of country pairs into this sub-sample depends on the vector $X_{i,j,t}$.

To see this, suppose, for instance, that for high values of $X_{i,j,t}$ (the specific level X^H in Figure 1) i-j pair-wise FDI flows are all positive. That is, for all pairs of countries potential $Y_{i,j,t}$ are higher than the threshold line. Thus, the observed average, for $X_{i,j,t} = X^H$ is also equal to the conditional population average, point R on the line AB. However, this does not hold for low values of $X_{i,j,t}$ (denoted by X^L). For those i-j pairs we observe positive values of $Y_{i,j,t}$ only in a non-random sample of the population. For instance, point S is excluded from the observed sub-sample of positive FDI flows. consequently, as predicted by our model, we observe only those with low setup cost (namely high $V_{i,j,t}$), among those with low $X_{i,j,t}$ .As seen in Figure 1, the observed conditional average is at

[9] Let $\alpha = -X_{i,j,t}\gamma$. Then the partial derivative of the inverse Mills ratio is:

$$\frac{\partial \lambda(\alpha)}{\partial a} = \delta_{i,j,t} = \lambda(\alpha)[\lambda(\alpha) - \alpha],$$

so that $\delta_{i,j,t} > 0$.

point $M^{'}$, which lies above point M. The sub-sample OLS regression line is shown by the line $A^{'}B^{'}$, which understates the influence of the income per capita differentials on the flows of FDI.

3.1.2 Selection Bias: Setup Costs Versus Measurement Errors

There is a long tradition in the international economics literature of log-linearizing the capital flow gravity model, and estimating the parameters of interest by ordinary least squares (OLS). In these statistical models the gravitational force can be very small, but not zero, whereas FDI flows for a $i-j$ source-host pair of countries is often zero. The empirical literature developed after Tinbergen (1962) has often either *ignored* pairs with no FDI flows, or treated these cases as *measurement errors,* or as *literally* indicating *zero flows.*[10] This view is consistent with models that ignore the role of setup costs. In such models pairs with zero flows do indeed represent zero flows; or they reflect measurement errors (which are common with a small volume of capital flows).[11]

In our theoretical model, setup costs play an important role in determining whether a source country i invest directly in a host country j. Moreover, the model predicts that there well could be a negative correlation between the error term in the FDI flows equation and the error term in the participation equation.

This prediction of the model distinguishes between the "setup cost model" and the "measurement errors hypothesis". While the the "measurement errors hypothesis" is consistent only with a positive ρ, the "setup cost model" is consistent also with a negative ρ.

[10] Recently, Silva and Tenreyro (2003) proposed the Poisson pseudo-maximum likelihood estimator to deal with zero values in the bilateral trade models.

[11] Note that if measurement errors (in the $Y_{i,j,t}$) are not correlated with the explanatory variables, then the estimated parameters are not biased; although they are imprecisely estimated.

3.1.3 Tobit and Setup Costs

The Tobit model [see Tobin (1958)] has been often used in the empirical international trade literature [e.g., Carr, Markusen and Muskus (2001)]. The model ignores setup costs. Let $Y^*_{i,j,t}$ denote the desired FDI flows from i to j in period t:

$$Y^*_{i,j,t} = X_{i,j,t}\beta + U_{i,j,t}, \tag{23}$$

Note that $Y^*_{i,j,t}$ could be negative (for instance, when the rate of return differential works in favor of country i). The latent variable $Y^*_{i,j,t}$ is observed only if it is positive. Thus, the actual dependent variable $Y_{i,j,t}$ is by the way the data is constructed,

$$Y_{i,j,t} = \max\left(0,\ Y^*_{i,j,t}\right) \tag{24}$$

The population regression function for equation (11) is:

$$E\left(Y_{i,j,t} \mid X_{i,j,t},\ D_{i,j,t} = 1\right) = X_{i,j,t}\beta + 1 \cdot \sigma_{U_F} \cdot \tilde{\lambda}_{i,j,t}. \tag{25}$$

where

$$\tilde{\lambda}_{i,j,t} = \frac{\phi\left(\frac{X_{F,i,j,t}}{\sigma_{U_F}}\right)}{\Phi\left(\frac{X_{F,i,j,t}}{\sigma_{U_F}}\right)}. \tag{26}$$

Comparing the set of equations (9) - (10) and the set of equations (240 - (25), the Tobit model can be viewed as a special case of the Heckman model for the particular case where $\rho = 1$. In this case, the flow equation serves also as the participation equation (up to a scale). Because the only difference between the participation and the flow equations is in the role played by the setup costs, the Tobit model is the correct model under the null hypothesis of no setup costs.

Endogeneity of the explanatory variables The large fraction of country pairs with zero flows makes it clear that the selection into source and host countries is the key issue the empirical framework must address. Yet, this is not the only problem that needs to be addressed in the empirical implementation. So far we treated the explanatory variables as exogenous to the FDI flows. Although bilateral FDI flows are only a subset

of the international capital flows that enter in the host countries from all sources, we should not ignore the possibility that foreign direct investment flows from source country i to host country j may affect both economies. If such influence exists, the explanatory variables, such as GDP per capita in the source and the host countries, are expected to be correlated with the error terms in the flow and in the participation equations. We address this endogeneity problem by instrumenting our explanatory variables using lagged values. Because our theory does not generate any prior about the time structure of the X_t time series, we estimate the full system using various time lags.

4 Data

Our data is drawn from OECD reports (OECD, various years) on a sample of 24 OECD countries, over the period from 1961 to 1998. The data on FDI flows are for the period from 1981 to 1998 only. The FDI data are based on the OECD reports of FDI exports from 17 OECD source countries to 24 OECD countries.[12]

We employ 3-year averages, so that we have six periods (each consisting of 3 years). The main variables we employ are: (1) standard country characteristics such as GDP or GDP per-capita, population, educational attainment (as measured by average years of schooling), language, financial risk rating, etc.; (2) $s-h$ source-host pairs, such as $s-h$ FDI flows, geographical distance, common language (zero-one variable), $s-h$ flows of goods, bilateral telephone traffic per-capita as a proxy for informational distance, etc. Appendix B provides more information on the data: Table B1 describes the list of the 24 countries in the sample, and whether observed in the sample (at least once) as a source or host country (but most source countries do not interact more than with few

[12] However, the OECD reports accurately on *all* rich and poor countries that are a host to OECD FDI exports. But data are missing on non-OECD countries as sources of FDI exports. This is the reason that we we restrict our sample to the group of OECD countries, as potential source and host countries, among themselves. In Razin, Rubinstein, and Sadka (2003) we present also results for a 45-country sample, which include also non-OECD countries as host to FDI originating from OECD countries. (see tables D.2, D.3 and D.4).

host countries), and Table B2 describes the data sources.

5 Estimation

Table 1 and Table 2 provide a first look at the direction and volume of FDI flows. While s-h differences in GDP per capita are good predictors of the direction of flows (the exstensive margin; see Table 1), they are not correlated with the volume of FDI flows for the subset of country paris with positive flows (the intensive margin; see Table 2).

We now turn to the estimation of the determinants of these flows. To estimate the effect of GDP per capita, education, and financial risk ratings, on FDI flows, we now control for country and time fixed effects. The dependent variable in all the flow (gravity) equations is the log of the FDI flow, deflated by the unit value of manufactured goods exports.

We estimate the model under three alternative econometric procedures. As a benchmark, we ignore the selection equation (17), and simply estimate the gravity equation (11) twice: (i) by treating all FDI flows in s-h pairs with no recorded FDI flows as "zeros"; (ii) excluding country pairs with no FDI flows.[13] The rationale for inserting "zeros" is as follows. Generally, when one observes no FDI flows between a pair of countries, it could be either because the two countries do not wish to have such flows, even in the absence of fixed costs, or because setup costs are prohibitive for low flows, or because of measurement errors. But in this benchmark case, which ignores setup costs and measurement errors, s-h pairs with no FDI flows "truly" indicate zero flows. This is why we assume a neglegible value as a common low value for the value of the FDI flows for the no-flows s-h pairs.[14] (All other positive flows have logarithmic value much exceeding zero.) The estimation results for this benchmark case are shown in panel A of Table 3.

As a second benchmark, we treat all FDI flows that are below a certain low threshold level (censor) as due to measurement errors, and employ a Tobit estimator. (Note that

[13] More precisely, the log of the FDI flow is set equal to log of the lowest observed flow between any s-h country pair in the sample.

[14] We choose this value to be the lowest observe flow between any s-h country pair in the sample

this estimator is appropriate also in the case where the desired FDI flows were actually negative, as in the case where a foreign subsidiary is liquidated, but were reported as zeros.) We report the results in Panel B in Table 3, with three censor levels (lowest, 0.0 and 3.00).

Against these two benchmarks, the complete picture, and especially the role played by the unobserved fixed set up costs, are brought to the limelight, when we employ the third econometric procedure. This procedure, the Heckman selection method, jointly estimates the maximum likelihood of the flow (gravity) equation and the selection equation. This estimation accommodates both measurement errors and a possible existence of set up costs.[15] Consider a binary variable $D_{i,j,t}$ which is equal to 1 if country i exports positive FDI flows to country j at time t. Assuming that setup cost are lower if country i already invested in the past in country j, then $D_{i,j,t-k}$ could serve as an instrument in the selection equation (exclusion restriction). The results are reported in Panel C in Table 3.

Both OLS and Tobit estimations conform to the notion that the volume of FDI flows is not affected by deviation from long-run averages in the source and host countries. GDP per capita is also not significant in Heckman selection equation.[16] Turning to the effect of the host country education level, relative to the source country counterpart: while educational gaps have no effect on the intensive margin, they do have a significant effect on the extensive margin. To test whether the effect is non-linear we estimate the paramters of interest, we provide OLS and Tobit estimates for different ranges of FDI flows.[17] The first two columns report the OLS coefficients for all country-pairs and for the sub-sample of country-pairs with positive FDI flows respectively. While the coefficient of the educational gaps is positive and significantly different from zero in the first column the point estimate is substantial smaller and insignificant when we estimate the effect of

[15] We have a few cases of negative flows in our sample. We control for that using a dummy variable in the selection equation. See Appendix.

[16] Recall that in the estimation we control for country fixed effects. In Appendix C Table C.1 we present also results of the estimation without controling for country fixed effects.

[17] We are indebted to Anil Kashyap for suggesting us to compare the coefficients over different volumes of FDI flows.

educational attainements gaps within the sub-sample of country-pairs with positive FDI flows (intensive margin). This suggests that differences between source and host country schooling levels are very important in explaining the differences between country-pairs with no FDI flows (imputed flows) and country-pairs with "true" positive flows rather than the variation among country-pairs with positive FDI flows.

The effect of education on the extensive margin is also well reflected in our estimates using the Tobit and Heckman models. We find significant effects in the Tobit and Heckman models. However, while the Tobit model predicts that FDI flows are positively related to host-source difference in education levels, the Heckman model predicts that the education level affect positively the likelihood of a non-zero source-host pair, but does not influence the volume of FDI flows within the pair. Note that by imposing the no fixed cost assumption (as in the Tobit model) we might mistakenly conclude that educational gaps affect FDI volumes while in fact they affect only the extensive margin.

Source-country financial risk ratings is important in all models; but we find evidence for the importance of the ratings only in Heckman's selection equation. Improvements in the source-country financial risk rating lead to a fall in the volume of FDI flows as expected.[18] In contrast to the OLS and Tobit models, where the effects of risk ratings is only on the volume of FDI flows, in the Heckman model the effect is only on the likelihood of a country becoming a source for FDI exports. The difference between the OLS and Tobit models, on the one hand, and the Heckman model, on the other hand, is sharp when we look at the effect of host country financial risk ratings. We find no effect whatsoever in the OLS and Tobit models. In contrast, the Heckman model shows that an improvement in the host-country financial risk ratings raises the volume of FDI flows.

As expected, and consistent with previous gravity equation literature, we find that common language raises, and distance reduces the volume of FDI flows. Deviations of population size from long run averages have no effect in the OLS and Tobit models. This

[18] Note, from Tables C.1 in Appendix C, that without controling for country fixed effects the coefficient of source country financial risk rating is implausibly positive. Without country fixed effects, the coefficient may reflect unobserved, time-invariant, country characteristics, rather than the effect of risk ratings on FDI flows.

is not surprising when we look at the Heckman estimations: host-country population size affects FDI flows negatively, but the selection equation coefficient is positive. The source country population size effect is insignificant in all models.[19]

The coefficient of the lagged FDI participation variable ($D_{i,j,t-2}$) in panel C is expressed in terms of standard deviations of the unobserved profits. Thus, a pairs of countries which already had positive FDI flows between them in period $t-2$ (six years before), have the equivalent saving in setup cost of investment in period t, of a 0.7 standard deviation of profits. Most importantly as a "smoking gun" for the existence of fixed costs in the data, we note that: The correlation between the error terms in the flow and the selection equations is negative and significant. This finding, on which we further elaborate in the next section , provides an additional evidence for the relevance of fixed set up costs.

In Table 4 we use past FDI liquidations as instruments. They are good instruments because they are correlated positively with past FDI flows (Liquidations, by definition, are generated from existing stocks) but not apriori correlated with current FDI flows.

6 Interpretation

The finding that there is a significant correlation (ρ) between the error terms in the gravity and participation equations indicates that the formation of an $s-h$ pair of countries and the size of the FDI flow between this pair of countries are not independent processes. Furthermore, with ρ being negative, this correlation is consistent with the setup costs hypothesis. If shocks drive jointly marginal productivity of capital and setup costs of capital formation, then shocks to the participation equation must be negatively correlated with shocks to the flow equation. That is, below-average productivity in a host country, raising the likelihood of non-zero exports of FDI, is also associated with above-average marginal productivity of capital, raising the flow of FDI to the country (if new investment takes place at all).

If education, as measured by the average years of schooling is indeed a “good” measure

[19] Note from Tables C.1 in Appendix C, that without country fixed effects, the coefficient is significant.

for host–source country differences in human capital, then education levels are important in predicting the volume of FDI flows. The Heckman estimations predict that, as a country improves the education level, it would raise the likelihood of becoming a host to FDI flows. Likewise, improvements in the host-country financial risk ratings (where a higher rating indicates less risk) is important for her. It allows the country to solicit inward FDI flows. As expected, as far as the source country is concerned, it is just the opposite. Better risk ratings crowd out FDI outflows, diverting the flows to domestic investments. The likelihood of a country with better ratings to become a source for FDI exports is therefore lessened.

7 Conclusion

The existence of setup costs of foreign direct investment must present foreign investors with a two-fold decision: whether to establish subsidiaries in a specific host country at all, and how much to invest in the subsidiary, if they decide to establish it. Invoking this simple idea we estimate in this paper a participation equation (the decision whether to invest at all) jointly with a flow equation (the decision how much to invest).

The FDI model works as follows. A comparative advantage for the source country is based on low setup costs of direct investment, relative to setup costs of domestic investors. This allows foreign investors to bid up for investment projects in the host country. An exogenous productivity shock in the host country may affect the decision of the FDI investors whether to invest at all, and how much to invest, in opposite directions. For instance, a positive productivity shock, ceteris paribus, improves both marginal and total profitability of new investment. But, it also raises the demand for labor and consequently wages. The rise in wages, in turn, mitigates the initial rise in the marginal profitability and in the total profitability of the new investment, through its adverse effect on variable costs. However, the increase in wage costs does not completely offset the initial rise in the marginal and total productivity of new investments. As a result, the positive productivity shock implies a net rise in the marginal profitability of new investment. This

may not be the case with total profitability. It is adversely affected by the rise in wages not only through the increase in the variable costs, but also through the increase in the wage bill associated with setup costs. Hence, it may well be the case that a positive productivity shock increases the marginal productivity and lowers the total profitability of new investments, at the same time. Our model therefore provides a *rationale* for the negative correlation between the residuals of the participation and flow equations, which our econometric study is able to detect.

Empirical international trade and international finance literature often fail to address the endogeneity issue of the selection of countries into source-host country pairs. Source-host country pairs with no recorded FDI flows are either ignored, treated as measurement errors, or as if they literally indicate zero flows. A standard procedure in the literature is to treat all FDI flows that are below a certain low threshold level (censor) as due to measurement errors, and to employ a Tobit estimator. Tobit estimation is indeed often employed in the analysis of international flows of goods and capital. Evidently, the Tobit model is the correct model under the null hypothesis of no setup costs. In such a case, the error terms in the participation equation and in the flow equation are restricted to be perfectly and positively correlated, in conflict with empirical evidence and theory. The Tobit estimator is also appropriate in the case where the desired FDI flows were actually negative, as in the case where a foreign subsidiary is liquidated, but were nonetheless reported as zeros. But the Tobit estimator fails in the presence of fixed costs.

To allow for the role played by unobserved fixed setup costs, which is at the center stage of our theoretical model, we employ the Heckman selection method. We estimate jointly the maximum likelihood of the volume of FDI flows (the gravity equation), and the selection of countries into source-host country pairs (the participation equation). Evidently, this estimation procedure accommodates both measurement errors and, crucially, the possible existence of setup costs in the data. Importantly, if setup costs play an important role in determining whether a source country invests directly in a host country, then we should expect a *negative* correlation between the error terms of the gravity and the participation equation.

We do indeed find that the correlation between the error terms is negative in our data set, indicating the importance of setup costs that governs the export of FDI in the data. We find that the important predictors of the likelihood of which pair of countries will be linked by host-source relationship selection are: (1) source country GDP per capita, (2) difference in education levels (as measured by average years of schooling), and (3) differences in financial risk ratings. These variables may also be interpreted as good proxies for setup costs because they are expected to determine the technological and financial ease by which a foreign subsidiary is established. Generally, these findings support an existence of setup costs of foreign direct investment. Furthermore, the evidence points to differing effects on FDI flows driven by the marginal productivity conditions and the setup cost conditions, as rationalized in our theoretical model. The paper also sheds light on the importance of several covariates, such as income per capita, education, and financial risk ratings as key determinants of volume of FDI flows. While the coefficients of both the source- and host-country average years of schooling are positive and significant in the flow equation, the magnitude of the source country coefficient is more than twice that of the host country. That is, the richer the source country is relative to the host country, the larger are the FDI flows which occur between them. Our findings therefore suggest that capital does flow from a high income country to a low income country, and from countries with high average years of schooling to countries with low average years of schooling, in the way suggested when one looks at marginal productivity conditions alone. The characteristics of the host-source country pair with respect to the setup costs are crucially important.

References

[1] Broner, Fernando, A., Guido Lorenzoni, and Sergio L. Schmukler (2003), "Why Do Emerging Markets Borrow Short Term?" Princeton University, http://www.princeton.edu/~guido/shortterm.pdf, July.

[2] Caballero, Ricardo and Eduardo Engel (1999), "Explaining Investment Dynamics in US Manufacturing: A Generalized (S;s) Approach," **Econometrica**, July, 741-82.

[3] _____ (2000), "Lumpy Adjustment and Aggregate Investment Equations: A "Simple" Approach Relying on Cash Flow "Information," mimeo.

[4] Carr, David L., Markusen, James R. and Keith E. Maskus (2001), "Estimating the Knowledge-Capital Model of the Multinational Enterprise," **American-Economic-Review**. 91(3): 693-708.

[5] Cecchini, Paolo (1988), **The European Challenge 1992: The Benefits of a Single Market**, "The Cecchhini Report", Wilwood House.

[6] Eaton, Jonathan, and Akiko Tamura (1994), "Bi-lateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment," **The Journal of The Japanese and International Economies**, Vol. 8, pp.478-510.

[7] Eichengreen, Barry, and Douglas Irwin, (1998), "The Role of History in Bilateral Trade Flows," in Jeffrey Frankel (ed.), **The Regionalization of the World Economy**, (Chicago and London: University of Chicago Press).

[8] Heckman, James, J. (1979), "Sample Selection Bias as a Specification Error", **Econometrica** 42: 153-168.

[9] Helpman, Elhanan, Mark Melitz, and Steve Yeaple (2004), "Exports vs FDI with Heterogeneous Firms," **American Economic Review**, 94 (1):

[10] Kyriazidou, Ekaterini (1996), "Estimation of a Panel Data Sample Selection Model," **Econometrica** 65 (6): 1335-1364.

[11] Loungani, Prakash, Ashoka Mody, and Assaf Razin (2002), "The Global Disconnect: The Role of Transactional Distance and Scale Economies in Gravity Equations," **Scottish Journal of Political Economy**, 49 (5): 526-543.

[12] Razin, Assaf, Efraim Sadka, and Tarek Coury (2003), "Trade Openness, Investment Instability, and Terms-of-Trade Volatility," **Journal of International Economics**, 59 (2): .

[13] Razin, Assaf, Yona Rubinstein, and Efraim Sadka (2003), "Which Countries Export FDI, and How Much?", NBER Working Paper No. 10145, December.

[14] Smanzyska, Beata K. and Shang-Jin Wei (2001), "Corruption and Cross-Border Investment," NBER Working Paper 8465.

[15] Santos Silva, J. M. C., and Tenreyro Silvana (2003), "Gravity-defying trade," Working Papers 03-1, Federal Reserve Bank of Boston.

[16] Tinbergen, Jan (1962), **The World Economy.: Suggestions for an International Economic Policy**. New York, NY: Twentieth Century Fund.

[17] Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables", **Econometrica** 26, 24-36.

[18] Zhang, Kevin Honglin and Markusen, James R. "Vertical Multinationals and Host-Country Characteristics." **Journal of Development Economics**, August 1999, 59 (2), pp.

[19] 233.252.

8 Appendix A: A Productivity Shock

For a fixed wage rate w_H, it follows from equation (8), for the case of positive FDI flows, that

$$\frac{\partial(FDI)}{\partial A_H} = \frac{\partial V^+}{\partial A_H} + \frac{\partial K^+}{\partial A_H}. \tag{A1}$$

Using the envelope theorem, it follows from equation (1) that

$$\frac{\partial V^+}{\partial A_H} = \frac{F(K,L)}{1+r} > 0. \tag{A2}$$

Total differentiation of equations (2) and (3) with respect to A_H (while still maintaining w_H constant) yields:

$$\frac{\partial K^+}{\partial A_H} = \frac{-F_K F_{LL} + F_L F_{KL}}{A_H(F_{KK}F_{LL} - F_{KL}^2)} > 0 \tag{A3}$$

and

$$\frac{\partial L^+}{\partial A} = \frac{-F_L F_{KK} + F_K F_{KL}}{A_H(F_{KK}F_{LL} - F_{KL}^2)} > 0, \tag{A4}$$

In equations (A3) and (A4) we assume that capital and labor are substitute to each other in the production function, namely that $F_{KL} > 0$. (Recall also that $F_{KK}F_{LL} - F_{KL}^2 > 0$, $F_{KK} < 0$, and $F_{LL} < 0$, by the concavity of F.) Equations (A1) - (A3) imply that $\partial(FDI)/\partial A_H > 0$.

Thus, for a given w_H, an increase in A_H raises FDI, and K^+ and V^+.

However, when new investment is made, equation (A4) implies that a rise in A_H increases the demand for labor. When no new investment is made, it follows from equation (4), for a given w_H, that

$$\frac{\partial L^-}{\partial A_H} = -\frac{F_L}{AF_{LL}} > 0.$$

Thus, the demand for labor rises in this case as well.

9 Appendix B: Data Description

Table B1: Frequency of Source-Host Interactions by Countries

Country	**Source**	**Host**	**Country**	**Source**	**Host**
Australia	0.43	0.41	Korea	0.09	0.39
Austria	0.66	0.38	Mexico	0.00	0.33
Belgium	0.03	0.56	Netherlands	0.68	0.54
Canada	0.62	0.41	New Zealand	0.00	0.34
Denmark	0.35	0.46	Norway	0.64	0.33
Finland	0.65	0.34	Portugal	0.00	0.49
France	0.94	0.52	Spain	0.02	0.51
Germany	0.98	0.54	Sweden	0.84	0.45
Greece	0.00	0.36	Switzerland	0.27	0.47
Ireland	0.00	0.49	Turkey	0.02	0.36
Italy	0.81	0.46	United Kingdom	0.91	0.58
Japan	0.96	0.41	United States	0.87	0.64

TABLE B.2: DATA SOURCE

Variables:	**Source:**
Import of Goods	*Direction of Trade Statistics*, IMF
FDI Inflows	*International Direct Investment Database*, OECD
Unit Value of Manufactured Exports	*World Economic Outlook*, IMF
Population	*International Financial Statistics*, IMF
Distance	Shang Jin Wei's Website: www.nber.org/~wei
Bilateral Telephone Traffic	*Direction of Traffic:*
	Trends in International Telephone Tariffs,
	International Communication Union
	International Telecommunications Union
Education Attainment	Barro-Lee Dataset: www.nber.org/N...
	....
Language	....
	....
ICRG index of financially	Ashoka Mody, IMF
sound rating (inverse of financial risk)	

10 Appendix C: OECD Countries - extensions

Insert Table C.1

11 Appendix D: OECD and Non-OECD Countries

Table D1: List of Countries, by Observed Source/Host Status

Country	**Country**
Argentina	Kuwait
Australia	Malaysia
Austria	Mexico
Belgium	Netherlands
Brazil	New Zealand
Canada	Nigeria
Chile	Norway
China	Peru
Columbia	Philippines
Denmark	Portugal
Ecuador	Saudi Arabia
Egypt	Singapore
Finland	South Africa
France	Spain
Germany	Sweden
Greece	Switzerland
Hong Kong	Taiwan
India	Thailand
Ireland	Turkey
Israel	United Kingdom
Italy	United States
Japan	Venezuela
Korea	

Insert Table D.2

Insert Table D.3

Figure 1: Selection Bias in theand Setup costs Presence of Setup Costs

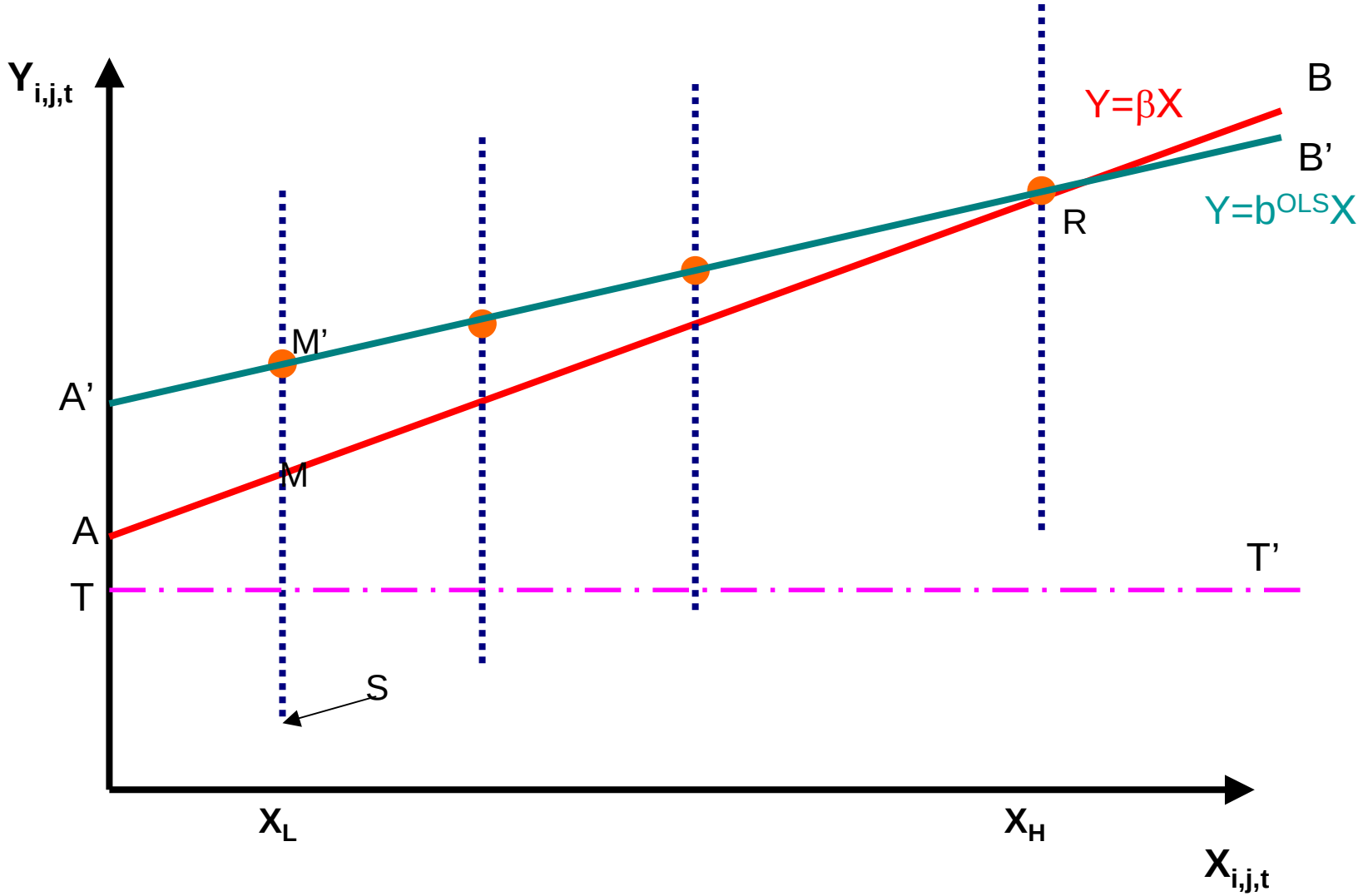

Table 1: Source-Host country pairs by GDP per capita

Country	Turkey	Mexico	Korea	Portugal	Greece	Spain	New Zealand	Ireland	Italy	UK	Canada	Australia	Finland	France	Germany	Netherlands	Sweden	Belgium	US	Austria	Norway	Denmark	Japan	Switzerland
Turkey		0	0	0	0	0	0	0	0	0.17	0	0	0	0	0.33	0	0	0	0	0	0	0	0	0
Mexico	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Korea	0	0		0	0	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0.83	0	0	0	0.67	0
Portugal	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Greece	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Spain	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0
New Zealand	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ireland	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Italy	0.83	0.17	0.5	1	1	1	0.33	1		1	1	0.67	0.5	1	1	1	0.83	1	0.83	0.67	0.83	0.83	1	0.67
UK	1	1	1	1	1	1	1	0.83	1		1	0.83	0.83	1	1	0.83	1	0.83	1	1	0.5	0.83	1	0.5
Canada	0	0.83	0.83	0.5	0.33	0.5	0.5	1	0.67	0.83		0.67	0	0.67	1	0.83	0.67	0.67	1	0.67	0.33	0.33	0.67	0.67
Australia	0.17	0	0.83	0	0	0	0.83	0.5	0.83	1	0.83		0	0.5	0.5	0.83	0	0.67	1	0	0	0	0.5	0.83
Finland	0.17	0	0	0.67	0	1	0	0.83	1	1	0.83	0.33		1	1	0.83	1	1	1	0.5	1	0.83	0.17	0.83
France	0.83	1	1	0.83	0.83	1	1	0.83	1	1	1	1	0.83		1	1	1	1	1	1	0.5	1	1	1
Germany	1	1	1	1	1	1	1	0.83	1	1	1	1	0.83	1		1	1	1	1	1	1	1	1	1
Netherlands	0.33	0.5	0.5	0.83	0.83	0.83	0	0.83	0.83	1	0.5	0.5	0.33	1	0.83		0.5	1	1	0.5	0.5	0.83	0.67	1
Sweden	1	0.67	0.67	0.83	0.83	0.83	0.5	0.83	0.67	1	0.67	0.83	1	1	0.83	1		0.83	1	0.83	1	0.83	1	0.67
Belgium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0.67	0	0	0	0	0
US	0.83	0.67	0.83	0.83	0.5	1	1	1	1	1	0.83	1	0.83	1	1	0.83	0.83	1		0.5	0.67	0.83	1	1
Austria	0.67	0.67	0.5	0.67	0.5	0.67	0.17	0.67	0.83	0.83	0.83	0.67	0.5	0.83	1	1	0.5	1	1		0.33	0.33	0.17	1
Norway	0.33	0.17	0.33	0.83	0.17	0.67	0.5	0.83	0.5	0.83	0.67	0.5	0.83	0.83	0.5	0.83	0.83	0.67	0.83	0.83		0.83	0.5	0.83
Denmark	0	0	0	0.83	0	1	0	0	0	0.83	0	0	0	0.83	1	1	0.83	0.83	0.83	0	0		0	0
Japan	0.83	1	1	1	0.83	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.67		0.83
Switzerland	0.33	0	0	0.33	0.33	0.33	0	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0	0.33	0.17	
Average	0.36	0.33	0.39	0.49	0.36	0.51	0.34	0.49	0.46	0.58	0.46	0.41	0.34	0.52	0.54	0.54	0.45	0.56	0.64	0.38	0.33	0.41	0.41	0.47

Table 2: Source-Host country Pairs by GDP per capita: FDI Flows in Percentage of GDP

Country	Turkey	Mexico	Korea	Portugal	Greece	Spain	New Zealand	Ireland	Italy	UK	Canada	Australia	Finland	France	Germany	Netherlands	Sweden	Belgium	US	Austria	Norway	Denmark	Japan	Switzerland
Turkey		0	0	0	0	0	0	0	0	0.03	0	0	0	0	0.02	0	0	0	0	0	0	0	0	0
Mexico	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Korea	0	0		0	0	0	0	0	0	0.26	0	0	0	0	0	0	0	0	0.32	0	0	0	0.03	0
Portugal	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Greece	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Spain	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0.09	0	0	0	0	0
New Zealand	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ireland	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Italy	0.66	0.29	0.13	3.64	1.53	2.51	0.05	5.73		2.7	0.49	0.19	0.26	2.24	0.42	12.2	0.75	20.1	0.41	1.15	0.15	0.27	0.08	7.82
UK	4.45	3.55	0.67	12	7.97	8.76	32.3	52.1	3.47		9.63	27.1	0.99	6.91	2.4	62.7	8.66	15.8	10.7	2.12	15.6	3.6	0.36	17.3
Canada	0	1.65	0.15	0.36	0.31	0.38	7.8	32.1	0.2	3.83		2.2	0	0.69	0.22	1.65	1.28	3.1	4	0.61	0.45	0.09	0.1	0.96
Australia	0	0	0.14	0	0	0	43.7	4.44	0.21	5.79	1.02		0	0.05	0.04	1.18	0	0.2	1.23	0	0	0	0.03	0.03
Finland	0.01	0	0	0.78	0	0.42	0	3.03	0.12	1.21	0.51	0.09		0.4	0.5	4.48	32.7	1.93	0.27	0.32	3.1	3.96	0.01	0.67
France	3.27	1.19	0.99	8.42	2.75	12.1	1.42	7.91	6.57	11	4.35	3.56	0.53		3.36	27.2	6.71	44.5	3.83	2.1	2.41	1.84	0.07	16.6
Germany	4.68	3.36	1.81	9.29	4.03	8.99	0.67	69	6.19	16.6	4.66	2.88	2.07	8.02		19.9	6.12	39.6	4.69	22.7	4.3	4.73	0.37	18.3
Netherlands	0.98	1.49	0.48	5.77	3.8	5.48	0	35.1	1.24	13.1	1.35	2.24	0.46	3.34	1.25		6.5	40	3.25	1.31	2.84	5.65	0.09	9.97
Sweden	0.18	0.46	0.27	0.78	0.11	0.79	0.14	21.1	0.52	4.31	0.31	0.43	35.4	1.56	0.56	9.93		2.73	0.99	0.6	15.4	6	0.02	3.34
Belgium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0.49	0	0	0	0	0
US	3.42	36.2	4.78	6.84	1.22	6.54	26.1	127	6.35	57	47	27.4	4.06	8	4.29	60.3	5.65	35.7		4.24	16.8	3.85	1.26	39.9
Austria	0.18	0.02	0.01	0.46	0.13	0.22	0.05	2.14	0.26	0.82	0.22	0.28	0.02	0.12	0.42	0.42	0.19	0.92	0.09		0.04	0.67	0	1.25
Norway	0.02	0	0	1.14	0.01	0.42	0.18	4.08	0.1	1.56	0.88	0.06	1.81	0.35	0.08	1.63	8.37	0.92	0.2	0.66		7.11	0	0.18
Denmark	0	0	0	0.81	0	1	0	0	0	3.1	0	0	0	0.39	0.21	2.37	7.52	0.9	0.21	0	0		0	0
Japan	1.75	4.15	7.71	1.2	0.54	2.69	16.7	19.1	0.82	19.1	7.66	34.2	0.64	2.65	1.29	28.3	0.26	18.2	15.7	0.97	3.32	0.16		4.48
Switzerland	0.68	0	0	0.88	1.39	0.44	0	5.5	0.51	4.88	1.3	1.43	1.64	0.63	1.02	3.23	2.84	3.3	1.01	1.42	0	0.51	0.01	
Average	0.88	2.28	0.75	2.28	1.03	2.21	5.62	16.9	1.15	6.32	3.45	4.44	2.08	1.54	0.7	10.2	3.81	9.9	2.07	1.66	2.8	1.67	0.11	5.26

Table 3:
Bilateral FDI Flows and Selection into Source-Host Pairs:
OLS, Tobit Hekcman Maximum Likelihood,
Controlling for Country Fixed Effects,
OECD Countries only

	Panel A: OLS		Panel B: Tobit Correction			Panel C: Heckman selection	
	Sample:		Low censored (in logs)			Equation:	
Variables	All^^	Intensive margin	lowest	0	3	FDI Flows	Selection
GDP per capita - host^	0.260	0.445	-0.151	-0.040	0.107	0.330	-0.421
	(0.997)	(0.689)	(2.294)	(1.172)	(1.016)	(0.683)	(0.769)
GDP per capita - source^	-0.653	0.640	-0.861	-0.174	-0.211	0.648	-0.338
	(0.797)	(0.576)	(2.421)	(1.231)	(1.059)	(0.558)	(0.841)
Difference between source and host years of schooling	**0.367**	**0.018**	**0.855**	**0.413**	**0.321**	**-0.020**	**0.273**
	(0.146)*	**(0.096)**	**(0.282)****	**(0.145)****	**(0.126)***	**(0.101)**	**(0.099)****
Common language	**0.749**	**1.021**	**1.599**	**1.193**	**1.146**	**0.975**	**0.303**
	(0.250)**	**(0.146)****	**(0.319)****	**(0.162)****	**(0.139)****	**(0.130)****	**(0.133)***
Distance (in logs)	**-0.830**	**-0.677**	**-1.547**	**-1.003**	**-0.902**	**-0.633**	**-0.382**
	(0.138)**	**(0.095)****	**(0.188)****	**(0.095)****	**(0.082)****	**(0.092)****	**(0.088)****
Population - host^	6.825	-1.943	15.543	5.511	3.269	-2.973	7.232
	(3.888)	(2.369)	(7.776)*	(3.959)	(3.417)	(2.373)	(2.592)**
Population - source^	5.023	-0.492	10.322	5.310	5.442	-1.289	2.013
	(3.232)	(3.029)	(9.094)	(4.648)	(4.040)	(2.938)	(2.669)
Financial risk rating - host	**-0.029**	**0.045**	**-0.048**	**-0.006**	**0.006**	**0.050**	**-0.029**
	(0.027)	**(0.017)****	**(0.062)**	**(0.032)**	**(0.027)**	**(0.017)****	**(0.021)**
Financial risk rating - source	**-0.098**	**-0.035**	**-0.235**	**-0.137**	**-0.118**	**-0.027**	**-0.066**
	(0.025)**	**(0.026)**	**(0.081)****	**(0.042)****	**(0.036)****	**(0.026)**	**(0.025)****
Export of FDI flows from i to j six years ago (=1 if yes)							0.838
							(0.124)**
Correlation (Ui,j, Vi,j)						**-0.429**	
						(0.196)	
Inverse Mills ratio						**-0.429**	
						(0.240)	
Observations	2116	995	2116	2116	2116	2116	2116
Left-censored observations	--	--	1121	1141	1174	--	--
Uncensored observations	--	--	995	975	942		

Note:
^ in logs
^^ Replacing the zeros by the lowest observed flow between any s-h country pair in the sample.
All specifications include year fixed-effects.
Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table 4
Bilateral FDI Flows and Selection into Source-Host Pairs:
OLS, Tobit Hekcman Maximum Likelihood,
Controlling for Country Fixed Effects and Past Liquidations
OECD Countries only

	Panel A: OLS		Panel B: Tobit Correction			Panel C: Heckman selection	
	Sample:		Low censored (in logs)			Equation:	
Variables	All^^	Intensive margin	lowest	0	3	FDI Flows	Selection
GDP per capita - host^	0.219	0.440	-0.287	-0.104	0.064	0.350	-0.475
	(0.987)	(0.690)	(2.288)	(1.171)	(1.016)	(0.682)	(0.759)
GDP per capita - source^	-0.543	0.584	-0.460	-0.017	-0.104	0.581	-0.202
	(0.796)	(0.580)	(2.418)	(1.232)	(1.060)	(0.562)	(0.845)
Difference between source and host years of schooling	**0.386**	**0.012**	**0.917**	**0.438**	**0.338**	**-0.029**	**0.288**
	(0.148)**	**(0.097)**	**(0.282)****	**(0.145)****	**(0.126)****	**(0.103)**	**(0.102)****
Common language	**0.762**	**1.014**	**1.655**	**1.217**	**1.162**	**0.965**	**0.315**
	(0.254)**	**(0.146)****	**(0.319)****	**(0.162)****	**(0.139)****	**(0.129)****	**(0.138)***
Distance (in logs)	**-0.836**	**-0.674**	**-1.572**	**-1.013**	**-0.909**	**-0.629**	**-0.393**
	(0.139)**	**(0.095)****	**(0.187)****	**(0.095)****	**(0.082)****	**(0.092)****	**(0.091)****
Population - host^	6.794	-1.967	15.401	5.460	3.237	-2.960	7.232
	(3.894)	(2.384)	(7.756)*	(3.956)	(3.417)	(2.393)	(2.626)**
Population - source^	5.395	-0.703	12.083	6.000	5.892	-1.536	2.828
	(3.220)	(3.032)	(9.102)	(4.659)	(4.050)	(2.933)	(2.724)
Financial risk rating - host	-0.028	0.045	-0.045	-0.005	0.007	0.050	-0.029
	(0.027)	(0.017)**	(0.061)	(0.032)	(0.027)	(0.017)**	(0.021)
Financial risk rating - source	-0.098	-0.034	-0.245	-0.141	-0.120	-0.025	-0.071
	(0.024)**	(0.026)	(0.081)**	(0.042)**	(0.036)**	(0.026)	(0.025)**
Negative flows from I to j three years ago (=1 if yes)^^^	**0.661**	**-0.169**	**1.592**	**0.610**	**0.418**	**-0.243**	**0.505**
	(0.423)	**(0.152)**	**(0.508)****	**(0.257)***	**(0.222)**	**(0.155)**	**(0.164)****
Export of FDI flows from i to j six years ago (=1 if yes)						**0.841**	
						(0.127)**	
Correlation (Ui,j, Vi,j)						**-0.425**	
						(0.206)	
Inverse Mills ratio						**-0.486**	
						(0.252)	
Observations	2116	995	2116	2116	2116	2116	2116
Left-censored observations	--	--	1121	1141	1174	--	--
Uncensored observations	--	--	995	975	942		

Note:
^ in logs
^^ Replacing the zeros by the lowest observed flow between any s-h country pair in the sample.
^^^ FDI flows from country i to country j being negative.

All specifications include year fixed-effects.
Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table C.1
Bilateral FDI Flows and Selection into Source-Host Pairs:
OLS, Tobit Hekcman Maximum Likelihood,
Without Country Fixed Effects,
OECD Countries only

	Panel A: OLS		Panel B: Tobit Correction			Panel C: Heckman selection	
	Sample:		Low censored (in logs)			Equation:	
Variables	All^^	Intensive margin	lowest	0	3	FDI Flows	Selection
GDP per capita - host^	0.164	0.366	0.084	0.232	0.192	0.365	-0.232
	(0.313)	(0.212)	(0.455)	(0.238)	(0.208)	(0.213)	(0.119)
GDP per capita - source^	3.923	0.905	9.034	4.611	3.857	0.630	1.166
	(0.265)**	(0.357)*	(0.571)**	(0.298)**	(0.259)**	(0.346)	(0.152)**
Difference between source and host years of schooling	-0.036	-0.050	-0.020	-0.040	-0.037	-0.053	0.012
	(0.052)	(0.031)	(0.080)	(0.042)	(0.037)	(0.031)	(0.020)
Common language	0.522	1.146	0.905	0.847	0.873	1.097	-0.038
	(0.387)	(0.241)**	(0.405)*	(0.210)**	(0.181)**	(0.231)**	(0.110)
Distance (in logs)	-0.780	-0.532	-1.482	-0.888	-0.802	-0.474	-0.128
	(0.129)**	(0.078)**	(0.147)**	(0.077)**	(0.067)**	(0.078)**	(0.041)**
Population - host^	0.720	0.662	1.348	0.882	0.812	0.614	0.089
	(0.129)**	(0.077)**	(0.150)**	(0.079)**	(0.068)**	(0.079)**	(0.040)*
Population - source^	2.117	0.799	3.278	1.908	1.686	0.680	0.378
	(0.089)**	(0.066)**	(0.155)**	(0.082)**	(0.071)**	(0.072)**	(0.045)**
Financial risk rating - host	0.115	0.109	0.220	0.145	0.141	0.103	0.028
	(0.031)**	(0.020)**	(0.051)**	(0.027)**	(0.024)**	(0.020)**	(0.013)*
Financial risk rating - source	0.050	0.086	0.262	0.144	0.132	0.077	0.026
	(0.027)	(0.027)**	(0.066)**	(0.035)**	(0.031)**	(0.027)**	(0.015)
Export of FDI flows from i to j six years ago (=1 if yes)							1.613
							(0.091)**
Correlation (Ui,j, Vi,j)						**-0.383**	
						(0.089)	
Inverse Mills ratio						**-0.383**	
						(0.089)	
Observations	2116	995	2116	2116	2116	2116	2116
Left-censored observations	--	--	1121	1141	1174	--	--
Uncensored observations	--	--	995	975	942		

Note:
^ in logs
^^ Replacing the zeros by the lowest observed flow between any s-h country pair in the sample.
All specifications include year fixed-effects.
Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table D.2:
Bilateral FDI Flows and Selection into Source-Host Pairs:
OLS, Tobit Hekcman Maximum Likelihood,
Controlling for Country Fixed Effects,
OECD and Non-OECD Countries.

	Panel A: OLS		Panel B: Tobit Correction			Panel C: Heckman selection	
	Sample:		Low censored			Equation:	
Variables	All^^	Intensive margin	lowest	0	3	FDI Flows	Selection
GDP per capita - host^	0.239	0.116	-1.463	-0.892	-0.367	0.148	-0.370
	(0.175)	(0.450)	(1.189)	(0.636)	(0.468)	(0.446)	(0.383)
GDP per capita - source^	0.066	0.437	0.637	0.577	0.619	0.388	0.399
	(0.083)	(0.457)	(1.993)	(0.982)	(0.624)	(0.446)	(0.707)
Difference between source and host years of schooling	**0.211**	**0.116**	**0.708**	**0.388**	**0.188**	**0.083**	**0.227**
	(0.064)**	**(0.086)**	**(0.237)****	**(0.120)****	**(0.084)***	**(0.088)**	**(0.075)****
Common language	**0.383**	**0.846**	**1.647**	**1.094**	**0.879**	**0.792**	**0.301**
	(0.133)**	**(0.123)****	**(0.257)****	**(0.126)****	**(0.080)****	**(0.111)****	**(0.099)****
Distance (in logs)	**-0.633**	**-0.800**	**-1.716**	**-1.113**	**-0.803**	**-0.745**	**-0.413**
	(0.068)**	**(0.077)****	**(0.149)****	**(0.073)****	**(0.046)****	**(0.074)****	**(0.073)****
Population - host^	2.961	1.585	17.966	7.797	3.452	0.577	5.396
	(0.768)**	(1.355)	(3.704)**	(1.872)**	(1.298)**	(1.359)	(1.229)**
Population - source^	-2.338	1.012	-8.367	-1.779	2.949	1.352	-5.542
	(0.459)**	(2.561)	(7.509)	(3.736)	(2.532)	(2.480)	(2.305)*
Financial risk rating - host	**-0.014**	**0.036**	**-0.027**	**0.001**	**0.019**	**0.038**	**-0.020**
	(0.008)	**(0.011)****	**(0.037)**	**(0.019)**	**(0.013)**	**(0.011)****	**(0.012)**
Financial risk rating - source	**-0.058**	**-0.056**	**-0.261**	**-0.162**	**-0.120**	**-0.044**	**-0.070**
	(0.007)**	**(0.025)***	**(0.070)****	**(0.035)****	**(0.024)****	**(0.025)**	**(0.021)****
Export of FDI flows from i to j six years ago (=1 if yes)							0.721
							(0.099)**
Correlation (Ui,j, Vi,j)						**-0.471**	
						(0.148)	
Inverse Mills ratio						**-0.559**	
						(0.192)	
Observations	6724	1482	6724	6724	6724	6724	6724
Left-censored observations	--	--	5242	5301	5605	--	--
Uncensored observations	--	--	1482	1423	1119		

Note:
^ in logs
^^ Replacing the zeros by the lowest observed flow between any s-h country pair in the sample.
All specifications include year fixed-effects.
Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table D.3:
Bilateral FDI Flows and Selection into Source-Host Pairs:
OLS, Tobit Hekcman Maximum Likelihood,
Without Country Fixed Effects,
OECD and Non-OECD Countries.

	Panel A: OLS		Panel B: Tobit Correction			Panel C: Heckman selection	
	Sample:		Low censored			Equation:	
Variables	All^^	Intensive margin	lowest	0	3	FDI Flows	Selection
GDP per capita - host^	0.602	0.811	1.479	0.990	0.774	0.757	0.040
	(0.089)**	(0.095)**	(0.213)**	(0.107)**	(0.073)**	(0.095)**	(0.050)
GDP per capita - source^	1.965	0.764	9.615	4.736	2.573	0.478	1.203
	(0.084)**	(0.306)*	(0.453)**	(0.225)**	(0.146)**	(0.298)	(0.099)**
Difference between source and host years of schooling	-0.131	-0.087	-0.153	-0.113	-0.079	-0.085	-0.001
	(0.027)**	(0.027)**	(0.064)*	(0.032)**	(0.022)**	(0.026)**	(0.016)
Common language	0.348	1.057	1.224	0.967	0.951	1.006	0.021
	(0.149)*	(0.172)**	(0.320)**	(0.159)**	(0.104)**	(0.165)**	(0.079)
Distance (in logs)	-0.621	-0.496	-1.660	-0.933	-0.625	-0.426	-0.136
	(0.085)**	(0.066)**	(0.133)**	(0.066)**	(0.043)**	(0.066)**	(0.034)**
Population - host^	0.470	0.681	1.680	1.005	0.729	0.616	0.151
	(0.063)**	(0.063)**	(0.119)**	(0.060)**	(0.040)**	(0.064)**	(0.030)**
Population - source^	1.483	0.857	3.859	2.125	1.364	0.707	0.433
	(0.058)**	(0.054)**	(0.125)**	(0.063)**	(0.043)**	(0.062)**	(0.035)**
Financial risk rating - host	0.047	0.063	0.202	0.116	0.084	0.055	0.021
	(0.009)**	(0.010)**	(0.028)**	(0.014)**	(0.010)**	(0.010)**	(0.007)**
Financial risk rating - source	0.037	0.106	0.339	0.172	0.127	0.093	0.035
	(0.007)**	(0.024)**	(0.050)**	(0.026)**	(0.018)**	(0.024)**	(0.008)**
Export of FDI flows from i to j six years ago (=1 if yes)							1.663
							(0.075)**
Correlation (Ui,j, Vi,j)						**-0.382**	
						(0.062)	
Inverse Mills ratio						**-0.618**	
						(0.111)	
Observations	6724	1482	6724	6724	6724	6724	6724
Left-censored observations	--	--	5242	5301	5605	--	--
Uncensored observations	--	--	1482	1423	1119		

Note:
^ in logs
^^ Replacing the zeros by the lowest observed flow between any s-h country pair in the sample.
All specifications include year fixed-effects.
Robust standard errors in parentheses
* significant at 5%; ** significant at 1%