Fixed Costs and Bilateral FDI Flows: Conflicting Effects of Country Specific Productivity Shocks*

Assaf Razin[†], Yona Rubinstein[‡] and Efraim Sadka[§]

March, 2005

Abstract

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. We estimate in this paper a selection equation (the decision whether to invest at all) jointly with a flow equation (the decision how much to invest). 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. In a sample of 24 OECD countries, over the period 1981-1998,

*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 from participants in the November 2004 NBER IFM meeting. 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.

we observe many pairs of countries with no FDI flows between them. Zero reported flows could indicate either measurement errors, or genuine no FDI flows that are due to fixed costs (if the total profitability condition dominates the marginal profitability condition). We employ the Heckman selection procedure and demonstrate how to get unbiased estimates of the unobserved fixed-costs.

1 Introduction

The paper develops an international capital flows' model, with fixed (lumpy) set up costs of new investment which govern the flow of FDI. As fixed costs are typically unobserved, it is an econometric challenge to bring their existence to the surface. In this paper we develop a methodology to test the importance of the role played by setup costs in forming and enhancing bilateral FDI flows.[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, a standard first-order condition). 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 whether to carry out new investments at all. This decision is governed by the total (rather than the marginal) profitability of the new investment. In the econometric terminology, this decision is described by a so-called selection equation. One would expect that productivity differences would be a key factor that drives FDI flows. Thus, a high level of productivity in the potential source country versus a low level of productivity in the potential host country would put adverse pressures on FDI flows. We

[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). However, empirical international macroeconomics, which focuses on country-specific characteristics, has to date not incorporated such costs.

point out that when threshold barriers, typical for FDI, are taken into account this simple prescription needs a substantial modification. We show that the productivity shocks manifest themselves differently in the two-fold (the selection and flow) FDI decisions. Furthermore, their effects depend also on whether FDI is in the form of M&A foreign investment or in the form of a greenfield foreign investment. We demonstrate that in the presence of fixed costs, a productivity shock in the host country may also, on the one hand, increase 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.

Controlling for the selection into source-host pairs of countries, and for time and country fixed effects, the paper sheds light on the importance of several driving forces, such as income per capita, education, and financial risk ratings as key determinants of volume of FDI flows. Inter-country differences in income per capita, average years of schooling, and finacial ratings, in ways suggested when one looks at marginal productivity conditions alone, are not sufficient to predict the direction and magnitude of FDI capital flows.

The organization of the paper is as follows. Section 2 presents our model of fixed setup costs of foreign direct investment. Section 3 include the analysis of the conflicting effects of productivity shocks. Section 4 presents the econometric approach. The data

[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 incomplete about the exports of FDI to non-OECD countries.

are described in Section 5. Estimation results of the determinants of FDI flows, and whether source-host flows are formed at all, are presented in Section 6. Evidence on the existence of unobserved fixed costs in the sample is interpreted in Section 7. Conclusions are drawn in Section 8.

2 A Model of Country-Specific Productivity Shocks

The stylized model serves to underpin the paper's econometric analysis. In a nutshell our the model of FDI works as follows. First, a potential FDI investor decides how much she would like to invest. This decision is governed by the marginal profitability considerations, so as to equate the marginal factor productivity to factor prices (that is, the standard first-order condition). In an econometric terminology, this decision is described by a *flow* (or gravity) equation. Second, the potential FDI investor must also decide whether to carry out at all new investments, because of fixed costs of new investments. The decision is governed by the total (rather than marginal) profitability of the new investment. In an econometric terminology, such decision is described by the so-called *selection* equation. A productivity shock in the host country may, on the one hand, increase the volume of the desired FDI flows to this country, but, on the other hand, and somewhat counter-intuitively, the shock may lower the likelihood of making new FDI flows at all, by the source country. A source-country positive productivity shocks has a negative effect in the likelihood of making a new FDI, but is inconsequential for the flow of FDI. As we focus on *aggregate* bilateral capital flows in the econometric analysis, we specify in the theory background the general productivity level of a country, and ignore for simplicity heterogeneity among firms within a country [3].

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. As before, there is a single good which serves both for consumption and investment. In a straightforward extension of the model to more than one industry every country becomes potentially both a source for

[3] For notational simplicity we also set number of firms in the industry to be equal to one.

FDI flows to several host countries, and a host for FDIflows from several source countries. But because of fixed costs, some of the source-host pairs are inactive.

As our focus here is on the country-specific productivity shocks, we would like to reckon with the possibility that a productivity change affects wages. If the setup cost is in part in domestic (host-country) inputs, we have to take into account the indirect effect of a productivity change on the setup cost. Therefore, we assume that the setup cost is of the form

$$C_H = C_{SH} + w_H L_H^C, \tag{1}$$

where C_{SH} is a cost incurred in the source country and L_H^C is a fixed input of domestic labor.

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

$$V^+(A_H, C_H, w_H) = \max_{(K,L)} \left\{ \frac{A_H F(K,L) - w_H L + K}{1+r} - [(K - K_H^0) + C_H] \right\}. \tag{2}$$

The demand of the firm for K and L are denoted by $K^+(A_H, w_H)$ and $L^+(A_H, w_H)$, respectively. They are defined by the marginal productivity conditions:

$$A_H F_K(K,L) = r, \tag{3}$$

and

$$A_H F_L(K,L) = w. \tag{4}$$

Note again that the firm may choose not to invest at all (that is, to stick to the existing stock of capital K_H^0) and thereby avoid the lumpy setup cost C_H. In this case its present value is:

$$V^-(A_H, K_H^0, w_H) = \max_{L} \left\{ \frac{A_H F(K_H^0, L) - w_H L + K_H^0}{1+r} \right\}, \tag{5}$$

and its labor demand, denoted by $L^-(A_H, K_H^0, w_H)$, is given by

$$A_H F_L(K_H^0, L^-(A_H, K_H^0, w_H)) = w_H. \tag{6}$$

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

$$V^+(A_H, C_H, w_H) \geq V^-(A_H, K_H^0, w_H). \tag{7}$$

That is, the firm makes the amount of investment that is called for by the marginal productivity conditions, (3) and (4), if and only if, a global *selection* condition (7), is met.

As before, 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 clearing equation:

$$\left.\begin{array}{ll} L_H^C + L^+(A_H, w_H) = L_H^0 & \text{if } V^+(A_H, C_H, w_H) \geq V^-(A_H, K_H^0, w_H) \\ L^-(A_H, w_H) = L_H^0 & \text{if } V^+(A_H, C_H, w_H) < V^-(A_H, K_H^0, w_H) \end{array}\right\} \tag{8}$$

This market clearing equation determines the wage rate in the host country, as a function $w_H(A_H)$ of the host-country productivity factor. See Appendix A for a derivation of the partial derivative of FDI with respect to the productivity shock.

3 Conflicting Effects of Source- and Host-Country Productivity Shocks

We now turn to discuss the determinants FDI flows from the source country S to the host country H. We treat as FDI the investment of source-country entrepreneurs in the mergers and acquisitions (M&A) 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 below the lumpy setup cost of investment, if 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 acquisition of the investing firms in the host country. The representative firm is purchased at its value which is $V^{+}[A_H, C_H, w_H(A_H)]$. This essentially assumes that competition among the foreign direct investors pushes the price of the acquired firm to a 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 $K^{+}[A_H, w_H(A_H)]$ to expand the capital stock of the acquired the firm. On the other hand, if the selection condition (7) does not hold, then there will be no FDI flows from country S to country H. Thus, aggregate foreign direct investment is equal to:

$$FDI = \begin{cases} V^{+}[A_H, C_H, w_H(A_H)] + K^{+}[A_H, w_H(A_H)] - K_H^0 + w_H(A_H)L_H^{*C} \\ \text{if } V^{+}[A_H, C_H, w_H(A_H)] \geq V^{-}[A_H, K_H^0, w_H(A_H)] \\ 0 \\ \text{if } V^{+}[A_H, C_H, w_H(A_H)] < V^{-}[A_H, K_H^0, w_H(A_H)] \end{cases} . \quad (9)$$

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 ($C_{SH} + w_H L_H^C$) is sufficiently low, then FDI flows from country S to country H are positive. Otherwise, the flow of FDI from country S to country H must be zero.

As a preamble to our empirical analysis in the next part, recall that 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, taking into account this cost. The hallmark of our empirical approach to follow 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_{NOF}) is govern by a "notional" flow equation:

$$FDI_{NOF} = V^{+}[A_H, C_H, w_H(A_H)] + K^{+}[A_H, w_H(A_H)] - K_H^0 + w_H(A_H)L_H^C. \quad (10)$$

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

$$V^+[A_H, C_H, w_H(A_H)] - V^-[A_H, K_H^0, w_H(A_H)] \geqq 0. \tag{11}$$

Consider now the effect of a postive productivity shock which raises the host country's productivity factor, A_H. As before, suppose initially that the wage rate in the host country (w_H) is fixed [that is, ignore the labor market clearing condition in equation (8)]. An increase in A_H raises the quantity of new investment (K^+), if the investment is carried out at all, the acquisition price (V^+) that FDI investors pay, the amount of FDI, and the demand for the labor in the host country. The increase in the demand for labor 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 volume of FDI flows from country S to country H that is governed by the flow equation. But, at the same time, the rise in A_H increases also the value of the domestic component of the setup cost, $w_H(A_H)L_H^C$. Thus, it may weaken the advantage of carrying out positive FDI flows from country S to country H at all. In other words, the gap between V^+ and V^- in the selection equation narrows down. Thus, a positive productivity shock (typically unobserved in the data) raises the observed FDI flows in the flow equation but, 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 selection equations.

[4]However, with fixed setup cost the equilibrium need not to 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).

The productivity level (A_S) in the source country comes into play in the selection decision, when we consider again the limited supply of entrepreneurs in the source country. This consideration is particularly relevant for *greenfield* FDI. A source-country entrepreneur then faces a discrete choice of whether to invest either at home or abroad, but not in both. In this case, in order for her to make greenfield FDI, it no longer suffices that V^+ exceeds V^-; rather V^+ must also exceed the value of alternative direct investment at home. The latter naturally depends on the source-country productivity level, A_S, and we denote it by $B(A_S)$. That is, the selection condition is:

$$V^+[A_H, C_H, w_H(A_H)] > Max\left\{V^-[A_H, K_H^0, w_H(A_H)], B(A_S)\right\}. \tag{12}$$

Thus, the source-country positive productivity shock affects negatively the selection decision, but it has no bearing on the flow decision.

The FDI flow mechanism 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 selection and flow equations, which our econometric study is able to detect in section 6.

4 Econometric Application

Our empirical investigation is in the tradition of an often used gravity model, but with adjustments for a selection bias of all potential country pairs into *source* and *host* countries. As Feenstra (2004) explains, "In its simplest form, the gravity equation states that the bilateral trade between two countries is directly proportional to the product of the country's GDP. Thus, larger countries will tend to trade more with each other, and countries that are more similar in their relative sizes will also trade more. This equation performs extremely well empirically, as has been known since the original work of Tinbergen (1992)." [5]The size of a country may be alternatively represented by the size of its population. Gravity models may postulate also that bilateral international flows between any two economies are negatively related to the distance (physical or others, such as tariff barriers, standards and regulations, information asymmetries, etc.) between them.[6]

With n countries in the sample, there are potentially $n(n-1)$ pairs of source-host (s,h) countries with positive bi-lateral flows. In fact, as we show in the data section below, the actual number of (s,h) pairs is much smaller. Therefore, the selection of (s,h) pairs, which is naturally non random and *endogenous*, cannot be ignored; that is, this selection cannot be taken as exogenous, as has been a standard practice in many gravity models in the literature.

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 $Y_{j,i,t}$.

[5]For pioneering works with gravity models of international trade in goods, see Eaton and Tamura (1994) and Eichengreen (1998).

[6]For instance, using population as the size variable, Loungani, Mody and Razin (2002) find that imports of goods are less than proportionately related to the host country population, while they are close to being proportionately related to the source country population. Correspondingly, FDI flows increase by more than proportionately with both the source and the host-country populations.

Note that with this notation, $Y_{i,j,t}$ is almost always non-negative[7]. But, it may well be literally zero, because typically, in a global economy, there are only a few countries which significantly export FDI to all, or even many countries[8].

4.1 Selection and Flow Equations

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:

$$Y_{i,j,t} = X_{F,i,j,t}\beta + U_{i,j,t}, \tag{13}$$

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 (10) in section 3. Such variables are, for example, per-capita income differentials between country i and j (reflecting differences in the capital-labor ratio), as well as language, geographical distance, legal system, communication cost, or transportation cost. The vector β represents the standard *ceteris paribus* effect of $X_{F,i,j,t}$ on $Y_{i,j,t}$.

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

[8]A correction for selection bias is rare in the international economics literature. Notable exceptions are Broner, Lorenzoni and Schmukler (2003), Smarzynska and Wei (2001), and Helpman, Melitz and Rubinstein (2004). Broner, Lorenzoni and Schmukler (2003) 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, however, 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 econometric problem reduces to a standard Tobit model. Smarzynska and Wei (2001) applied Heckman method in a study of the effects of corruption on FDI in transition economies. Helpman, Melitz and Rubinstein (2004) study the selection of countries into trading partners in goods, using the Heckman selection method.

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

Let $Z_{i,j,t}$ be a *latent* variable, which represents total profits from the direct investment made in host country j, by firms in source country i, in period t.[9] We assume that $Z_{i,j,t}$ exhibits a linear form:

$$Z_{i,j,t} = X_{S,i,j,t}\gamma + V_{i,j,t}, \tag{14}$$

where $X_{S,i,j,t}$ and γ are, respectively, a regressor row vector and a coefficient vector, which affect the normalized profits, and $V_{i,j,t}$ is the error term. Note that all the variables in the vector X_F are also included in the vector X_S. But the vector X_S includes also fixed-cost variables.

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

$$(U_{i,j,t}, V_{i,j,t})\tilde{}N(0,\Omega), \tag{15}$$

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

The covariance matrix Ω is given by

$$\Omega = \begin{vmatrix} \sigma_U^2 & \rho\cdot\sigma_U \\ \rho\cdot\sigma_U & 1 \end{vmatrix}, \tag{16}$$

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

[9]To simplify, we assume that profits (excluding the setup costs) are a linear function of the flows 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 the normalized variable $Z_{i,j,t} = \tilde{Z}_{i,j,t}/\sigma_{\tilde{Z}}$, where $\sigma_{\tilde{Z}}$ is the standard deviation of $\tilde{Z}$.

4.1.1 Setup Costs and Selection Bias

The (statistical) population-regression function for equation (1) is:

$$E(Y_{i,j,t}|X_{F,i,j,t}) = X_{F,i,j,t}\beta. \tag{17}$$

According to our model, FDI flows ($Y_{i,j,t}$) are positive, if and only if, $Z_{i,j,t} \geqslant 0$ and otherwise they are zero. We can accordingly define a binary variable $D_{i,j,t}$:

$$D_{i,j,t} = \begin{Bmatrix} 1 & \text{if } Z_{i,j,t} = X_{F,i,j,t}\gamma + V_{i,j,t} \geqslant 0 \\ 0 & \text{otherwise} \end{Bmatrix}. \tag{18}$$

Note that whereas profits ($Z_{i,j,t}$) are not observed, the binary variable $D_{i,j,t}$, which indicate whether or not flows are positive, is indeed observed. The related probit equation exhibits the following form:

$$\Pr(D_{i,j,t} = 1|\cdot) = \Pr(X_{S,i,j,t}\gamma \geqslant -\, V_{i,j,t}) = \Phi(X_{S,i,j,t}\gamma), \tag{19}$$

where Φ is the cumulative distribution function of the unit normal distribution. Equation (18) or its probit version, equation (19), are analogous to equation (11) in section 3.

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

$$E(Y_{i,j,t}|X_{F,i,j,t}, D_{i,j,t} = 1) = X_{F,i,j,t}\beta + E(U_{i,j,t}|X_{F,i,j,t}, D_{i,j,t} = 1). \tag{20}$$

Note that the last term, the conditional expectation of $U_{i,j,t}$, is not equal to zero and is dependent on $X_{F,i,j,t}$. This upsets the classical assumptions concerning regression functions for random samples.

To see this, one can substitute equation (18) into equation (20) to get:

$$E(Y_{i,j,t}|X_{F,i,j,t}, D_{i,j,t} = 1) = X_{F,i,j,t}\beta + E(U_{i,j,t}|V_{i,j,t} \geqslant -X_{S,i,j,t}\gamma). \tag{21}$$

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 (21) is equal to:

$$E(Y_{i,j,t}|X_{F,i,j,t}, D_{i,j,t} = 1) = X_{F,i,j,t}\beta + \rho \cdot \sigma_U \cdot \lambda_{i,j,t}, \tag{22}$$

where the inverse Mills ratio, $\lambda_{i,j,t}$ is defined by:

$$\lambda_{i,j,t} \equiv E(U_{i,j,t}|V_{i,j,t} \geqslant -X_{S,i,j,t}\gamma) = \frac{\phi(-X_{S,i,j,t}\gamma)}{1 - \Phi(-X_{S,i,j,t}\gamma)} = \frac{\phi(X_{S,i,j,t}\gamma)}{\Phi(x_{S,i,j,t}\gamma)}, \tag{23}$$

and where ϕ and Φ are the density and the cumulative of the unit normal distribution function, respectively. The bias term is equal to the partial derivative of the *conditional* expectations of $U_{i,j,t}$ with respect to $X_{F,i,j,t}$. That is:

$$bias = -\gamma \cdot \rho \cdot \sigma_U \cdot \delta_{i,j,t}, \tag{24}$$

where $\delta_{i,j,t}$ is a positive number[10].

[Figure 1 about here]

Figure 1 provides the intuition for the case where $\rho > 0$. Suppose, for instance, that $X_{F,i,j,t}$ measures the per-capita income differential between the ith source country and the potential jth host country, holding all other variables constant. Our theory predicts that the parameter β is positive. This is shown by the upward sloping line AB. Note that the slope is an estimate of the "true" marginal effect of $X_{i,j,t}$ on $Y_{i,j,t}$. But recall that flows could also be equal to zero if the setup costs are sufficiently high. A flow threshold, which is derived from decisions in the presence of setup costs, is shown as line TT' in Figure 1. However, if the econometrician employs only those country pairs for which $Y_{i,j,t}$ is positive the sub-sample is no longer random. As equation (18) makes clear, the *selection* of country pairs into the sub-sample depends on the vector $X_{F,i,j,t}$.To illustrate, suppose, that for high values of $X_{F,i,j,t}$ (say, X^H in Figure 1), (i, j) pair-wise

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

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

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

FDI flows are all positive. That is, for all pairs of countries in the subsample the flow threshold line is surpassed and the observed average for $X_{F,i,j,t} = X^H$ is also equal to the conditional population average, point R on line AB. However, this does not hold for low values of $X_{F,i,j,t}$ (say, X^L). For these (i,j) pairs, we observe positive values of $Y_{i,j,t}$ only for a subset of country pairs in the population. Point S is, for instance, excluded from the sub-sample of positive FDI flows. Consequently, we observe only flows between country pairs with low setup cost (namely with high $V_{i,j,t}$'s), for low $X_{F,i,j,t}$'s. As seen in Figure 1, the regression line for the subsample is the A'B' line, which underestimates the effect of per-capita income differentials on bilateral FDI flows.

4.1.2 Setup Cost Bias Vs. Measurement Errors

Most of the empirical literature developed after Tinbergen (1962) has either *omitted* pairs with no FDI flows, or treated reported zero flows as *measurement errors* was *literally* indicating *zero flows*[11]. This view ignores the existence of setup costs.[12].

In the stripped-down model of section 3, setup costs play an important role in determining whether a source country i invests directly in a host country j. Moreover, the model may be interpreted as implying that there could be a negative correlation between the error term, of the FDI flows equation and the error term of the selection equation. This implication of the model distinguishes between the "setup-cost model" and the "measurement errors hypothesis". Note that whereas the "measurement-errors hypothesis" is consistent only with a positive ρ, the "setup-cost model" is also consistent with a negative ρ. The Tobit method is typically used in the former, whereas the Heckman method is used the second.

[11] A notable exception in trade-based literature is Helpman, Melitz and Rubinstein (2004). Recently, Silva and Tenreyro (2003) proposed the Poisson pseudo-maximum likelihood estimator to deal with zero values in the bilateral trade models.

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

4.1.3 Tobit and Setup Costs

Previous empirical works on the determinants of FDI flows frequently makes use of the Tobit procedure. But this procedure, which is proper to handle measurement errors when negative values are not reported, collapses , in effect, the flow and selection equations into just one equation. In contrast, the Heckman (1979) selection procedure with the help of the two equations, which are jointly estimated, yield unbiased estimates of the two equations separately[13]. The Tobit model [see Tobin (1958)] has been often used in the empirical international trade literature [e.g., Carr, Markusen and Muskus (2001)]. This model is originally developed to deal with situations where negative, or small positive values of the dependent variable in the data are reported (censured) as zero values, thus artificially truncating the sample distribution. However, the Tobit model ignores setup costs that give rise to genuine zero values for the dependent variable as a result of selection decisions.

The Tobit method works as follows. Let $Y^*_{i,j,t}$ denote the desired FDI flows from i to j in period t:

$$Y^*_{i,j,t} = X_{F,i,j,t}\beta + U_{i,j,t}. \tag{25}$$

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

$$Y_{i,j,t} = \max(0, Y^*_{i,j,t}). \tag{26}$$

The population regression function for equation (7.1) is given by:

$$E(Y_{i,j,t}|X_{F,i,j,t}, D_{i,j,t} = 1) = X_{F,i,j,t}\beta + \sigma_U \cdot \tilde{\lambda}_{i,j,t}, \tag{27}$$

where

[13] See also Kyriazidou (1996).

$$\tilde{\lambda}_{i,j,t} = \frac{\phi(\frac{X_{F,i,j,t}}{\sigma_U})}{\Phi(\frac{X_{F,i,j,t}}{\sigma_U})}. \qquad (28)$$

Comparing equation (22) with equation (27), the Tobit model can is seen as a special case of the Heckman model, with $\rho = 1$. Therefore, in the Tobit procedure, the flow equation serves also as the selection equation (up to a scale), because the error terms of the two equations are perfectly correlated. Because the only difference between the selection and the flow equations is in the role of the fixed costs played by the setup costs, the Tobit model is a correct method under the null hypothesis of no setup costs, but it yields biased estimates in the presence of setup costs.

4.2 Endogeneity Issues

Although bilateral FDI flows are only a subset of the international capital flows that enter in the host countries from all sources, one cannot ignore the possibility that foreign direct investment flows from source country i to host country j may affect both economies. If such interdependence 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 both the flow and selection equations. 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. Lagged *negative* flows while rare in the data may have some bearing on the setup costs making new investments and, consequently, on the selection process. Our theory does not generate any prior about the time structure of the X_t time series. But we also estimate the full system using various time lags, as instruments.

5 Data and Country-Specific Variables

Data are drawn from OECD reports (OECD, various years) on a sample of 24 OECD countries, over the period from 1981 to 1998. The FDI data are based on the OECD reports of FDI exports from 17 OECD source countries to 24 OECD countries.

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 size, educational attainment (as measured by average years of schooling), language, financial risk rating, etc.; (2) (s, h) source-host characteristics, 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. Table B.1 describes the list of the 24 countries in the sample, and indicates for each country whether positive flows are observed in the sample, at least once, as a source or host country (but most source countries do not interact more than with few host countries). Table B.2 summarizes the data sources.

6 Estimation

Table 1 and Table 2 provide a "first look" at the direction and volume of FDI flows. Whereas source-host differences in GDP per capita look as good predictors of the direction of flows (the extensive margin; see Table 1), they are not correlated with the volume of FDI flows for the subset of country pairs with positive flows (the intensive margin; see Table 2).

We now turn to the estimation of the determinants of bilateral FDI flows. We consider several potential explanatory variables of the two-fold decisions on FDI flows. These variables include standard "mass" variables (the source and host population sizes); "distance" variables (physical distance between the source and host countries and whether or not the two countries share a common language); and "economic" variables (source and host GDP per capita, source-host differences in average years of schooling, and source and host financial risk rating). We also 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 (8), and simply estimate the gravity equation (1) twice: (i) by treating all FDI flows in (s, h) pairs with no recorded FDI flows as "zeros";[14] (ii) excluding country pairs with no FDI flows. The rationale for inserting "zeros" in the first benchmark case 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 negligible value as a common low value for the value of the FDI flows for the no-flows (s, h) pairs.[15] (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 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 recorded as zeros.) We present the results in Panel B of 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 setup costs, can now brought to the limelight when we employ the Heckman selection method. We jointly estimate the maximum likelihood of the flow (gravity) equation and the selection equation. The Heckman estimation method accommodates both measurement errors and a possible existence of setup costs.[16] 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; zero otherwise. Assuming that setup costs are lower if country i already in-

[14] 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.

[15] We choose this value to be the lowest observed flow between any (s, h) country pair in the sample.

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

vested in country j in the past, then $D_{i,j,t-k}$ could serve as an instrument in the selection equation (the exclusion restriction). The results are described in Panel C of Table 3.

Both OLS and Tobit estimations conform to the notion that the volume of FDI flows is not affected by deviations from long-run averages in the source and host countries. The coefficient of GDP per capita is not significant in Heckman selection equation.[17] Turn to the effect of the host country education level, relative to the source country counterpart. While cross-country educational gaps have no effect on the intensive margin (the flow equation), they do have a significant effect on the extensive margin (the selection equation). To test whether the effect on FDI flows is *non-linear*, we estimate the parameters of interest in the OLS method for different ranges of FDI flows. That is, the OLS regression in the first benchmark has different coefficient than in the OLS regression of the second benchmark. 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. Whereas the coefficient of the educational gaps is positive and significantly different from zero in the first column, the point estimate is substantially smaller and insignificant when we estimate the effect of educational attainments 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 and country-pairs with "true" positive flows rather than the variation among country-pairs with positive FDI flows.

The effect of the education variable on the extensive margin is also well reflected in our estimates using the Tobit and Heckman methods. We find significant effects in the two methods. However, whereas the Tobit method predicts that FDI flows are positively related to host-source difference in education levels, the Heckman method predicts that the education level affects 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 erroneously conclude that

[17] Recall that in the estimation we control for country fixed effects. In Table 5 we present also results of the estimation without controlling for country fixed effects.

cross-country educational gaps affect FDI volumes, whereas in Heckman estimation 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 host-country financial risk 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, the effect in the Heckman model 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 sharpened when we look at the effect of host country financial risk ratings. We find no effect 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" 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 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 selection variable $(D_{i,j,t-2})$, indicating whether exports of FDI in the past have been positive or zero, in panel C of Table 3 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

[18]Note from Tables 5, that without controlling for country fixed effects, the coefficient of source-country financial risk ratings 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.

[19]Note from Table 1 that without country fixed effects, the coefficient is significant.

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. The conclusions are similar to those presented in Table 3.

7 Evidence for Fixed Costs

The finding that there is a significant negative correlation (ρ) between the error terms in the gravity and selection equations indicates that the formation of an (s, h) pair of positive-FDI countries, and the size of the FDI flows between this pair of countries are not independent processes. A negative ρ is consistent with the setup costs hypothesis. If productivity shocks jointly drive marginal productivity of capital and setup costs of FDI, as in section 3, then shocks to the selection equation may be indeed negatively correlated with shocks to the flow equation. That is, above-average general productivity level in a host country, which may yield below-average likelihood of non-zero exports of FDI (because it may yield above-average setup costs), is also associated with above-average marginal productivity of capital, which yields above-average flow of FDI to the country (if new investment takes place at all); see section 3.

If education, as measured by the average years of schooling is indeed a “good” measure of host–source country differences in human capital, then education levels are important in predicting the volume of FDI flows. The Heckman estimation predicts 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.

8 Conclusion

The FDI flow mechanism 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 selection and flow equations, which our econometric study is able to detect.

To allow for the role played by the unobserved fixed setup costs, which is at the center stage of our model (see section 3), we employ the Heckman selection method. We jointly estimate the maximum likelihood of the volume of FDI flows (the gravity equation), and the selection of countries into source-host country pairs (the selection equation). Only if setup costs play an important role in determining whether a source country invests directly in a host country, we could expect a *negative* correlation between the error terms of the gravity and the selection 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.

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] Helpman, Elhanan, Mark Melitz, and Yona Rubinstein. (2004), "Trading Partners and Trading Volumes," Harvard University, 2004.

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

[12] 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.

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

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

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

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

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

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

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

9 Appendix A: Partial Equilibrium Effect of A Productivity Shock on FDI

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)	

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

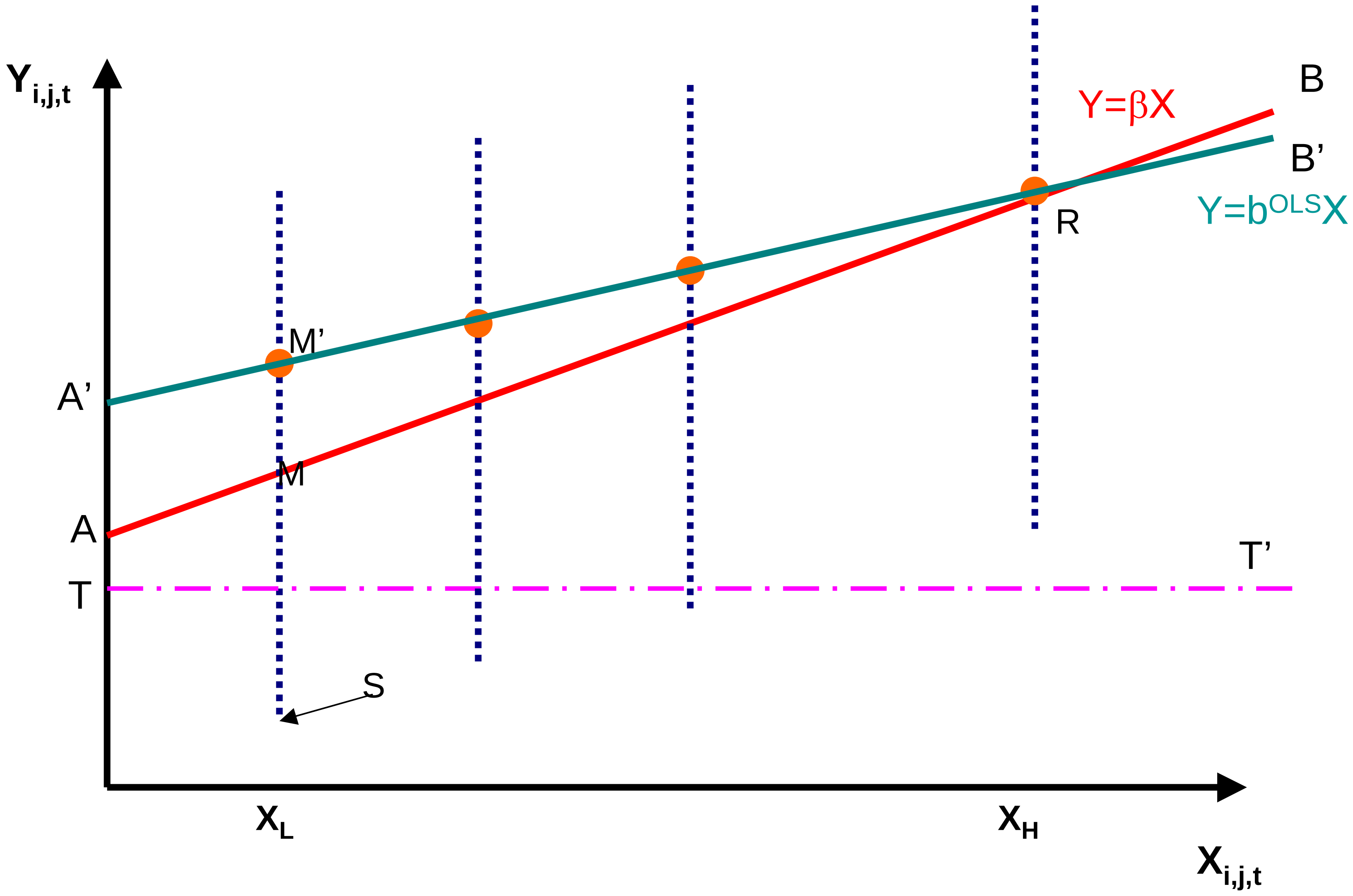

Table 1: Source-Host country pairs by GDP per capita

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 (0.987)	0.440 (0.690)	-0.287 (2.288)	-0.104 (1.171)	0.064 (1.016)	0.350 (0.682)	-0.475 (0.759)
GDP per capita - source^	-0.543 (0.796)	0.584 (0.580)	-0.460 (2.418)	-0.017 (1.232)	-0.104 (1.060)	0.581 (0.562)	-0.202 (0.845)
Difference between source and host years of schooling	**0.386** **(0.148)****	**0.012** **(0.097)**	**0.917** **(0.282)****	**0.438** **(0.145)****	**0.338** **(0.126)****	**-0.029** **(0.103)**	**0.288** **(0.102)****
Common language	**0.762** **(0.254)****	**1.014** **(0.146)****	**1.655** **(0.319)****	**1.217** **(0.162)****	**1.162** **(0.139)****	**0.965** **(0.129)****	**0.315** **(0.138)***
Distance (in logs)	**-0.836** **(0.139)****	**-0.674** **(0.095)****	**-1.572** **(0.187)****	**-1.013** **(0.095)****	**-0.909** **(0.082)****	**-0.629** **(0.092)****	**-0.393** **(0.091)****
Population - host^	6.794 (3.894)	-1.967 (2.384)	15.401 (7.756)*	5.460 (3.956)	3.237 (3.417)	-2.960 (2.393)	7.232 (2.626)**
Population - source^	5.395 (3.220)	-0.703 (3.032)	12.083 (9.102)	6.000 (4.659)	5.892 (4.050)	-1.536 (2.933)	2.828 (2.724)
Financial risk rating - host	-0.028 (0.027)	0.045 (0.017)**	-0.045 (0.061)	-0.005 (0.032)	0.007 (0.027)	0.050 (0.017)**	-0.029 (0.021)
Financial risk rating - source	-0.098 (0.024)**	-0.034 (0.026)	-0.245 (0.081)**	-0.141 (0.042)**	-0.120 (0.036)**	-0.025 (0.026)	-0.071 (0.025)**
Negative flows from I to j three years ago (=1 if yes)^^^	**0.661** **(0.423)**	**-0.169** **(0.152)**	**1.592** **(0.508)****	**0.610** **(0.257)***	**0.418** **(0.222)**	**-0.243** **(0.155)**	**0.505** **(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

Variables	Panel A: OLS		Panel B: Tobit Correction			Panel C: Heckman selection	
	Sample:		Low censored (in logs)			Equation:	
	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%