HONG KONG INSTITUTE FOR MONETARY RESEARCH

WHICH COUNTRIES EXPORT FDI, AND HOW MUCH?

Assaf Razin, Yona Rubinstein and Efraim Sadka

HKIMR Working Paper No.15/2004

August 2004

Which Countries Export FDI, and How Much?*

Assaf Razin†
Yona Rubinstein‡
and
Efraim Sadka§

August 2004

Abstract

The paper develops a model with "lumpy" setup costs, which govern the flow of bilateral foreign direct investment (FDI). Every country is potentially both a source for FDI flows to several host countries, and a host for FDI flows from several source countries. But technologically-advanced countries have a comparative advantage in setting up foreign subsidiaries. Thus, the model generates two-way, rich-rich and rich-poor, FDI flows. 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 either true zeros stemming from marginal productivity conditions, measurment errors, or true zeroes that are due to fixed costs (which dominate marginal productivity conditions). Previous empirical literature on the determinants of FDI flows imposes a no-fixed cost assumption on the estimation procedure (Tobit). In contrast, by employing the Heckman selection procedure, we show that the Tobit restriction is not consistent with the data, and yields biased estimates. Controlling for the selection into source-host pairs of countries, and for time and country fixed effects, we find: (1) FDI flows respond positively to advances in host country level of education relative to the source country level of education, whereas the source-country level of education is a predictor of the formation of source-host country pairs; (2) FDI flows respond positively to improvements in host country financial risk ratings relative to the source country ratings; (3) existence of rich-poor pairs hinge on surpassing an education-income threshold, whereas rich-rich FDI flow volumes depend on education and income levels.

* We wish to thank Elhanan Helpman, Gita Gopinath, Galina Hale, Anil Kashyap, Pierre-Olivier Gourinchas, Maury Obstfeld, Philip Lane, Jim Markusen, John Romalis and Guido Tabellini for many useful comments and suggestions. We benefitted from seminars in the University of Berkeley, the University of Boconni, the University of Chicago, Trinity College (Dublin), Tel-Aviv University and the Hebrew University. We are indebted to Ashoka Mody, IMF, for developing a comprehensive data set used in this paper.

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

‡ Tel-Aviv University

§ Tel-Aviv University and CESifo.

1. Introduction

In an influential paper, Lucas (1990) asks: "Why doesn't capital flow from rich to poor countries?" Indeed, the law of diminishing returns implies that the marginal product of capital is high in poor countries and low in rich countries. With standard constant-returns-to-scale production functions, when the wage (per efficiency unit of labor) is higher in a rich country than in a poor country (due to poor-rich country differences in relative supplies of capital and labor), then the return to capital must be lower in the rich country than in the poor country. Therefore, capital is expected to flow from rich to poor countries. In practice, however, this is hardly the case.[1] Even though barriers to international capital mobility are by and large being eliminated, the wage gap is still in force, and migration quotas from poor to rich countries have to be enforced.[2] Lucas reconciled this paradox (in theory and skillful calibration) by appealing to a human capital *externality* that generates a Hicks-neutral productivity advantage for rich countries over poor countries.

Motivated by the Lucas paradox, we focus on foreign direct investment (FDI), as a key channel of international capital flows, which is expected to be closely associated with international productivity differences; more closely than foreign portfolio investment. We develop a model with "lumpy" setup costs of new investments that govern the flow of bilateral foreign direct investment. Every country in this model may potentially be a source for FDI flows to several host countries; and each country may be a host for FDI flows from several source countries. But rich and technologically-advanced countries have a comparative advantage in setting up foreign subsidiaries. The model may generate two-way rich-rich, and rich-poor FDI flows.

With setup costs of investment, it does not pay a firm to make a "small" foreign investment, even though such an investment is called for by marginal productivity conditions (that is, the standard first-order conditions for profit maximization). Put differently, the foreign investment decision is two-fold now: marginal productivity conditions determine how much to invest, whereas a "participation" condition determines whether to invest at all. In such a framework, it is also possible to have a coexistence of equal rates of return to capital and a wage gap.[3]

Our sample consists of 24 OECD countries over the period 1981-1998.[4] When one looks at data on international capital flows of FDI, one is immediately struck by the lack of flows from many source

---

1 Maurice Obstfeld and Alan M. Taylor (2003) note: "A century ago, world income and productivity levels were far less divergent than they are today, so it is all the more remarkable that so much capital was directed to countries at or below the 20 percent and 40 percent income levels (relative to the United States). Today, a much larger fraction of the world's output and population is located in such low-productivity regions, but a smaller share of global foreign investment reaches them."

2 Note also that despite the expansion of international trade in goods, still the Stolper-Samuelson (1941) factor price equalization theorem does not manage to eliminate the wage gap.

3 The international trade literature in appeals often to fixed costs. Fixed costs play a very important role in determining the extent of trade-based foreign direct investment. References include Zhang and Markusen (1999), Carr, Markusen and Maskus (2001) and Helpman, Melitz, and Yeaple (2003).

4 Our original sample had 45 countries, both developed and developing countries. However, data on exports of FDI from non-OECD countries are not available in the OECD data set we use. But data on imports of FDI by these non-OECD countries are available. Therefore, the Appendix extends our estimation to the full 45 country sample.

countries to many host countries. Only 12 countries are a source for FDI outflows, and each one of them exports FDI to only a few host countries. This first-look finding provides a *prima facia* evidence for the existence of fixed setup costs of investment that shut off the potential of "small" capital flows, which may have been called for by marginal productivity conditions.

Whether a potential source-host country pair becomes "active" or "inactive", and how much flow is recorded between this pair of countries, in case it becomes "active", are jointly and simultaneously determined. Thus, the selection of pairs of countries into "active" and "inactive" pairs should not be assumed to be exogenous. Indeed, if one treats this selection as exogenous, the estimates of the determinants of FDI flows are biased. We therefore employ a Heckman selection-bias method in order to simultaneously estimate the determinants of FDI flows and the selection of countries into source-host pairs.[5]

The organization of the paper is as follows. Section 2 explains the simple way Lucas reconciles the paradox of the inadequacy of capital flows from rich to poor countries. Section 3 presents our model of fixed setup costs of foreign direct investment. This model is used in Section 4 to provide an alternative theoretical reconciliation for the Lucas paradox. Section 5 presents the econometric approach. The data are described in Section 6. A first-look examination of the potential for a selection bias in the data is described in Section 7. Estimation results of the determinants of FDI flows, and whether source-host flows are formed at all, are presented in Section 8. The results are interpreted in Section 9. Section 10 reports on findings for an extended OECD and non-OECD country sample. Conclusions are drawn in Section 11.

2. The Lucas Paradox

The widespread pressure of migration from poor to rich countries is undoubtedly indicative of a higher marginal productivity of labor in rich relative to poor countries (over and above the attractiveness of the rich welfare states to migrants from poor countries). However, *ceteris paribus*, a relatively lower marginal product of labor is usually associated with a relatively high marginal product of capital. In the wake of globalized capital markets, capital should flow from rich to poor countries so as to mitigate these differentials in marginal productivity of capital, and also of labor, assuming constant-returns-to scale and identical technologies (via globalization). This is the essence of the Lucas paradox.

Lucas (1990) employs a standard constant-returns-to-scale production function:

$$Y = AF(K, L), \tag{1}$$

where Y is output, K is capital and L is **effective** labor. The latter is used in order to allow for differences in the human capital content of labor between developed and developing countries. The parameter A is a productivity index which may reflect the average level of human capital in the country, external to the firm as in Lucas (1990). In addition, parameter A may reflect the stock of public capital (roads and other infrastructure) that is external to the firm. In per effective-labor terms, we have:

5 See Heckman (1979).

$$y \equiv Y/L = AF(K/L, 1) \equiv Af(k). \tag{2}$$

The return to capital is:

$$r = Af'(k), \tag{3}$$

whereas the wage per effective unit of labor is:

$$w = A[f(k) - kf'(k)]. \tag{4}$$

Let a variable subscripted by "*S*" stand for a rich (source) country and a variable subscripted by "*H*" for a poor (host) country. The function f is common to all countries. Initially (before capital is freed to flow from one country to another), $r_S^0 < r_H^0$. But when capital can freely move from rich to poor countries, then rates of return are equalized, so that:

$$r_S = A_S f'(k_S) = A_H f'(k_H) = r_H. \tag{5}$$

Lucas explains the paradox by appealing to a human-capital externality. This externality makes A_S larger than A_H. Hence, it follows from equation (5) that $k_S > k_H$ (because of a diminishing marginal product of capital). Therefore, employing equation (4), it follows that $w_S > w_H$. Moreover, Lucas was able to simulate the observed difference between k_H and k_S and between w_H and w_S by calibrating the difference between A_H and A_S.

Thus, at the calibrated equilibrium, workers can earn higher wages (per effective labor) in the rich country than in the poor country, and administrative means (migration quotas) are employed to impede the flow of labor from poor to rich countries. Yet there is no pressure on capital to flow in the opposite direction because rates-of-return on capital are already equalized.

Lucas focuses on capital flows from rich to poor countries. This was a major direction of flows in the era of free capital mobility that preceded World War I.[6] However, the increased mobility of capital that followed World War II and accelerated with the end of the Cold War is of a different nature: "globalized capital markets are back, but with a difference: capital transactions seem to be mostly a rich-rich affair, a process of 'diversification finance'rather than 'development finance'. The creditor-debtor country pairs involve more rich-rich than rich-poor, and today's foreign investment in the poorest developing countries lag far behind the levels attained at the start of the last century" [Obstfeld and Taylor (2003)].

In essence, the driving force in Lucas analysis of capital flows is cross-country differences in marginal productivity of capital. We supplement this *marginal* analysis with *average* analysis.[7] This marginal-average analysis can generate both rich-poor and rich-rich FDI related transactions.

---

6 Most notable were the flows of capital from imperialist countries to their colonies.

7 Capital flows are governed not only by cross-country differences in marginal productivities of capital but also by cross-country differences in average productivity of capital.

3. Lumpy Adjustment Cost of Investment

We employ a "lumpy" adjustment cost for new investment, in the form of a fixed setup cost of investment. This specification, which has been recently supported empirically by Caballero and Engel (1999, 2000), creates economies of scale in investment.[8]

Consider again a pair of countries, "host" and "source", in a world of free capital mobility which fixes the world rate of interest, denoted by r. We will now describe the host country, whose economic variables will be subscripted by "*H*". The description of the source country is similar with a subscript "*S*". Variables that are not subscripted are identical for the two countries. There is a single industry whose product serves both for consumption and investment. For simplicity suppose that existing firms will last for two periods. In the first period there exists a continuum of N_H firms which differ from each other by a productivity index ε. We denote a firm which has a productivity index of ε by an ε-firm. The cumulative distribution function of ε is denoted by $G(\cdot)$, with a density function $g(\cdot)$.

We assume for simplicity that the initial net capital stock of each firm is the same and denote it by $(1-\delta)K_H^0$. This consists of the net initial stock, K_H^0, of the preceding period, multiplied by one minus the depreciation rate, δ. If an ε-firm invests I in the first period, it augments its capital stock to $K=(1-\delta)K_H^0+I$ and its gross output in the second period will be $A_H F(K,L)(1+\varepsilon)$, where L is the labor input (in effective units). Naturally, $\varepsilon \geq -1$ so. so that $G(-1)=0$.

We assume that there exists a fixed setup cost of investment, C_H, which is the same for all firms (that is, independent of ε): In order for the firm to be able to incur such a cost, we no longer assume that *F* exhibits constant returns to scale. We assume instead that, due to some (suppressed) fixed factor, *F* exhibits diminishing returns to scale in *K* and *L*. Consider an ε-firm which does invest in the first period an amount $I=K-(1-\delta)K_H^0$ in order to augment its stock of capital to K. Its present value becomes:

$$V^+(A_H, C_H, K_H^0, \varepsilon, w_H, r) = \underset{(K,L)}{Max}\left\{\frac{A_H F(K,L)(1+\varepsilon) - wL + (1-\delta)K}{1+r} - [K-(1-\delta)K_H^0 + C_H]\right\} \quad (6)$$

The demands of such a firm for K and L are denoted by $K^+(A_H, \varepsilon, w_H, r)$ and $L^+(A_H, \varepsilon, w_H, r)$ and given by the marginal productivity conditions:

$$A_H F_K(K,L)(1+\varepsilon) = r+\delta, \quad (7)$$

[8] Economies of scale 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.

and

$$A_H F_L(K, L)(1+\varepsilon) = w_H. \tag{8}$$

Note, however, that an ε-firm may chose not to invest at all $\left[\right.$that is, to stick to its existing stock of capital $(1-\delta)K_H^0)\left.\right]$ and avoid the setup lumpy cost C_H. In this case its labor input, denoted by $L^-(A_H, K_H^0, \varepsilon, w_H)$ is defined by:

$$A_H F_L[(1-\delta)K_H^0, L](1+\varepsilon) = w_H. \tag{9}$$

Note that L^- depends on the initial stock of capital. Naturally, a firm with a low ε may not find it worthwhile to incur the setup cost C_H. In this case its present value is:

$$V^-\left(A_H, K_H^0, \varepsilon, w_H\right) = \underset{L}{Max}\{\frac{A_H F\left[(1-\delta)K_H^0, L\right](I+\varepsilon) - w_H L + (1-\delta)^2 K_H^0}{(1+r)}\}. \tag{10}$$

Therefore, there exists a cutoff level of ε, denoted by ε_0, such that an ε-firm will make a new investment if and only if $\varepsilon > \varepsilon_0$. This cutoff level of ε depends on $A_H,\ C_H,\ K_H^0,\ w_H$ and r. We write it as $\varepsilon_0(A_H, C_H, K_H^0, w_H, r)$ and define it implicitly by:

$$\begin{aligned} & V^+\left(A_H, C_H, K_H^0, \varepsilon_0, w_H, r\right) \\ = \ & V^-\left(A_H, K_H^0, \varepsilon_0, w_H\right). \end{aligned} \tag{11}$$

Note that as the setup cost rises, fewer firms will choose to make new investments, that is $\varepsilon_0(\cdot)$ is increasing in C_H. We continue to assume that labor is confined within national borders. Denoting the country's endowment of labor in effective units by $\tilde{L}_H^0$, we have the following labor market clearance equation:

$$N_H \int_{-1}^{\varepsilon_0(A_H,C_H,K_H^0,w_H,r)} L^-\left(A_H, K_H^0, \varepsilon, w_H\right) g(\varepsilon)d\varepsilon + N_H \int_{\varepsilon_0(A_H,C_H,K_H^0,w_H,r)}^{\bar{\varepsilon}} L^+(A_H, \varepsilon, w_H, r) g(\varepsilon) d\varepsilon = \tilde{L}_H^0,$$

where $\bar{\varepsilon}$ is the upper productivity level. Dividing the latter equation through by N_H, yields:

$$\int_{-1}^{\varepsilon_0(A_H,C_H,K_H^0,w_H,r)} L^-\left(A_H, K_H^0, \varepsilon, w_H\right) g(\varepsilon)d\varepsilon + \int_{\varepsilon_0(A_H,C_H,K_H^0,w_H,r)}^{\bar{\varepsilon}} L^+\left(A_H, \varepsilon, w_H, r\right) g(\varepsilon) = L_H^0, \tag{12}$$

where $L_H^0 \equiv \tilde{L}_H^0 / N_H$ is the effective labor per firm.

Note that no similar market clearance equation is specified for capital, as we continue to assume that capital is freely mobile internationally and its rate of return is equalized internationally. The same description with the subscript "*S*" replacing "*H*" holds for the source country.

Note that differences in labor abundance between the two countries are manifested in the wage differences. To see this, suppose that the two countries are identical, except that effective labor per firm is more abundant in the poor-host country than in the rich-source country, that is: $L_H^0 > L_S^0$. If wages were equal in the two countries, then effective labour demand per firm would be equal and the market clearing condition [equation (12)] could not hold for both countries. Because of the diminishing marginal product of labor, it follows that the wage in the relatively labor-abundant country is lower than in the relatively labor-scarce country, that is: $w_H < w_S$[9]. Thus, equal returns to capital (through capital mobility) coexist with unequal wages, as in Lucas (1990).[10]

4. Foreign Direct Investors: Expansion and Greenfield Investments

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 C_H^* which is below C_H, 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. Each such firm $\left[\text{that is, each firm whose } \varepsilon \text{ is above } \varepsilon_0(A_H, C_H^*, K_H^0, w_H, r)\right]$ is purchased at its value, which is $V^+\left(A_H, C_H^*, K_H^0, \varepsilon, w_H, r\right)$. 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 $K^+\left(A_H, \varepsilon, w_H, r\right) - (1-\delta)K_H^0$ in the firm. Thus, aggregate foreign direct investment is equal to:

$$FDI = \int_{\varepsilon_0(A_H, C_H^*, K_H^0, w_H, r)}^{\bar{\varepsilon}} \left\{V^+(A_H, C_H^*, K_H^0, \varepsilon, w_H, r) + K^+(A_H, \varepsilon, w_H, r) + C_H^* - (1-\delta)K_H^0\right\} g(\varepsilon)d\varepsilon \tag{13}$$

Suppose first that w_H is fixed. Note that it follows from equation (6) that $\frac{\partial V^+}{\partial\left[(1-\delta)K_H^0\right]} = 1$, by the envelope theorem. Therefore $\frac{\partial(FDI))}{\partial\left[(1-\delta)K_H^0\right]} = 0$, by equation (13). Thus, the amount of FDI in a firm whose ε is above ε_0 does not depend on the initial capital stock $\left[(1-\delta)K_H^0\right]$: an increase of \$1 in the initial stock of capital of such a firm increases the value of the firm by \$1,[11] but decreases the required new investment by the same amount, so that FDI does not change. However, the aggregate amount of FDI diminishes, when the initial stock of capital $(1-\delta)K^0$ rises. This is because fewer firms will make new investment

9 The equilibrium wage gap implies that the poor country employ more workers per firm than the rich country. Thus, even though the productivity distribution across firms is assumed equal, the rich country in equilibrium is effectively more productive.

10 See also Amiti (1998) who studies the effect of agglomeration effects on cross-regional wage differences. See also Melitz (2003) for the role of fixed costs in intra-industry reallocations in reaction to industry productivity shocks.

11 This is because, in the absence of a marginal adjustment cost of investment, the marginal Tobin's q is identically equal to one.

and be purchased by foreign direct investors, that is, the cutoff ε_0 rises when $(1-\delta)K^0$ rises. To see this, differentiate equation (13) with respect to K^0 to get:

$$\begin{aligned} \frac{\partial(FDI)}{\partial K_H^0} &= (1-\delta)-(1-\delta) \\ &\quad \{-\{V^+\left[A_H, C_H^*, K_H^0, \varepsilon_0(A_H, C_H^*, K_H^0, w_H, r), w_H, r\right] \\ &\quad +K^+\left[A_H, \varepsilon_0(A_H, C_H^*, K_H^0, w_H, r), w_H, r\right] \\ &\quad +C_H^* \\ &\quad -(1-\delta)K_H^0\}\cdot g[\varepsilon_0(A_H, C_H^*, K_H^0, w, r)]\frac{\partial\varepsilon_0}{\partial K_H^0} \\ &< 0 \end{aligned} \tag{14}$$

because, by equation (11), we have $\frac{\partial\varepsilon_0}{\partial K_H^0} > 0$.

Similarly, it follows from equation (12) that:

$$\begin{aligned} \frac{\partial(FDI)}{\partial C_H^*} &= \frac{\partial V^+}{\partial C_H^*}+1 \\ &\quad -\{V^+\left[A_H, C_H^*, K_H^0, \varepsilon_0(A_H, C_H^*, K_H^0, w_H, r), w_H, r\right] \\ &\quad +K^+\left[A_H, \varepsilon_0(A_H, C_H^*, K_H^0, w_H, r), w_H, r\right] \\ &\quad +C_H^* \\ &\quad -(1-\delta)K^0\}\cdot g[\varepsilon_0(A_H, C_H^*, K_H^0, w, r)]\frac{\partial\varepsilon_0}{\partial C_H^*} \end{aligned} \tag{15}$$

Because $\frac{\partial V^+}{\partial C_H^*} = -1$, by equation (6), and $\frac{\partial\varepsilon_0}{\partial C_H^*} > 0$, by equation (11), it follows that $\frac{\partial(FDI)}{\partial C_H^*} < 0$.

Thus, a lower level of an initial stock of capital in the host country attracts more foreign direct investment. Similarly, a lower level of the setup cost of investment in the host country for the FDI investors promotes more FDI. These conclusions were drawn under the assumption that the wage (w_H) in the host country is fixed. When it is not fixed, then lower K_H^0 and/or C_H^* attract more FDI and push the wage upward, thereby mitigating the initial increase in FDI but not eliminating it altogether.

Observe that FDI flows constitute only a part of the international capital transactions of the host and source countries. In a globalized world capital market, where the world rate of interest is given to our pair of countries, domestic saving and domestic investment are not equal to each other and neither FDI is equal to either saving or investment.

So far, FDI has taken the form of acquisitions of existing firms. Consider now the possibility of establishing a new firm (that is, a greenfield FDI where $K^0 = 0$). Suppose that the newcomer entrepreneur does not know in advance the productivity factor (ε) of the potential firm. She therefore takes G (·) as the cumulative probability distribution of the productivity factor of the new firm. However, ε is revealed to the new owner before she decides whether or not to make a new investment. The expected value of the new firm is therefore

$$V(A,C,w,r)=\int_{-1}^{\bar{\varepsilon}} Max\left\{V^{+}(A,C,\varepsilon,w,r),0\right\} g(\varepsilon)d\varepsilon. \tag{16}$$

Note when K^0 is equal to zero, only the firms with an ε high enough to justify a greenfield investment have a positive value. This explains equation (16).

Now, suppose that greenfield entrepreneurship is in limited supply and capacity. An entrepreneur in a source country (and there is a limited number of them) may have to decide whether to establish a new firm at home (the source country) or abroad (the host country). Her decision is determined by where V(·), as defined in equation (16), is higher. She will invest in the host country rather than in the source country if, and only if,

$$V\left(A_H, C_H^{*}, w_H, r\right) > V\left(A_S, C_S, w_S, r\right), \tag{17}$$

Naturally, the lower wage rate in the host country is a pull factor for that country, that is, it works in the direction of satisfying condition (17). Thus, the lower wage rate in the host country attracts greenfield FDI. On the other hand, the total factor productivity in the source country (namely, A_S), is expected to be higher than its counterpart in the host country (namely, A_H) and this discourages FDI. Assuming that the wage differential dominates the total factor productivity differences, the poor-host country attracts greenfield FDI from the rich-source country.

Assuming that newcomer entrepreneurs evolve gradually over time and that technology spillover equates total factor productivity, eventually this process may end up with full factor price equalization. Naturally, the capital-labor ratios and $L \equiv \tilde{L}/N$ are equalized in this long-run steady state. All this happens even though labor is not internationally mobile. The establishment of new firms in the global economy may be an engine for FDI flows by multinationals.

Our two-country model, which generates capital flows from the rich-source to the poor-host country, can be extended to explain two-way FDI flows. By assuming more than one industry, we may have two-way flows between two rich countries, when each country has a setup-cost advantage in a different industry.

5. The Econometric Approach

The preceding section presents a model of bilateral foreign direct investment flows distin-guished by setup costs of investment.[12] FDI takes place either in order to acquire existing firms and investing in them (Mergers and Acquisitions FDI), or to establish new firms (Greenfield Investment FDI). Our empirical

[12] For early works with gravity models for trade in goods, see Eaton and Tamura (1994) and Eichengreen and Irwin (1998). Helpman, Melitz, and Yeaple (2003) pose the question of how a source country can simultaneously make both FDI and exports 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.

investigation is in the tradition of the often used gravity models,[13] but with adjustments for a selection bias of 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 much 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 is the 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: FDI flows from source country j to host country i are denoted by $Y_{i,j,t}$: Note that with this notation, $Y_{i,j,t}$ is always non-negative. But, it may well be zero, because typically, in a global economy, there are only a few countries which significantly export FDI.

The existence of a setup cost of investment makes investment "lumpy". This means that the conventional determinants of FDI flows (such as rate of return differentials) have to be strong enough in order to generate a large FDI flow that surpasses a certain *unobserved* threshold. Otherwise, the observed FDI flow is practically zero. We argue that the sub-sample of FDI source countries is not a random sample of the country's 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 and correct for selection bias in the analysis of FDI flows.[14]

5.1 The Selection 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 1974, and Heckman 1979).

To simplify, but without losing generality, let us assume that, in an imaginary world with *no setup costs*, potential FDI flows exhibit the following linear form:

$$Y_{i,j,t} = X_{F,i,j,t}\beta + U_{F,i,j,t}, \tag{18}$$

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). Such variables are, for example, per-capita income differencess between country

[13] 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 to the distance (physical or other such as tari 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 more than proportionately with both the source and the host-country populations.

[14] 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 the Heckman method to study the effects of corruption on FDI in transition economies.

i and country j (reflecting differences in the capital-labor ratio), as well as language, geographical distance and communication/transportation costs. The vector β is 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 the persistent gap between the (marginal) productivity of capital in the i source and the j host countries $(\varepsilon_{i,j})$, and (ii) a random shock term, which is $i-j$ pairwise-specific $(\eta_{i,j,t})$, reflecting both deviations from the "long-run" productivity gap $(\Delta\varepsilon_{i,j,t})$ as well as other fluctuations in macroeconomic policy, political events, etc., that are unique to the $i-j$ pair. Note that according to our theory, the larger the unobserved (to the econometrician) productivity gap between the source country and the host country is, the lower is also the FDI volume. Hence, both $\theta_{i,j}$ and $\eta_{i,j,t}$ are decreasing functions of $\varepsilon_{i,j}$ and $\Delta\varepsilon_{i,j,t}$ respectively. That is:

$$U_{F,i,j,t} = \theta_{i,j}\left(\varepsilon_{i,j}\right) + \eta_{i,j,t}\left(\Delta\varepsilon_{i,j,t}\right), \tag{19}$$

Let $Z_{i,j,t}$ be a *latent* variable, indicating the profit (payoff, hereafter: π) from the direct investment made in host country j, by a firm in the source country i ; in period t. Following our model we allow $Z_{i,j,t}$ to be determined by unobserved (to the econometrician) setup costs $(C_{i,j,t})$, in addition to $X_{S,i,j,t}$ (and $i-j$ pairwise-specific random shocks). Note that $X_{S,i,j,t}$ is a vector which includes $X_{F,i,j,t}$ and a vector of instruments, that is, observed variables which are associated with setup cost, and therefore with the selection of pairs into $s-h$ pairs (hence the S subscript), but not with the marginal productivity of capital. The setup costs are the sum of two elements: (a) time-invariant costs $(C_{i,j})$, reflecting the persistent features of the $i-j$ pair, such as the "long-run" differences between the (marginal) productivity of capital in the i source and the j host countries $(\varepsilon_{i,j})$, and (b) the time-variant setup costs $(c_{i,j,t})$. We assume that the vector $Z_{i,j,t}$; as a function of $X_{S,i,j,t}$, exhibits the following linear form:

$$Z_{i,j,t} = X_{S,i,j,t}\gamma + U_{S,i,j,t}, \tag{20}$$

where the vector γ is the ceteris paribus effect of $X_{S,i,j,t}$ on $Z_{i,j,t}$. Note that in the case where $X_{S,i,j,t}$ affects the value of maximized profit, the same way as it influences $Y_{i,j,t}$ in equation (18) ; then = $\gamma = \beta$. The error term $V_{i,j,t}$, in the profit equation, is a composite of (i) the unobserved time invariant setup costs $(C_{i,j})$ and (ii) a random shock term which is $i-j$ pairwise-specific $(\nu_{i,j,t})$, reflecting both transitory deviations from the "long-run" setup costs $(c_{i,j,t})$ as well as fluctuations in macroeconomic policy, political events, etc., that are unique to the $i-j$ pair $(\Delta\varepsilon_{i,j,t})$.

The larger the productivity gap between the source and the host country is, the lower is also the relative cost of investment in the source country. This also holds for setup costs. If source country i is more productive than host country j then it is less expensive for a country i entrepreneur to establish a new firm in a host country j, than for a country j entrepreneur. Following this logic, the larger the productivity gap is, the lower are also the setup costs. That is:

$$U_{S,i,j,t} = C_{i,j}\left(\varepsilon_{i,j}\right) + \nu_{i,j,t}\left(\Delta\varepsilon_{i,j,t}\right), \tag{21}$$

where, $C'_{i,j}(\varepsilon_{i,j}) < 0$. We further assume that, for a *random* sample, the classical assumptions regarding the error term do hold. In particular, we assume that:

$$\begin{aligned} E\left(U_{F,i,j,t} \mid X_{F,i,j,t}\right) &= E\left(U_{F,i,j,t}\right) = 0, \\ &and \\ E\left(U_{S,i,j,t} \mid X_{S,i,j,t}\right) &= E\left(U_{S,i,j,t}\right) = 0, \end{aligned} \tag{22}$$

and that:

$$\begin{aligned} U_{F,i,j,t} &\sim N\left(0,\ \sigma^2_{U_F}\right), \\ U_{S,i,j,t} &\sim N\left(0,\ \sigma^2_{U_S}\right). \end{aligned} \tag{23}$$

We further assume that the cross-equation error terms $U_{F,i,j,t}$ and $U_{S,i,j,t}$ follow a bivariate normal distribution:

$$\left(U_{F,i,j,t},\ U_{S,i,j,t}\right) \sim N\left(0, \Omega\right),$$

with respective variances σ^2_U and σ^2_V :

$$\Omega = \begin{vmatrix} \sigma^2_{U_F} & \rho \cdot \sigma_{U_F} \cdot \sigma_{U_S} \\ & \sigma^2_{U_S} \end{vmatrix}, \tag{24}$$

whereas, ρ is the correlation between the cross-equation error terms.

Now, according to our model, FDI flows $(Y_{i,j,t})$ are positive, if and only if $Z_{i,j,t} \geq 0$. Denote by a binary $D_{i,j,t}$, whether or not country i exports positive FDI flows to country j at time t (the FDI participation variable) by:

$$D_{i,j,t} = \left\{ \begin{array}{lll} 1 & if & Z_{i,j,t} \geq 0 \\ 0 & otherwise. & \end{array} \right\}. \tag{25}$$

Note that whereas $Z_{i,j,t}$ *is not* observed, the binary variable $Z_{i,j,t}$ is indeed observed. Assuming that the errors in the underlying latent equation are distributed normally:

$$U_{S,i,j,t} \sim N\left(0,\ \sigma^2_{U_S}\right),$$

then the probability setup for the probit equation exhibits the following form.

$$\Pr(D_{i,j,t} = 1 \mid \cdot) = \Phi(X_{S,i,j,t}\gamma). \tag{26}$$

where Φ is the cdf of the unit normal distribution.

5.1.1 Setup Costs and Selection bias

The (statistical) population-regression function for equation (18) is:

$$E\left(Y_{i,j,t} \mid X_{F,i,j,t}\right) = X_{F,i,j,t}\beta. \tag{27}$$

Many previous studies aimed at estimating the effects of X on Y in the context of international capital mobility (and also, similarly, in the context of goods mobility through international trade) typically ignore the effect of unobserved setup costs on the observed capital flows. However, the regression function for the *sub-sample* of countries, for which we do indeed observe positive FDI flows is :

$$E\left(Y_{i,j,t} \mid X_{F,i,j,t},\ D_{i,j,t} = 1\right) = X_{F,i,j,t}\beta + E\left(U_{F,i,j,t} \mid D_{i,j,t} = 1\right) \tag{28}$$

Note that the last term, the conditional expectation of $U_{i,j,t}$ is no longer equal to zero. Furthermore, the term $E\left(U_{F,i,j,t} \mid X_{F,i,j,t},\ D_{i,j,t} = 1\right)$ depends on $X_{F,i,j,t}$, unlike the classical assumptions concerning regression functions applied to random samples.

To see this more clearly, one can substitute equations (19) and (25) into (28), to get:

$$E\left(Y_{i,j,t} \mid X_{F,i,j,t},\ D_{i,j,t} = 1\right) = X_{F,i,j,t}\beta + E\left(U_{F,i,j,t} \mid U_{S,i,j,t} \geq -X_{S,i,j,t}\gamma\right) \tag{29}$$

If $U_{F,i,j,t}$ and $U_{S,i,j,t}$ are correlated then the mean of the error term in the flow equation (18) will be a function of the error term in the profit equation (20). Consequently, since $X_{F,i,j,t}$ affects both $Y_{i,j,t}$ and $Z_{i,j,t}$, at the same time that $U_{F,i,j,t}$ and $U_{S,i,j,t}$ are correlated, a specification bias will affect the OLS estimates of β. If $U_{F,i,j,t}$ and $U_{S,i,j,t}$ follow a bivariate normal distribution with correlation ρ and with respective variances $\sigma^2_{U_F}$ and $\sigma^2_{U_S}$ then the expected volume of FDI flows from the source country i into the host country j in equation (29) equals:

$$E\left(Y_{i,j,t} \mid X_{F,i,j,t},\ D_{i,j,t} = 1\right) = X_{F,i,j,t}\beta + \rho \cdot \frac{\sigma_{U_F}}{\sigma_{U_S}} \cdot \lambda_{i,j,t}. \tag{30}$$

The $\lambda_{i,j,t}$, the inverse Mill's ratio, equals:

$$\lambda_{i,j,t} \equiv E\left(U_{S,i,j,t} \mid U_{S,i,j,t} \geq -X_{S,i,j,t}\gamma\right) = \frac{\phi\left(\frac{-X_{S,i,j,t}\gamma}{\sigma_{U_S}}\right)}{1-\Phi\left(\frac{-X_{S,i,j,t}\gamma}{\sigma_{U_S}}\right)} = \frac{\phi\left(\frac{X_{S,i,j,t}\gamma}{\sigma_{U_S}}\right)}{\Phi\left(\frac{X_{S,i,j,t}\gamma}{\sigma_{U_S}}\right)}, \tag{31}$$

where ϕ and Φ are the unit normal density and distribution functions respectively. The bias term, in the population, equals the partial derivative of the *conditional* expectation of U with respect to X :

$$bias = \frac{\partial E\left(U_{F,i,j,t} \mid U_{S,i,j,t} \geq -X_{S,i,j,t}\gamma\right)}{\partial x_{F,i,j,t}} = \rho \cdot \frac{\sigma_{U_F}}{\sigma_{US}} \cdot \frac{\partial \lambda_{i,j,t}}{\partial x_{F,i,j,t}}$$

or:

$$bias = -\gamma \cdot \rho \cdot \frac{\sigma_{U_F}}{\sigma_{US}} \cdot \delta_{i,j,t}, \tag{32}$$

where $\delta_{i,j,t}$ is a positive number.[15]

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 i th source country and the j th potential host country, holding all other variables constant, namely per-capita income differencess between the i th 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 $X_{F,i,j,t}$ on $Y_{i,j,t}$. But, recall that flows could be equal to zero if the setup 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 $X_{i,j,t}$ is positive. This sub-sample is, therefore, no longer random . Moreover, as equation (29) makes clear the *selection* of country pairs into this sub-sample depends on the vector $X_{F,i,j,t}$.

To see this, suppose, for instance, that for high values of $X_{F,i,j,t}$ (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 $X_{F,i,j,t} = X^H$ is also equal to the conditional population average, point R on the line AB. However, this does not hold for low values of $X_{F,i,j,t}$ (denoted by X^L). For those $i-j$ pairs we observe positive values of $X_{i,j,t}$ 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 $U_{S,i,j,t}$), among those with low $X_{F,i,j,t}$. As seen in Figure 1, the observed conditional average is at 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.

5.1.2 Two hypotheses for a selection bias

There is a long tradition in the international economics literature of log-linearizing the capital flow gravity statistical models, and estimating the parameters of interest by ordinary least squares. In these statistical models the gravitational force can be very small, but not zero, whereas trade between $i-j$ pairs of countries is quite often zero. Most of the empirical literature developed after Tinbergen (1962) has either

[15] Let $\alpha = -X_{S,i,j,t}\gamma/\sigma_{U_S}$; then the partial derivative of the inverse Mils ratio is:

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

and $\delta_{i,j,t} > 0$

ignored pairs with no FDI flows, or treated these cases as either *measurement errors* or *literally* indicating *zero flows*.[16] This view is consistent with models that ignore the role of setup cost. In such models the empty flow cells are either zeros, or a consequence of measurement errors, likely with a small volume of capital flow. Note that if measurement errors (in the $X_{i,j,t}$) are not correlated with the explanatory variables then the estimated parameters are not biased, although they are imprecisely estimated.

This does not hold in a model with setup costs of investment. If setup costs play an important role in determining whether a source country i invests directly in a host country j; then we should expect a negative correlation between the error term in the FDI flows equation and the error term in the profit equation.

Hence, by introducing the profit equation into our statistical model we do not ignore the investment decision. Moreover, by doing so we are able to estimate the correlation between the likelihood of source country i to invest directly in host country j and the FDI flows.

This setup generates important testable implications. While the "measurement errors" hypothesis implies that the cross equation correlation between the $i-j$ pairs FDI estimated residuals, obtained from the flow equation (18) and their residuals in the selection equation (26) is positive, the setup-cost model points to a negative correlation between the likelihood of (potential) source country i to invest in host country j and the volume of the FDI flows.

Hence, by estimating the full system we correct for the selection bias, and also recover the relevancy of setup costs in the generation of FDI exports.

5.1.3 The correlation between the error terms

Two methods are available to address the problem of missing data points: the Heckman selection method and Tobit.

Allowing for measurement errors and setup costs: The Heckman Model

To overcome the selection bias we employ Heckman's (1979) well known model, but instead of using Heckman's two stage procedure we obtain consistent and efficient estimates of the unknown parameters, $\beta,\ \gamma,\ \rho,\ \sigma_{U_F}$, (normalizing to $\sigma_{U_S} = 1$) using maximum likelihood (ML). A critical parameter is ρ, the correlation between the error terms. We estimate ρ from a cross equation correlation between the $i-j$ pair estimated residuals, obtained from the flow equation (18) ; and their residuals in the selection equation (26). Note that this correlation can be estimated using only the cross-sectional dimension of the data.

Allowing for measurement errors but ignoring setup costs: Tobit

Tobin (1958) observed that variables of interest are characterized by lower or upper limits, and takes on the limiting value for a substantial mass of observations. Tobin proposed an hybrid of a probit analysis

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

and a standard linear regression in estimating the model with a limited number of dependent variables, known since then as the Tobit model. The Tobit model is often used in the empirical international trade literature (e.g., Carr, Markusen, and Muskus 2001). The Tobit model is based on the following latent-variable model:

$$Y^*_{i,j,t} = X_{F,i,j,t}\beta + U_{F,i,j,t}, \tag{33}$$

whereas $X_{F,i,j,t}$ is a vector of regressors, and the error term $U_{F,i,j,t}$ is $N\left(0,\ \sigma^2_{U_F}\right)$ distributed, conditionally on $X_{F,i,j,t}$.

The latent variable $Y^*_{i,j,t}$ is only observed if $Y^*_{i,j,t} > 0$ (or any other lower level), thus the actual dependent variable $Y_{i,j,t}$ is:

$$Y_{i,j,t} = \max\left(0,\ Y^*_{i,j,t}\right) \tag{34}$$

Note that in this model, by definition $X_{F,i,j,t} = X_{S,i,j,t}$.

The population regression function for equation (34) is:

$$E\left(Y_{i,j,t} \mid X_{F,i,j,t},\ D_{i,j,t} = 1\right) = X_{F,i,j,t}\beta + \sigma_{U_F} \cdot \tilde{\lambda}_{i,j,t}. \tag{35}$$

where

$$\tilde{\lambda}_{i,j,t} = \frac{\phi\left(\frac{X_{F,i,j,t}}{\sigma_{U_F}}\right)}{\Phi\left(\frac{X_{F,i,j,t}}{\sigma_{U_F}}\right)}. \tag{36}$$

Comparing the set of equations (30) and (31) and the set of equations (35) and (36) the Tobit model can be viewed as a special case of the Heckman model, for the particular case where the flow equation serves also as the participation equation (up to a scale). Because the only difference between the participation and the FDI 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. In such a case, the participation equation error term and the flow equation error term are perfectly positively correlated $(\rho = 1)$, and the parameters in the selection equation have the same values as the parameters in the FDI flow equation.

What do we learn fromthe panel data?

There are three main advantages of using panel data, relative to cross-sectional data. First, having panel data we can decompose the cross-equation correlation between the estimated residuals of $i-j$ FDI pairs into the correlation between the $i-j$ specific time-invariant effect and the transitory $i-j$ FDI pair shocks. Second, the panel data can improve the efficiency of the estimates, if, for instance, the error term is serially correlated. Third, if the $i-j$ time-invariant pair specific terms are correlated with the regressors, then the panel data may reduce the asymptotic bias, which obtains with a single, or repeated, cross-section data.[17]

[17] Note that our model assumes the the lumpy costs of investment recur every period. Thus theory does not suggest that FDI flows are serially correlated.

Following the decomposition of the error terms into (i) $i-j$ pairs time invariant heterogeneity and (ii) transitory shocks, we have the following error structure:

$$U_{F,i,j,t} = \theta_{i,j} + \eta_{i,j,t,} \tag{37}$$

$$U_{S,i,j,t} = C_{i,j} + \nu_{i,j,t}. \tag{38}$$

We assume that:

$$\left(\theta_{i,j},\ C_{i,j},\ \eta_{i,j,t},\ \nu_{i,j,t}\right) \sim N\left(0, \Omega^*\right),$$

where:

$$\Omega^* = \begin{vmatrix} \rho\frac{\sigma_{U_F}}{\sigma_{U_S}} & \rho^L\sigma_{\theta C} & & \\ & \sigma_C^2 & & \\ & & \sigma_\eta^2 & \rho^S\sigma_{\eta\nu} \\ & & & \sigma_{U_S}^2 \end{vmatrix},$$

and $\rho^L = -\frac{\sigma_{\theta C}}{\sigma_\theta \sigma_C}$ is the correlation between the “permanent” terms and $\rho^S = \frac{\sigma_{\eta\nu}}{\sigma_\eta \sigma_\nu}$ stands for the correlation between the deviations. Hence, by controlling for fixed effects we can decompose the correlation between the error terms into long-run and short-run components.

Endogeneity of the explanatory variables

The large fraction of empty cells 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 have treated the explanatory variables as exogenous to FDI flows. Although bilateral FDI flows are only a subset of the international capital flows that enter 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 countries’ 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. Since 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.

Selection equation: instruments and functional form

The existence of fixed cost implies that the likelihood that a country i will export FDI to country j is non-linear in the explanatory variables because this function is a flat 0 below a threshold and a flat 1 above it, even if the flow equation is linear in these explanatory variables. Therefore, the difference in the functional forms stem from the fixed cost, and identification based on differences in functional form is quite natural. In addition, there are also a set of variables that are plausibly associated with fixed cost but not with marginal productivity of capital. We use two variables as instruments in the selection equation: (i) lagged FDI participation variable $\left(D_{i,j,t-1}\right)$ and (ii) source-country financial risk ratings. That is, setup costs of a capacity expansion of a pre-existing foreign subsidiary are lower than the corresponding setup costs for a greenfield investment in a new subsidiary; and (iii) source-country financial risk ratings. Similarly, the source-country capital market seems to be important for acquisitions of green-field foreign establishments, but less so for expansion investments in pre-existing foreign subsidiaries (which are typically heavily financed by host-country financial institutions).

6. Data

Our original data were drawn from OECD reports on a sample of 45 countries, both developing and developed 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 12 OECD source countries to 45 OECD and non- OECD countries. However, the OECD reports accurately on *all* rich and poor countries that are a host to OECD FDI exports. To overcome the missing reported data on non-OECD countries as sources of FDI exports, we restrict our sample to the group of OECD countries, as potential source and host countries among themselves. In the appendix we also present results for the 45-country sample.

We employ three-year averages, so that we have six periods (each consisting of three years). The main variables we employ are: (1) standard country characteristics such as GDP or GDP per-capita, population, educational attainment, geographical longitude and altitude, language, road length per country's area, telephone lines per-capita, etc.; (2) $s-h$ pair data such as $s-h$ FDI flows, geographical distance, common language (zero-one variable), $s-h$ flows of goods, bilateral telephone traffic, etc. The appendix provides more information on the data: Table A1 describes the list of the 45 countries in the sample and whether observed in the sample (at least once) as a source or host country; Table A2 describes the sources for our data.

7. A First Look at the Selection Bias Problem in the Data

As was already pointed out, the selection of countries into $s-h$ pairs is a key feature in the data. Out of 3,312 (=24 x 23 x 6, recalling that there are six periods) potential $s-h$ pairs, we observe only 1,291 (about 39%) $s-h$ pairs with positive flows in the data. Only 12 countries serve as a source for FDI exports. Expanding on this remarkable feature, we take a first look into the $s-h$ selection pattern in this section.

Table 1 describes the number of periods (out of 6) in which a country i exports FDI to a country j. The countries are arranged by an increasing order of GDP per capita. The table clearly indicates that only rich countries export FDI, both to other rich countries and to poor countries. But all 24 countries (including the very rich countries) serve as hosts to inward FDI flows. Table 2 describes the volume (as a percentage of the host country GDP) of FDI flows from the 12 source countries to all the 24 host countries. The Table suggests that the bulk of FDI export does not go to low GDP per capita countries, but rather to middle GDP per capita countries, such as Ireland, Belgium, and the Netherlands. Thus, GDP per capita in the host-source country pairs is presumably not a good predictor of the volume of FDI flows, conditional on having positive FDI flows. In contrast, Tobit estimators, which apply to all 24 countries (including the 12 non-source countries, which have low GDP per capita) treat them as potential source countries. Thus the Tobit estimators would suggest that source-country GDP per capita is a good predictor for FDI flows.

Figure 2 describes the fraction of OECD countries which serve as a host to FDI inflows from other OECD countries, ranked by GDP per capita. Observe that the figure exhibits a “discontinuity”: there is a certain income level below which a country does not export FDI at all; above such income level, almost of all of them export FDI at least to some countries.

This eyeballing amounts to *a prima facia* evidence for the existence of fixed costs.

8. Estimation

Table 1 demonstrates that FDI flows from rich to poor countries. We now turn to the estimation of the determinants of these flows. To estimate the effect of a change in 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 first benchmark, we ignore the selection equation (20), and simply estimate the gravity equation (18) ; by treating all FDI flows in $s-h$ pairs with no recorded FDI flows as “zeros”.[18] The rationale for inserting inzeroslgis 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, 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 one-dollar value (with the log equalling zero) as a common low value for the value of the FDI flows for the no-flows $s-h$ pairs. (All other positive flows have logarithmic values greatly 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 reported as zeros.) Tobit estimation is indeed often employed in the analysis of international flows of goods and capital. We report the results in Panel B in Table 3, with two censor levels (0.01 and 1.00).

Against these two benchmarks, the complete picture, and especially the role played by the unobserved fixed set up costs, is 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 setup costs. The results are reported in Panel C in Table 3, for the case where the source-country financial risk ratings and the lagged FDI participation variable $D_{i,j,t-1}$ are not excluded from the flow equation (identification is still possible due to the different functional forms of the flow and selection equations); and in Panel D in Table 3 for the case where both, source-country financial risk

[18] More precisely, the log of the FDI flow is set equal to zero. We performed robustness tests by replacing the zeros by large negative numbers. The conclusions are not meaningfully changed.

ratings and the lagged FDI participation variable $D_{i,j,t-1}$ are excluded from the flow equation and serve as instruments in the selection equation.

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 not significant in the Heckman selection equation.[19] Turning to the effect of the host country education level, relative to the source country counterpart, there is no effect in the OLS model. We do 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 positively affects the likelihood of a non-zero source-host pair, but does not influence the volume of FDI flows within the pair.

Source country financial risk ratings are 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.[20] In contrast to the OLS and Tobit models, where the effect 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 the 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 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.[21]

The coefficient of the lagged FDI participation variable (0.619) in panel D is expressed in terms of standard deviations of the unobserved profits (20). Thus, a pair of countries which already had positive FDI flows between them in period $t-1$; have the equivalent saving in setup cost of investment in period t; of a 0.619 standard deviation of profits. Most importantly as a "smoking gun" for the existence of fixed costs in the data, we note that:

1. We find no difference between the results in panels C and D. That is, the results do not depend on the exclusion restrictions, pointing to the role of setup cost which generate the difference in the functional forms between the flow and the selection equation.

---

[19] Remember that in the estimation we control for country fixed effects. In the Appendix we also present results of the estimation without controlling for country fixed effects.

[20] Note, from Tables A2 in the Appendix, that without controlling 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.

[21] Note from Tables A2 in the Appendix, that without country fixed effects, the coefficient is significant.

2. 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 additional evidence for the relevance of fixed set up costs.

9. 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 hypothesis of setup costs of investment. If technology shocks drive jointly marginal productivity of capital and setup costs of capital formation, then shocks to setup costs must be negatively correlated with shocks to marginal productivity. That is, below-average setup costs for a technologically advanced source country is associated with above-average marginal productivity of capital in the country. If the setup costs which govern the formation of a certain $s - h$ pair of countries, is high, it becomes unlikely to observe such a pair. The error term in the participation equation is negative in this case. But, then the error term in the gravity equation is positive because the marginal productivity in the source country is low. (Recall that the unobserved heterogeneity in the gravity equation is affected only by the marginal productivity of capital.) However, the unobserved heterogeneity in the selection equation is affected both by the marginal productivity of capital *and* by the setup costs of investment. The negative correlation implies that a source country with low setup costs (and, therefore, with positive error in the selection equation) is also the source country with high marginal productivity of capital (and, therefore, with negative errors in the gravity equation).

If education, as measured by the average years of schooling is indeed a “good” measure for host-source country differences in human capital, then education levels are important in predicting the volume of FDI flows. The Heckman model predicts that, as a country improves the education level, it would raise the likelihood of becoming a host to FDI flows. This finding is in line with the Lucas hypothesis. Likewise, improvements in the host country financial risk ratings are 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 becoming a source for FDI exports is therefore lessened.

10. More Information From the Full Sample

So far, our results are derived from the OECD country sample. The data source, *International Direct Investment Database*, OECD, is reporting data on OECD source countries and OECD and non-OECD host countries. The fact that 29 non-OECD countries serve as a host to OECD country exports of FDI provides additional information that could be exploited. Results (reported in the Appendix) are broadly in line with Table 3 for the OECD sub-sample.

11. Conclusion

The existence of setup costs of foreign direct investment presents the investors with a two-fold decision: whether to invest at all and how much to invest. Invoking this idea 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).

The FDI model works as follows. A comparative advantage for the rich country concerns both relatively low startup costs of direct investment, and relatively high marginal productivity of capital. As a result, the ere is a wage gap between the rich and poor countries. The lower wage rate in the (poor) host country becomes a pull factor for positive FDI inflows into the country. But the total factor productivity in the source (technologically advanced) country which is higher than in the host country, tend to reduce the volume of FDI flows. Assuming that the source-host wage differences dominates the total factor productivity differences, the poor-host country does attract FDI from the rich-source country, but among source countries, the more technologically advanced one will export FDI in a relatively small volume. Consequently, the model predicts that an above-average technologically-advanced country is more frequently a source for FDI exports in the sample. But, at the same time, among countries which serve as a source for FDI exports, the volumes of FDI flows from an above-average technologically-advanced country are relatively small.

Empirical international trade and international finance literature often failed 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 literally indicating zero flows. A standard procedure 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. 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 nonetheless reported as zeros. 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 selection-equation error term and the flow-equation error term are restricted to be perfectly positively correlated.

To allow for the role played by unobserved fixed setup costs, we employ the Heckman selection method. We estimate jointly the maximum likelihood of the volume of FDI flows (the so-called gravity equation), and the selection of countries into source-host country pairs (the selection equation). Evidently, this estimation procedure accommodates both measurement errors and, crucially, the possible existence of setup costs in the data. 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 selection equation.

We do indeed find that the correlation between the error terms is negative in our data set, and various data subsets, indicating the importance of setup costs that govern the export of FDI in the data. We find that important predictors of a selection of a pair of countries as a host-source pair are: source country GDP per capita, difference in education levels, and differences in financial risk ratings. These variables may 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 and is expanding. Generally, these findings support the existence of setup costs of foreign direct investment. Furthermore, the evidence points to differing effects on FDI flows of the marginal productivity conditions and the setup cost conditions.

To sum up, the paper brings forth new evidence that the source country GDP percapita, and average years of schooling, are significant determinants for the selection of countries into source-host pairs for FDI flows. 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 rich to poor countries, and from countries with high average years of schooling to countries with low average years of schooling, but in a somewhat more subtle way than may be inferred from the marginal productivity conditions alone.

References

Amiti, Mary (1998), "Inter-Industry Trade in Manufactures: Does Country Size Matter?" *Journal of International Economics*, 44(2): 231-55.

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

Caballero, Ricardo and Eduardo Engel (1999), "Explaining Investment Dynamics in US Manufacturing: A Generalized (S;s) Approach," *Econometrica*, 67: 783-826.

________________ (2000), "Lumpy Adjustment and Aggregate Investment Equations: A 'Simple' Approach Relying on Cash Flow 'Information'," mimeo.

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

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

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*, 8: 478-510.

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

Heckman, James J. (1979), "Sample Selection Bias as a Specification Error," *Econometrica* 42: 153-68.

Helpman, Elhanan (2003), *The Mystery of Economic Growth*, manuscript, August.

Helpman, Elhanan, Mark Melitz, and Steve Yeaple (2003), "Exports vs FDI," NBER Working Paper No. 9439, Cambridge MA: National Bureau of Economic Research.

Kyriazidou, Ekaterini (1996), "Estimation of a Panel Data Sample Selection Model," *Econometrica* 65(6): 1335-64.

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): 526-43.

Lucas, Robert E. (1990), "Why Doesn't Capital Flow from Rich to Poor Countries?" *American Economic Review: Papers and Proceedings*, 80(2): 92-6.

Melitz, Marc J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6): 1695-725.

Mody, Ashoka, Assaf Razin and Efraim Sadka (2003), "The Role of Information in Driving FDI Flows: Host-Country Tranparency and Source Country Specialization," NBER Working Paper No. 9662, Cambridge MA: National Bureau of Economic Research.

Obstfeld, Maurice and Alan M. Taylor (2003), "Globalization and Capital Markets," in Bordo, Michael D., Alen M. Taylor and Jeffrey G. Williamson, eds., *Globalisation in Historical Perspective*, University of Chicago Press.

Roberts, Mark J. and James R. Tybout (1991), "Size Rationality and Trade Exposure in Developing Countries," The World Bank, mimeo.

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

Smanzyska, Beata K. and Shang-jin Wei (2001), "Corruption and Cross-Border Investment," NBER Working Paper No. 8465, Cambridge MA: National Bureau of Economic Research.

Stolper, Wolfgang and Paul A. Samuelson (1941), "Protection and Real Wages," *The Review of Economic Studies*, 9: 56-73.

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

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

Zhang, Kevin Honglin and James R. Markusen (1999), "Vertical Multinationals and Host-Country Characteristics." *Journal of Development Economics*, 59(2): 233-52.

12. Appendix

12.1 List of Countries and Data Sources

Table A1.1: List of Countries by Source-Host Status

Country	Source	Host	Country	Source	Host
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*		+			

Notes: * OECD countries

Table A1.2: 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 Traffics,
	International Communication Union
	International Telecommunications Union
Education Attainment	Barro-Lee Dataset: www.nber.org/N...
	....
Language	....
	....
ICRG index of financially sound rating (inverse of financial risk)	Ashoka Mody, IMF

12.2 Bilateral FDI Flows for the OECD Countries Excluding Country Fixed Effects

Table A.2: Bilateral FDI Flows and Selection into Source-Host Pairs: OLS, Tobit and Heckman Maximum Likelihood Without Country Fixed Effects OECD Countries only

	Panel A: OLS	Panel B: Tobit Correction		Panel C: Heckman selection		Panel D: Heckman selection	
		Low censored		Equation:			
		0.01	1	FDI flows	selection	FDI flows	selection
Variables							
GDP per capita - host^	0.026	-0.183	-0.091	0.357	-0.111	0.400	0.068
	(0.156)	(0.260)	(0.229)	(0.216)	(0.140)	(0.214)	(0.103)
GDP per capita - source^	1.589	3.910	3.251	0.261	1.432	0.487	1.419
	(0.155)**	(0.323)**	(0.282)**	(0.442)	(0.146)**	(0.480)	(0.150)**
Difference between source and	-0.033	0.013	-0.009	-0.062	-0.006	-0.095	0.003
host years of schooling	(0.028)	(0.046)	(0.041)	(0.035)	(0.028)	(0.036)**	(0.027)
Financial risk rating - host	0.068	0.128	0.131	0.104	0.027	0.104	
	(0.017)**	(0.029)**	(0.026)**	(0.020)**	(0.015)	(0.019)**	
Financial risk rating - source	0.047	0.244	0.210	0.126	0.050	0.055	
	(0.014)**	(0.042)**	(0.037)**	(0.033)**	(0.017)**	(0.019)**	
Common language	0.590	1.075	1.099	1.242	0.128	1.186	0.152
	(0.224)**	(0.230)**	(0.200)**	(0.220)**	(0.150)	(0.214)**	(0.152)
Distance (in logs)	-0.349	-0.779	-0.744	-0.541	-0.111	-0.505	-0.101
	(0.082)**	(0.085)**	(0.075)**	(0.090)**	(0.053)*	(0.098)**	(0.052)
Population - host^	0.327	0.642	0.621	0.643	0.096	0.611	0.105
	(0.078)**	(0.086)**	(0.076)**	(0.079)**	(0.055)	(0.086)**	(0.053)
Population - source^	1.350	2.344	2.111	0.844	0.790	0.749	0.773
	(0.059)**	(0.092)**	(0.082)**	(0.136)**	(0.065)**	(0.206)**	(0.069)**
Correlation ($U_{i,j}$, $V_{i,j}$)				0.055		-0.275	
				(0.214)		(0.323)	
SD of $U_{i,j}$ (flow equation)				1.462		1.504	
				(0.055)		(0.083)	
Inverse Mills ratio				0.081		-0.414	
				(0.313)		(0.503)	
Observations	2116	2116	2116	2116	2116	2116	2116

Note:
^ in logs
All specifications include year and country fixed-effects.
Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

12.3 The OECD and the Non-OECD country sample

Table A.3.1: Bilateral FDI Flows and Selection into Source-Host Pairs: OLS, Tobit and Heckman Maximum Likelihood Controlling for Country Fixed Effects OECD and Non-OECD Countries

	Panel A: OLS	Panel B: Tobit Correction		Panel C: Heckman selection		Panel D: Heckman selection	
		Low censored		Equation:			
		0.01	1	FDI flows	selection	FDI flows	selection
Variables							
GDP per capita - host^	0.104	-0.593	-0.611	0.297	-0.206	0.346	-0.292
	(0.077)	(0.543)	(0.517)	(0.381)	(0.394)	(0.386)	(0.369)
GDP per capita - source^	0.087	0.668	0.367	-0.141	0.715	-0.264	-0.112
	(0.034)*	(1.036)	(0.911)	(0.466)	(1.249)	(0.483)	(1.171)
Financial risk rating - host	-0.001	0.013	0.022	0.034	-0.009	0.036	-0.017
	(0.004)	(0.017)	(0.015)	(0.010)**	(0.014)	(0.010)**	(0.014)
Financial risk rating - source	-0.028	-0.207	-0.189	-0.076	-0.097		-0.127
	(0.003)**	(0.042)**	(0.037)**	(0.027)**	(0.034)**		(0.036)**
Difference between source and	0.081	0.316	0.262	0.060	0.222	0.038	0.216
host years of schooling	(0.026)**	(0.111)**	(0.101)**	(0.083)	(0.081)**	(0.084)	(0.079)**
Common language	0.187	0.962	0.925	0.721	0.352	0.715	0.278
	(0.070)**	(0.115)**	(0.102)**	(0.101)**	(0.124)**	(0.101)**	(0.107)**
Distance (in logs)	-0.321	-1.059	-1.007	-0.762	-0.503	-0.754	-0.399
	(0.039)**	(0.069)**	(0.061)**	(0.071)**	(0.083)**	(0.068)**	(0.075)**
Population - host^	1.289	9.334	7.529	-0.317	7.904	-0.488	6.606
	(0.356)**	(1.687)**	(1.529)**	(1.438)	(1.505)**	(1.285)	(1.470)**
Population - source^	-0.929	-7.247	-5.797	2.899	-7.142	1.608	-5.665
	(0.197)**	(3.506)*	(3.127)	(2.442)	(2.655)**	(2.345)	(2.613)*
Lagged FDI participation variable							0.809
							(0.098)**
Correlation between the error terms				-0.520		-0.632	
$\rho(U_F, U_S)$				-(0.213)		(0.100)	
$\sigma(U_F)$				1.071		1.091	
				-(0.051)		(0.0381)	
Inverse Mills ratio				-0.557		-0.689	
				-(0.252)		(0.128)	
Observations	6724	6724	6724	6724	6724	6724	6724

Note:
^ in logs
All specifications include year and country fixed-effects.
Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table A.3.2: Bilateral FDI Flows and Selection into Source-Host Pairs: OLS, Tobit and Heckman Maximum Likelihood Without Country Fixed Effects OECD and Non-OECD Countries

	Panel A: OLS	Panel B: Tobit Correction		Panel C: Heckman selection		Panel D: Heckman selection	
		Low censored		Equation:			
		0.01	1	FDI flows	selection	FDI flows	selection
Variables							
GDP per capita - host^	0.273	0.639	0.655	0.764	0.107	0.729	0.264
	(0.047)**	(0.111)**	(0.099)**	(0.088)**	(0.059)	(0.103)**	(0.048)**
GDP per capita - source^	0.944	4.264	3.588	0.319	1.402	0.401	1.361
	(0.048)**	(0.223)**	(0.193)**	(0.353)	(0.100)**	(0.391)	(0.102)**
Difference between source and	-0.063	-0.078	-0.087	-0.093	-0.022	-0.121	-0.012
host years of schooling	(0.014)**	(0.034)*	(0.030)**	(0.027)**	(0.019)	(0.028)**	(0.019)
Financial risk rating - host	0.024	0.104	0.099	0.056	0.031	0.055	
	(0.005)**	(0.014)**	(0.013)**	(0.010)**	(0.008)**	(0.010)**	
Financial risk rating - source	0.012	0.226	0.190	0.136	0.053		0.059
	(0.004)**	(0.028)**	(0.025)**	(0.027)**	(0.009)**		(0.011)**
Common language	0.333	1.060	1.108	1.067	0.159	0.981	0.158
	(0.084)**	(0.165)**	(0.146)**	(0.161)**	(0.095)	(0.155)**	(0.093)
Distance (in logs)	-0.243	-0.732	-0.688	-0.490	-0.123	-0.432	-0.126
	(0.052)**	(0.070)**	(0.062)**	(0.077)**	(0.043)**	(0.079)**	(0.043)**
Population - host^	0.207	0.766	0.728	0.635	0.175	0.567	0.206
	(0.036)**	(0.062)**	(0.055)**	(0.065)**	(0.037)**	(0.073)**	(0.035)**
Population - source^	0.809	2.321	2.094	0.863	0.727	0.713	0.704
	(0.036)**	(0.067)**	(0.060)**	(0.130)**	(0.044)**	(0.142)**	(0.046)**
Correlation ($U_{i,j}$, $V_{i,j}$)				0.059		-0.379	
				(0.229)		(0.205)	
SD of $U_{i,j}$ (flow equation)				1.443		1.517	
				(0.042)		(0.079)	
Inverse Mills ratio				0.086		-0.575	
				(0.331)		(0.337)	
Observations	6724	6724	6724	6724	6724	6724	6724

Note:
^ in logs
All specifications include year and country fixed-effects.
Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table 1. Number of Source-Host country Pairs by GDP per capita

Host / Source	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	0	0	0	0	0	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	0	0	0	0	0	0	0	0	0	0	0	0	0	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	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	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
UK	1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1
Canada	0	1	1	1	1	1	1	1	1	1		1	0	1	1	1	1	1	1	1	1	1	1	1
Australia	1	0	1	0	0	0	1	1	1	1	1		0	1	1	1	0	1	1	0	0	0	1	1
Finland	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0
France	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1
Germany	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1
Netherlands	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1
Sweden	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1
Belgium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0
US	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1
Austria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1
Norway	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1
Denmark	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0
Japan	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1
Switzerland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Average	0.48	0.48	0.52	0.48	0.48	0.48	0.48	0.52	0.48	0.48	0.48	0.48	0.43	0.48	0.48	0.48	0.43	0.52	0.48	0.43	0.43	0.48	0.48	0.52

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

Host / Source	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
Italy	0.7	0.3	0.1	3.6	1.5	2.5	0.0	5.7		2.7	0.5	0.2	0.3	2.2	0.4	12.2	0.8	20.1	0.4	1.1	0.2	0.3	0.1	7.8
UK	4.5	3.5	0.7	12.0	8.0	8.8	32.3	52.1	3.5		9.6	27.1	1.0	6.9	2.4	62.7	8.7	15.8	10.7	2.1	15.6	3.6	0.4	17.3
Canada	0.0	1.7	0.1	0.4	0.3	0.4	7.8	32.1	0.2	3.8		2.2	0.0	0.7	0.2	1.6	1.3	3.1	4.0	0.6	0.4	0.1	0.1	1.0
Australia	0.0	0.0	0.1	0.0	0.0	0.0	43.7	4.4	0.2	5.8	1.0		0.0	0.0	0.0	1.2	0.0	0.2	1.2	0.0	0.0	0.0	0.0	0.0
Finland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
France	3.3	1.2	1.0	8.4	2.7	12.1	1.4	7.9	6.6	11.0	4.4	3.6	0.5		3.4	27.2	6.7	44.5	3.8	2.1	2.4	1.8	0.1	16.6
Germany	4.7	3.4	1.8	9.3	4.0	9.0	0.7	69.0	6.2	16.6	4.7	2.9	2.1	8.0		19.9	6.1	39.6	4.7	22.7	4.3	4.7	0.4	18.3
Netherlands	1.0	1.5	0.5	5.8	3.8	5.5	0.0	35.1	1.2	13.1	1.3	2.2	0.5	3.3	1.3		6.5	40.0	3.3	1.3	2.8	5.6	0.1	10.0
Sweden	0.2	0.5	0.3	0.8	0.1	0.8	0.1	21.1	0.5	4.3	0.3	0.4	35.4	1.6	0.6	9.9		2.7	1.0	0.6	15.4	6.0	0.0	3.3
Belgium	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0
US	3.4	36.2	4.8	6.8	1.2	6.5	26.1	127.3	6.4	57.0	47.0	27.4	4.1	8.0	4.3	60.3	5.7	35.7		4.2	16.8	3.9	1.3	39.9
Austria	0.2	0.0	0.0	0.5	0.1	0.2	0.0	2.1	0.3	0.8	0.2	0.3	0.0	0.1	0.4	0.4	0.2	0.9	0.1		0.0	0.7	0.0	1.3
Norway	0.0	0.0	0.0	1.1	0.0	0.4	0.2	4.1	0.1	1.6	0.9	0.1	1.8	0.3	0.1	1.6	8.4	0.9	0.2	0.7		7.1	0.0	0.2
Denmark	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0
Japan	1.8	4.1	7.7	1.2	0.5	2.7	16.7	19.1	0.8	19.1	7.7	34.2	0.6	2.6	1.3	28.3	0.3	18.2	15.7	1.0	3.3	0.2		4.5
Switzerland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
Average	0.9	2.3	0.7	2.2	1.0	2.1	5.6	16.5	1.1	5.9	3.4	4.4	2.0	1.5	0.6	9.8	1.9	9.6	2.0	1.6	2.7	1.5	0.1	5.2

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

Variables	Panel A: OLS	Panel B: Tobit Correction		Panel C: Heckman selection		Panel D: Heckman selection	
		Low censored		Equation:			
		0.01	1	FDI flows	selection	FDI flows	selection
GDP per capita - host^	0.283	0.454	0.351	0.207	0.237	0.158	0.082
	(0.459)	(1.093)	(0.971)	(0.684)	(0.830)	(0.697)	(0.853)
GDP per capita - source^	0.324	-0.192	-0.186	-0.123	-1.178	-0.232	-1.882
	(0.316)	(1.347)	(1.181)	(0.609)	(1.412)	(0.624)	(1.587)
Difference between source and	0.086	0.338	0.263	-0.052	0.276	-0.073	0.261
host years of schooling	(0.060)	(0.136)*	(0.122)*	(0.099)	(0.112)*	(0.100)	(0.106)*
Financial risk rating - host	0.002	0.001	0.015	0.053	-0.014	0.058	-0.018
	(0.012)	(0.029)	(0.026)	(0.017)**	(0.025)	(0.016)**	(0.026)
Financial risk rating - source	-0.040	-0.201	-0.171	-0.057	-0.134		-0.153
	(0.011)**	(0.054)**	(0.047)**	(0.032)	(0.046)**		(0.048)**
Common language	0.529	1.222	1.163	0.892	0.491	0.881	0.410
	(0.135)**	(0.152)**	(0.134)**	(0.121)**	(0.211)*	(0.123)**	(0.187)*
Distance (in logs)	-0.448	-0.926	-0.883	-0.667	-0.482	-0.663	-0.408
	(0.077)**	(0.090)**	(0.079)**	(0.084)**	(0.107)**	(0.083)**	(0.098)**
Population - host^	1.949	6.878	4.119	-4.952	10.438	-5.331	8.949
	(1.608)	(3.715)	(3.286)	(2.489)*	(2.767)**	(2.366)*	(2.792)**
Population - source^	0.202	-2.695	-2.629	-0.136	-0.016	-1.170	0.514
	(1.490)	(4.468)	(3.959)	(2.864)	(3.278)	(2.817)	(3.400)
Lagged FDI participation variable							0.619
							(0.145)**
Correlation between the error terms				-0.577		-0.605	
$\rho(U_F, U_S)$				(0.177)		(0.157)	
$\sigma(U_F)$				1.068		1.070	
				(0.051)		(0.048)	
Inverse Mills ratio				-0.616		-0.648	
				-(0.213)		-(0.191)	
Observations	2116	2116	2116	2116	2116	2116	2116

Note:
^ in logs
All specifications include year and country fixed-effects.
Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Figure 1. Selection Bias and Setup costs Presence of Setup Costs

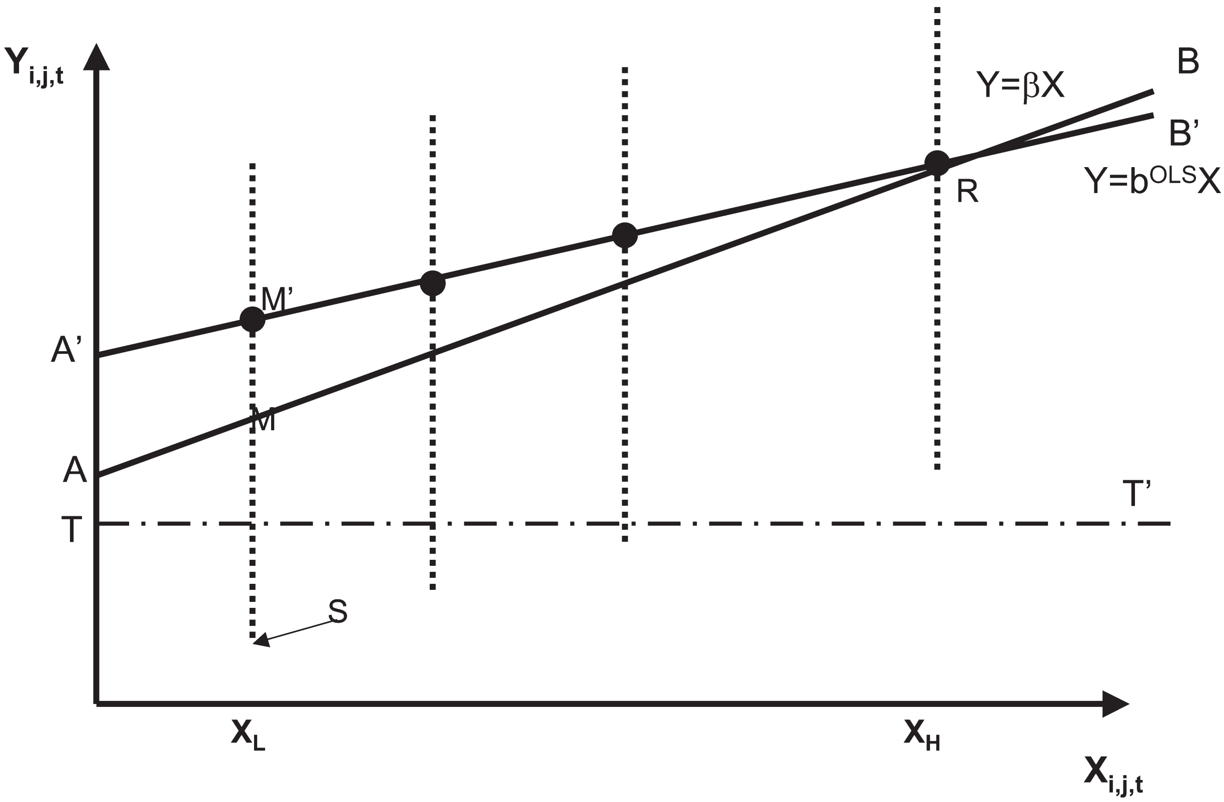

Figure 2. Fraction of OECD Countries which Serve as a Host to FDI Inflows from other OECD Countries, Ranked by GDP per Capita.