Barriers to Global Capital Allocation

Bruno Pellegrino[*] University of Maryland

Enrico Spolaore[†] Tufts University and NBER

Romain Wacziarg[‡] UCLA and NBER

PDF: click here for the most up-to-date draft.

December 2020

Abstract

We quantify the impact of barriers to international investment, using a novel multi-country overlapping generations model with heterogeneous investors and imperfect capital mobility. Our model yields a gravity equation for foreign asset demand. We estimate this gravity equation using recently-developed foreign investment data that has been restated to account for the presence of offshore investment and financing vehicles. We show that a parsimonious implementation of the model, with four barriers (capital controls, geographic distance, institutional distance and cultural distance) can account for a large share of the observed variation in bilateral Foreign Direct Investment (FDI) and Foreign Portfolio Investment (FPI) positions. Our model predicts a significant home bias and higher rates of return on capital in emerging markets. In our benchmark calibration, we estimate that capital misallocation induced by these barriers reduces World GDP by 8.8%, compared to a situation without barriers. We also find that barriers to global capital allocation contribute significantly to cross-country inequality: the standard deviation of log capital per employee is 70% higher than it would be in a world without barriers to international investment, while the dispersion in output per employee is 39% higher.

JEL Codes: E22, E44, F2, F3, F4, G15, O4

Keywords: Capital Flows, Foreign Investment, Culture, Institutions, Geography, Gravity, International Macroeconomics, International Finance, Open Economy

[*] `bpellegr@umd.edu` – 880 P Street, NW, Washington DC 20001; [†] `enrico.spolaore@tufts.edu` – Department of Economics, Tufts University, Medford, MA, 02155; [‡] `wacziarg@ucla.edu` – UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles CA 90095. We thank Pablo Roa Prieto for outstanding research assistance.

1 Introduction

Even though international capital flows have greatly increased in recent decades (Lane and Milesi-Ferretti, 2018), observed features of international investment data are consistent with significant barriers to capital investment across countries. Such features include the lack of large flows from capital-abundant to capital-scarce economies (Lucas, 1990; Alfaro, Kalemli-Ozcan, and Volosovych, 2008) and a disproportionate portfolio allocation towards domestic assets, commonly-referred to as the "home bias" of international financial markets (French and Poterba, 1991; Coeurdacier and Rey, 2013). In this paper, we investigate what factors prevent capital from freely flowing from one country to another, and provide a quantitative assessment of how much these barriers affect the efficient allocation of capital across countries, as well as cross-country inequality.

To address these issues, we propose a multi-country general equilibrium model of international investment, where heterogeneous investors choose to allocate capital among different destinations with varying objective and subjective returns. Intermediation costs vary across projects and countries, and depend on a vector of physical, cultural, and institutional distances between societies as well as policies that affect capital account openness. Such distances capture policy-induced barriers, costs in communicating and traveling to the destination country for the purpose of monitoring investments, as well as cultural factors such as differences in norms, values and beliefs.

The model yields a gravity equation for foreign assets demand. Equilibrium allocations of capital from each origin country to each destination country are characterized as a function of intermediation costs and other fundamentals. The effect of these variables on the equilibrium allocation can be directly estimated empirically. We do so using recently-developed Foreign Direct Investment (FDI) and Foreign Portfolio Investment (FPI) data that has been restated by (respectively) Damgaard, Elkjaer, and Johannesen (2019) and Coppola, Maggiori, Neiman, and Schreger (2020) from a residency to a nationality basis. This novel data accounts for the presence of offshore investment and financing vehicles located in tax havens such as Bermuda and the Cayman Islands. Barriers to capital flows are captured using measures of capital controls, geographic distance, cultural distance, and institutional distance.

We obtain three sets of empirical results. First, we find evidence that a parsimonious specification of our model, with four types of barriers (geographic distance, cultural distance, institutional distance, and capital controls) exert quantitatively substantial effects on international financial positions, controlling for origin and destination fixed effects. The effects are similar for FDI and FPI, are robust to using different specifications, and remain quantitatively large independently of the estimation method (OLS regressions, Poisson regression, and Instrumental Variable regressions). Our IV approach, used to ensure that our estimates are not amplified by reverse-causality, is based on the assumption that long-term factors, such as measures of ancestral and religious distance between populations, have an impact on contemporary measures of cultural and institutional distance. These in turn act as current barriers to the global allocation of capital.

Second, we find that a conservative calibration of our model predicts, out-of-sample, allocations of domestic capital that are consistent with the home bias in international investment documented in the literature (French and Poterba, 1991; Coeurdacier and Rey, 2013). Our estimates also match independently-measured differences in rates of return across countries. In particular, our model predicts higher rates of return on capital in emerging markets.

Third, we conduct a counterfactual analysis, using the model to study the quantitative implications of removing barriers to global capital allocation. We find that our four barriers introduce a significant misallocation of capital across countries. Compared to a situation without barriers, World GDP is 8.8% lower. Barriers also generate substantial increases in the dispersion of output per employee (38.8%) and of capital per employee (70.7%) compared to the zero-gravity benchmark.

This paper builds on an extensive theoretical and empirical literature on international capital flows. A seminal theoretical contribution in this area is Martin and Rey (2004), who provided a two-country, two-period model of international investment capturing a number of features of empirical gravity relations. Our model differs from Martin and Rey's in several dimensions. We extend the analytical framework to a multi-country, multi-period setting, and integrate assets markets with the real economy. In our framework, as in Martin and Rey's, investors have an incentive to diversify, but while Martin and Rey derive such diversification motive from risk aversion, we model it in terms of heterogeneity in intermediation costs, which in turn depend on a vector of geographical, cultural and institutional distances between societies. The theoretical section of this paper also shares some modeling choices with Head and Ries (2008)'s theory of cross-border mergers and acquisitions (M&A), although there are several differences here also: our theory is built to describe cross-border investment in a much broader sense than just international M&A, and it is embedded in a multi-country general equilibrium model - with consumption, saving, and production - that we use to perform quantitative counterfactuals. Moreover, unlike Head and Ries (2008), our model produces a fully-fledged gravity equation, where bilateral investment flows are proportional to the GDP of investor and destination countries.

Empirically, our paper contributes to a large literature that estimates the determinants of international flows using a gravity approach, originally suggested by Tinbergen (1962) for trade. Ghosh and Wolf (1999), De Ménil (1999) and Di Giovanni (2005) were among the first to use gravity regressions to study international assets flows. Applications of gravity models to financial flows have benefited from advances in the trade gravity literature (Eaton and Kortum, 2002; Anderson and Van Wincoop, 2003; Santos Silva and Tenreyro, 2006; Helpman et al., 2008). Based on the micro-founded gravity model of assets trade developed by Martin and Rey (2004), Portes and Rey (2005) showed systematic geographical patterns in gross cross-border equity portfolio flows. Subsequent empirical analyses of cross-border investment flows considered the role of historical and cultural factors (Eichengreen and Luengnaruemitchai, 2008; Guiso et al., 2009; Lane and Milesi-Ferretti, 2008; Rose and Spiegel, 2009; Blonigen and Piger, 2014).

Leblang (2010) found that diaspora networks affect international investment, and argued that cultural ties increase trust and reduce informational frictions. More recently, Burchardi et al. (2019) documented a causal effect of the ancestry composition of US counties on foreign direct investment sent and received by local US firms to and from the immigrants' nations of origin, and interpreted this effect as resulting from lower information frictions. Our paper also relates to the literature on historical and cultural barriers to international exchanges and the spread of innovations and development across countries (Spolaore and Wacziarg, 2009; Guiso et al., 2009; Felbermayr and Toubal, 2010; Spolaore and Wacziarg, 2012; Fensore et al., 2017; Bove and Gokmen, 2018; Spolaore and Wacziarg, 2018)).

While our paper is also concerned with studying the determinants of international flows, we attack the question from a rather different angle compared to previous empirical contributions: our study connects theory and empirics in a unified structural framework; moreover, we address a new question in the literature, by quantifying the extent of capital misallocation induced by international investment barriers. Thus, our work also contributes to the literature on that studies the misallocation of capital and other factors of production (Hopenhayn, 2014; Gopinath, Kalemli-Özcan, Karabarbounis, and Villegas-Sanchez, 2017; David and Venkateswaran, 2019; Baqaee and Farhi, 2020).

Our paper is also related to the recent trade literature on multinational enterprises (MNEs) (Arkolakis et al., 2018; Ramondo and Rodríguez-Clare, 2013; Tintelnot, 2016). This literature uses gravity models to study a specific aspect of foreign *direct* investment – multi-national production – and how this can act as a substitute for imports and exports. While we do share the structural modeling approach that is common within this literature, our research question is rather different, as we aim to understand foreign investment in a much broader sense; moreover (unlike this literature) we relate capital misallocation to geopolitical factors as well as cross-country variation in the rate of return on capital.

Finally, because we find that barriers to international investment amplify cross-country dispersion of capital and output per worker, our study also provides new evidence on the fundamental question of the origin of cross-country income differences (Hall and Jones, 1999; McGrattan and Schmitz Jr, 1999 - among many others).

The rest of the paper is organized as follows. In Section 2, we present our theoretical gravity model. In Section 3, we describe the specification and data used in the empirical analysis of the determinants of foreign investment. In Section 4, we present econometric estimates of the effect of geographical, cultural, institutional distance on foreign direct investment and portfolio investment. In Section 5, we use our framework to perform counterfactual analysis. Section 6 concludes.

2 A Multi-country Gravity Model of Foreign Investment

2.1 Production

In this section, we present a multi-country, general equilibrium overlapping generations (OLG) model with heterogenous investors and imperfect capital mobility that describes the allocation of capital investment across countries.

Time is discrete and indexed by t. There is a set of n countries $i \in \{1, 2, , \ldots, n\}$. Each country has fixed labor supply ℓ_i and a representative firm (also called i) that acts competitively and produces a perfectly-substitutable and tradable good using a Cobb-Douglas production function:

$$y_{it} \;=\; \omega_{it}\, k_{it}^{\theta_i}\, \ell_{it}^{1-\theta_i} \tag{2.1}$$

where y_{it} is the level of output and k_{it} and ℓ_{it} denote capital and labor input. ω_i is country i's total factor productivity. The production function parameter θ_i, which in equilibrium is equal to the capital income share, is allowed to vary across countries.

The final good can either be used for consumption or saved and transformed into units of capital to be used for production in the next period. Hence, the global resource constraint is:

$$\sum_{i=1}^{n} y_{it} \;=\; \sum_{i=1}^{n} (c_{it} + k_{it+1}) \tag{2.2}$$

where c_{it} is the current-period consumption of country i's agents. The final homogeneous good is assumed to be the numéraire of the economy (we normalize its price to one).

Investors from any country can purchase claims to country i's capital stock; by doing so, they are entitled to a proportional share of the capital income from the next period's production. Going forward, we will use the index i to refer to the country where production takes place (the *destination* country), and the index j to refer to the country that provides the capital (the *investor* country).

Assuming that the representative firm acts as a price-taker in input markets as well, the equilibrium rate of return on capital and wage rate are determined as usual:

$$r_{it} \;=\; \theta_i \,\frac{y_{it}}{k_{it}} \qquad w_{it} \;=\; (1-\theta_i)\,\frac{y_{it}}{\ell_{it}} \tag{2.3}$$

We make an additional assumption about the structure of production, which will be helpful later on when we model consumers' investment choices: production is carried out in *plants*. We identify individual plants with the index x. We assume that plants can be built and decommissioned costlessly, but each plant can contain a maximum capital stock of λ. In short, plants are simply a discretization of country i's capital stock. This implies that the number of active plants in each country is k_i/λ, and that there are total K/λ plants distributed over the n countries, where:

$$K \stackrel{\text{def}}{=} \sum_i k_i \tag{2.4}$$

2.2 Consumption and Saving

In each country j, a continuum of agents $z \in [0,1]$ is born every period t. They live for two periods and are endowed with ℓ_i units of labor in period t. Saving is their only source of income in period $t+1$.

The preferences of agent z, who is born in country j at time t, are described by the following intertemporal utility function:

$$U(z) \;=\; (1-\alpha)\log c_t(z) + \alpha \log c_{t+1}(z) \tag{2.5}$$

where $c_t(z)$ is agent z's consumption at time t. The amount of final good saved by investor z at time t is $s_t(z)$. Thus, agent z's intertemporal budget constraint is defined by the following two equations:

$$w_{jt}\ell_j \;=\; c_t(z) + s_t(z) \tag{2.6}$$

$$c_{t+1}(z) \;=\; R_{t+1}(z) \cdot s_t(z) \tag{2.7}$$

where $R_{t+1}(z)$ is the subjective return earned by investor (z) on their worldwide investment. Then, the Euler equation for agent z is:

$$\frac{\alpha}{c_{t+1}(z)} \cdot R_{t+1}(z) \;=\; \frac{1-\alpha}{c_t(z)} \tag{2.8}$$

We look for a steady-state equilibrium, with constant consumption, output, capital and saving $(\mathbf{c}_t, \mathbf{y}_t, \mathbf{k}_t, \mathbf{s}_t)$. Therefore, going forward, we drop time subscripts when referring to steady-state solutions. By plugging (2.6) and (2.7) inside (2.8), we have that, in equilibrium, all investors save a constant share ϕ of their labor earnings:

$$s_t(z) \;=\; s_{jt} \;=\; \alpha\, w_{jt}\ell_j \;=\; \alpha(1-\theta_j)\, y_j \qquad \forall j \in \{1, 2, ..., n\} \tag{2.9}$$

We assume, without loss of generality, that the claims to capital are denominated in the same units as physical capital, so that:

$$\sum_{i=1}^{n} k_{it+1} \;=\; \sum_{j=1}^{n} s_{jt} \tag{2.10}$$

Define a_{ij} as the assets purchased in country i by investors from country j. Thus, the following two accounting relationships hold: 1) Country i's supply of physical capital k_i equals the sum of all units of financial capital invested from all countries j

$$k_i \;=\; \sum_{j=1}^{n} a_{ij} \;; \tag{2.11}$$

2) The total financial capital supplied by country j to all countries i must equal total country j's total savings a_j:

$$s_j \;=\; \sum_{i=1}^{n} a_{ij} \tag{2.12}$$

2.3 Asset Allocation

Define σ_{ij}, the share of capital invested in country i as a percentage of country$j$'s aggregate saving:

$$\sigma_{ij} \;\stackrel{\text{def}}{=}\; \frac{a_{ij}}{s_j} \tag{2.13}$$

Writing k_i, s_j and σ_{ij} in linear algebra notation, we have the following equation describing the flow of capital across countries:

$$\mathbf{k} \;=\; \mathbf{\Sigma}\,\mathbf{s} \;\;: \qquad \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} \;=\; \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{n1} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} \tag{2.14}$$

We next describe how agents allocate capital across countries – that is, how the matrix $\mathbf{\Sigma}$ is determined in the steady-state equilibrium.

We assume that capital investment is lumpy: atomistic investor z in country j invests their savings by buying claims to the return on the capital from of one plant x, anywhere in the world.

Asset markets are imperfect: investment is intermediated by an agent that collects a fee from investors. This fee depends on the investors' asset allocation choice and is then rebated back to investors. Specifically, if investor z chooses to invest in plant x (located in country i), they receive the following subjective return $R(x, z)$:

$$R(x,z) \;\stackrel{\text{def}}{=}\; r_i(x) \cdot e^{-\tau(x,z)} \tag{2.15}$$

where $r_i(x)$ is the return on capital in country i where plant x is located, while $\tau(x, z)$ is the intermediation cost incurred by investor z to invest in plant x. We intend this term to capture the costs associated with, monitoring and enforcing and acquiring information about cross-border investments.

We assume that $\tau(x, z)$ can be decomposed linearly into four components: 1) one component that is systematic at the country pair-level (i, j) and which depends on a metric of distances between the investor and the destination country; 2) a second component that also varies by country pair (i, j) and which captures the severity of regulatory restrictions to capital inflows in country i and to capital outflows in country j; 3) a random idiosyncratic component that varies at the investor-plant pair level (x, z); 4) an investor (j) country-level rebate:

$$\tau(x, z) = \mathbf{d}'_{ij} |\beta| - \log(\phi_{ij}) - \xi(x, z) - T_j \tag{2.16}$$

The components of the vector

$$\mathbf{d}_{ij} = \begin{bmatrix} d^1_{ij} & d^2_{ij} & \cdots & d^D_{ij} \end{bmatrix}' \tag{2.17}$$

are bilateral measures of distance between country i and country j. By allowing the distance vector $\mathbf{d}$ to be multi-dimensional (where D is the dimensionality), we can allow cross-border investment costs to depend on measures of distance other than physical, such as cultural and institutional.

ϕ_{ij} is defined over the interval $[0, 1]$. Therefore $\phi_{ij} \to 0$ implies $\tau \to \infty$: that is, capital controls prevent investing from j to i. $\phi_{ij} = 1$ implies that investment from j to i is unrestricted. For domestic investors $(i = j)$ ϕ_{ij} is always 1 by definition.

$\xi(x, z)$ is an idiosyncratic component that is specific to the (x, z) investor-plant pair and is assumed to be drawn from an Extreme Value Type I (i.e., Gumbel) distribution.

The term T_j is a proportional rebate that is equal for all investors from country j, and is determined in such a way that the investment intermediary of country j makes zero profits. Hence, τ_{ij} does not (directly) affect the aggregate resource constraint: it distorts asset allocation, but it does not actually destroy capital.

In the model, there is no explicit uncertainty.[1] Hence, each agent z invests their savings s_j in the plant that yields the highest subjective return:

$$x^*(z) = \underset{x}{\operatorname{argmax}} \left\{ \log r_i(x) + \log \phi_{ij}(x, z) + \mathbf{d}_{ij}(x, z)' \beta + \xi(x, z) \right\} \tag{2.18}$$

Next, we aggregate the choices of individual investors. Following the seminal result of McFadden (1973), we have that the share of country j assets invested in plant x is:

$$\sigma_j(x) = \frac{\phi_{ij} \cdot r_i \cdot \exp\left(\mathbf{d}'_{ij} \beta\right)}{\sum_{x=1}^{K/\lambda} \phi_{cj} \cdot r_c \cdot \exp\left(\mathbf{d}'_{cj} \beta\right)} \tag{2.19}$$

It is worth pausing to analyze this equation. The idiosyncratic component $\xi(x, z)$, which has been averaged out in the aggregate, has the effect of making the portfolio shares imperfectly inelastic with respect to the rate of returns r_i. In other words, while at the investor-level $\xi(x, z)$ captures investor

[1]For simplicity, we abstract from modeling randomness explicitly. Investors' incentives to diversify across countries do not depend on uncertainty and risk aversion, but on their idiosyncratic preferences over different locations. The framework could be extended to allow for explicit risk, with broadly similar results.

heterogeneity, at the *aggregate* level it acts as a motive for diversification. We see this as a parsimonious way to introduce diversification in a setting where data on individual investors' portfolios and individual assets' returns, risks and covariances are unavailable (we only see country-level investment positions).

Because the probability of investing in any of the k_i/λ plants in country i is the same, we can then sum these probabilities at the level of destination country:

$$\sigma_{ij} \overset{\text{def}}{=} \sum_{x \in i} \sigma_j(x) = \frac{k_i}{\lambda} \cdot \sigma_j(x) \qquad \forall\, x \in j \in \{1, 2, ...n\} \tag{2.20}$$

By summing across plants (within destination country) we obtain the following equilibrium expression for σ_{ij} - the share of country j's foreign asset positions in destination country i:

$$\sigma_{ij} = \frac{\phi_{ij} \cdot r_i k_i \cdot \exp\left(\mathbf{d}'_{ij}\beta\right)}{\sum_{c=1}^{n} \phi_{cj} \cdot r_c k_c \cdot \exp\left(\mathbf{d}'_{cj}\beta\right)} \tag{2.21}$$

which mimics Eaton and Kortum (2002)'s equation for international trade shares. Note that this expression does not depend on plant size (λ). Hence, for all practical purposes, we can ignore the fact that k_i may not be divisible by λ, because we can make plants arbitrarily small ($\lambda \to 0$) without affecting country portfolios.

2.4 Efficient Allocation of Capital

Let Y be World GDP:

$$Y \overset{\text{def}}{=} \sum_{i=1}^{n} Y_i \tag{2.22}$$

and let us call a vector $\mathbf{k} = (k_1, k_2, ..., k_n)'$, a *capital allocation.* Because labor is fixed and the production function is exogenous, Y is a function of $\mathbf{k}$ alone. We say that an allocation $\mathbf{k}$ is *efficient* if it maximizes World GDP Y for a given level of world capital $K \overset{\text{def}}{=} \sum_{i=1}^{n} k_i$.

Int this sub-section, we show that equilibrium in input and asset markets, combined with the absence of systematic distortions to international investment ($\phi_{ij} = 1$ for all ij and $\mathbf{d}'_{ij}\beta$ is a constant) coincides with World GDP maximization for a given level of world capital K. Input markets equilibrium implies that the marginal product of capital in country i is equal to the objective rate of return on capital r_i.

We start by showing that a necessary and sufficient condition for World GDP maximization is that the rates of returns on capital are equalized across countries. To show necessity, consider the first-order Taylor approximation for the change in Y following a change $\Delta\mathbf{k}$ such that $\sum_i \Delta k_i = 0$:

$$\Delta Y \approx \sum_{i=1}^{n} r_i \Delta k_i \tag{2.23}$$

then, if $r_i > r_j$ for some (i,j), we can construct a $Y-$increasing $\Delta\mathbf{k}$ by simply reallocating an arbitrarily-small amount of capital from j to i.

To show sufficiency, notice that we can write country i's capital stock as a strictly-decreasing function of the common rate of return r:

$$k_i \;=\; r^{-\frac{1}{1-\theta_i}} \left(\theta_i \omega_i\right)^{\frac{1}{1-\theta_i}} \ell_i \tag{2.24}$$

This implies that K and Y are also strictly-decreasing functions of r. As a consequence, it is not possible to change r and increase Y without also increasing K.

Next, we consider the relationship between asset markets equilibrium and efficient capital allocation. First, notice that, if asset markets are unaffected by distance ($\mathbf{d}'_{ij}\beta$ is invariant across i,j), there are no capital controls ($\phi_i = 1$ for all i) and investors are optimizing, then equation (2.21) implies that all countries have their capital invested in identical destination country portfolios:

$$\sigma_{ij} \;=\; \frac{r_i k_i}{\sum_{c=1}^{n} r_c k_c} \;=\; \frac{k_i}{K} \quad \forall \; (i,j) \tag{2.25}$$

The right hand side of the equation is a consequence of the fact that the share of capital invested in country i (σ_{ij}) is independent of the origin country j. Because the right part of equation (2.25) is true if and only if rate of returns are equalized, we have thus shown that, provided that companies and investors are optimizing, the following three statements are equivalent: 1) capital is efficiently allocated; 2) rates of returns are equalized across countries; 3) asset markets are undistorted by distance ($\mathbf{d}'_{ij}\beta$ is invariant across i,j) and capital controls ($\phi_i = 1$ for all i). Notice that this is not simply a re-statement of the first Welfare Theorem, since this is a statement about GDP, not welfare.

Moreover, we've also shown that, if asset markets are in equilibrium and undistorted, all origin countries j hold identical portfolios of foreign assets (σ_{ij} is independent of j) - implying that there can be no domestic bias in the aggregate portfolio such case.

Having shown that efficient capital allocation is equivalent to rates of return being equalized, the next step is to show formally that capital *misallocation* manifests itself as cross-country dispersion in the rate of return on capital. We do that by considering a second-order Taylor approximation of the change in World GDP around an efficient $\mathbf{k}$[2]:

$$\Delta Y \quad \approx \quad r \sum_{i=1}^{n} \Delta k_i - \frac{1}{2} \sum_{i=1}^{n} \left(1-\theta_i\right) \frac{r}{k_i} \left(\Delta k_i\right)^2 \tag{2.26}$$

Because we want to study capital misallocation, we consider a $\Delta\mathbf{k}$ that leaves K unaffected: this implies that the first-order term of the equation above is zero. We can then divide both sides by world GDP and

[2]To derive this expression, recall that, in equilibrium, the rate of return r is equal to the marginal product of capital

rearrange the second-order term as:

$$\Delta \log Y \;\approx\; -\frac{1}{2}\sum_{i=1}^{n}\left(1-\theta_i\right)\frac{rk_i}{Y}\left(\Delta \log k_i\right)^2 \tag{2.27}$$

We then use the following facts:

$$\Delta \log r_i \;=\; -\left(1-\theta_i\right)\Delta \log k_i \tag{2.28}$$

$$r_i k_i \;=\; \theta_i y_i \tag{2.29}$$

to re-write the second-order change in World GDP as:

$$\Delta \log Y \quad \approx \quad -\frac{1}{2}\sum_{i=1}^{n}\frac{\theta_i}{1-\theta_i}\cdot\frac{y_i}{\mathrm{Y}}\cdot\left(\Delta \log r_i\right)^2 \tag{2.30}$$

This expression can be seen as a weighted measure of dispersion of the rate of return r_i across countries.

2.5 Gravity

Since, in equilibrium, the capital share of income is equal to GDP times θ_i, Equation (2.21) can be rearranged as follows:

$$\sigma_{ij} \;=\; \frac{\phi_{ij}\,\theta_i y_i\cdot\exp\left(\mathbf{d}'_{ij}\beta\right)}{\sum\limits_{c=1}^{n}\phi_{cj}\,\theta_c y_c\cdot\exp\left(\mathbf{d}'_{cj}\beta\right)} \tag{2.31}$$

The denominator of this expression can be interpreted as a distance-discounted measure of the total size of the global market for capital that is available to country j investors. We shall call this G_j:

$$G_j \;\stackrel{\text{def}}{=}\; \sum_{c=1}^{n}\phi_{cj}\,\theta_c y_c\cdot\exp\left(\mathbf{d}'_{cj}\beta\right) \tag{2.32}$$

Multiplying both sides by s_j and using the fact that $s_j=\alpha\left(1-\theta_j\right)y_j$, equation (2.31) can be rearranged as a gravity equation:

$$\boxed{\textbf{Gravity}}\;:\quad a_{ij} \;=\; \phi_{ij}\cdot\theta_i\left(1-\theta_j\right)\cdot\frac{\alpha}{G_j}\cdot\frac{y_i\cdot y_j}{\exp\left|\mathbf{d}'_{ij}\beta\right|} \tag{2.33}$$

2.6 Global Capital Markets Clearing

To close the model, we find the vector of capital stocks $\mathbf{k}$ that simultaneously clears the market for inputs and assets. First, the matrix of country shares $\mathbf{\Sigma}$ is a function of the vector of country output $(\mathbf{y})$ and the matrices of distortions $\mathbf{\Delta} = [\mathbf{d}'\beta_{ij}]$ and $\mathbf{\Phi}$– that is:

$$\mathbf{\Sigma} \;=\; \mathbf{\Sigma}\left(\mathbf{y};\mathbf{\Theta},\mathbf{\Delta},\Phi\right) \tag{2.34}$$

$$\text{with} \qquad \mathbf{\Delta} \overset{\text{def}}{=} \begin{bmatrix} \mathbf{d}'\beta_{11}, & \mathbf{d}'\beta_{12}, & \dots & \mathbf{d}'\beta_{1n} \\ \mathbf{d}'\beta_{21} & \mathbf{d}'\beta_{22} & \dots & \mathbf{d}'\beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{d}'\beta_{n1} & \mathbf{d}'\beta_{n2} & \dots & \mathbf{d}'\beta_{nn} \end{bmatrix} \tag{2.35}$$

$$\mathbf{\Theta} \overset{\text{def}}{=} \begin{bmatrix} \theta_1, & 0, & \dots & 0 \\ 0 & \theta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_n \end{bmatrix} \quad \text{and} \quad \Phi \overset{\text{def}}{=} \begin{bmatrix} 1, & \phi_{12}, & \dots & \phi_{1n} \\ \phi_{21} & 1 & \dots & \phi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n1} & \phi_{n2} & \dots & 1 \end{bmatrix} \tag{2.36}$$

Since $\mathbf{s}$ can be written as $\alpha\left(\mathbf{I}-\mathbf{\Theta}\right)\mathbf{y}$ and $\mathbf{y}$ in turn can be written as a function of $\mathbf{k}$, equation (2.14) can be re-written as:

$$\mathbf{k} \quad = \quad \mathbf{\Sigma}\left(\mathbf{y}\left(\mathbf{k}\right);\mathbf{\Theta},\mathbf{\Delta},\Phi\right)\cdot\alpha\left(\mathbf{I}-\mathbf{\Theta}\right)\mathbf{y}\left(\mathbf{k}\right) \tag{2.37}$$

where $\mathbf{I}$ is the identity matrix. The market-clearing equilibrium vector of capital stocks $\mathbf{k}$ is then determined as the fixed point of equation (2.37). There is a trivial equilibrium at $\mathbf{k} = 0$. When we solve equation (2.37) numerically, we can rule out the trivial equilibrium by taking logs of both sides of the equation.

The consumption of final good by country j (by old and young agents) balances the domestic consumers' budget:

$$c_j \;=\; \mathbf{r}'\mathbf{a}_j + \left(1-\theta_j\right)w_j\ell_j \tag{2.38}$$

where

$$\mathbf{a}_j \;=\; \begin{bmatrix} a_{1j} & a_{2j} & \cdots & a_{nj} \end{bmatrix}' \tag{2.39}$$

and the following equation balances country j's current account:

$$\underbrace{c_j - \left(y_j - s_j\right)}_{\text{Net Imports}} \quad = \quad \mathbf{r}'\mathbf{a}_j - \alpha w_j\ell_j - r_jk_j + s_j \quad = \quad \underbrace{\mathbf{r}'\mathbf{a}_j - r_jk_j}_{\text{Net Foreign Income}} \tag{2.40}$$

That is, all consumption in excess of production (net of savings) is financed by a positive net foreign capital income. Vice-versa, a negative balance in net foreign income has to be balanced by a trade surplus.

3 Specification and Data

3.1 Econometric Model

Our regression equation can be estimated using a linear-in-logs specification. We assume that capital flows are observed with a multiplicative error term:

$$\hat{a}_{ij} \; = \; a_{ij} \cdot \exp\left(\varepsilon_{ij}\right) \qquad \text{with} \qquad \mathbf{E} \exp\left(\varepsilon_{ij}\right) = 1 \tag{3.1}$$

implying the following logit demand system for international assets:

$$\hat{\sigma}_{ij} \; = \; \frac{\phi_{ij}\,\theta_i y_i \cdot \exp\left(\mathbf{d}'_{ij}\beta + \varepsilon_{ij}\right)}{\sum_{c=1}^{n} \phi_{cj}\,\theta_c y_c \cdot \exp\left(\mathbf{d}'_{cj}\beta + \varepsilon_{cj}\right)} \tag{3.2}$$

By defining:

$$\gamma_j \quad \stackrel{\text{def}}{=} \quad \log s_j - \log \sum_{c=1}^{n} \left[\phi_c\,\theta_c y_{ct} \cdot \exp\left(\mathbf{d}'_{cj}\beta + \varepsilon_{cj}\right)\right] \tag{3.3}$$

we can re-write equation (3.2) as the following fixed effects linear regression model for the log of foreign investment:

$$\log \hat{a}_{ij} \; = \; \mu_i + \gamma_j + \mathbf{d}'_{ij}\beta + \log \phi_{ij} + \varepsilon_{ij} \tag{3.4}$$

where μ_i is a country of origin fixed effect and γ_j is a country of origin fixed effect. y_i, θ_i and the mean of ε_{ij} are absorbed by origin and destination country fixed effects. The parameter ϕ_{ij} is a function of observed capital controls between i and j according to the following equation:

$$\phi_{ij} \; = \; \exp\left(\text{Capital Controls}_{ij}\beta_\phi\right) \tag{3.5}$$

Empirically, our measure of capital controls is the sum of an $i-$level variable and a $j-$level variable, and therefore it is absorbed by i and j fixed effects. To address this issue, we exploit the fact that our foreign investment and capital control data have a panel structure, and expand the gravity equation above to include the time dimension:

$$\log \hat{a}_{ijt} \; = \; \mu_i + \gamma_j + \psi_t + \mathbf{d}'_{ij}\beta + \text{Capital Controls}_{ijt}\beta_\phi + \varepsilon_{ijt} \tag{3.6}$$

where ψ_t is a year fixed effect.

This is our main econometric specification. The dependent variable is measured using data on *Foreign Direct Investment*, *Foreign Portfolio Investment*, and the sum of the two (*Foreign Total Investment*). To capture $\mathbf{d}$, we propose a parsimonious specification based on three measures of distance, hypothesized to capture major impediments to international investments: *Geographic Distance, Cultural Distance*, *Institutional Distance.* To capture ϕ_{ij}, we turn to a widely-used index of capital controls.

Because the vector of distances $\mathbf{d}$ varies at the level of the *undirected* country pair, while ϕ_{ij} varies by *directed* country pair and over time (but is serially correlated over time), in our regression analysis we compute double-clustered standard errorrs (by directed and undirected country pair); we have also computed standard errors using alternative clustering structures[3].

Additional bilateral variables are used either as instruments or control variables, depending on the specific empirical model under consideration. We now turn to describing these variables in more detail.

3.2 Dependent Variables: Restated Foreign Investment Data

Our analysis improves upon the empirical literature on the determinants of foreign investment by using recently-developed foreign investment data that accounts for the existence of tax havens. These tax havens may serve as indirect conduits between the origin and destination countries. For instance, the Cayman Islands are often used to transit funds between origin and destination countries in a way that is tax efficient. In recent work, Damgaard, Elkjaer, and Johannesen (2019) combined FDI data from the IMF's Coordinated Direct Investment Survey (CDIS) and the OECD's Foreign Direct Investment statistics. They restated the data in order to account for the fact that some countries act as offshore investment centers. In such countries, there is a high concentration of investment companies that only act as investment vehicles, and do not actually engage in productive activities. Damgaard, Elkjaer, and Johannesen (2019) used cross-border entity ownership data from Bureau Van Dijk's Orbis to reallocate asset ownership from country of residence of the investment vehicle to the nationality country of the ultimate investor, thereby correcting for artificially inflated numbers pertaining to offshore tax havens.

Regarding portfolio investment, our main source is the IMF's Coordinated Portfolio Investment Survey (CPIS). We restate CPIS data to account for the presence of shell companies in tax havens - often used to issue securities. To do so, we use *reallocation matrices* produced by Coppola, Maggiori, Neiman, and Schreger (2020) based on fund holdings data from Morningstar. Using these reallocation matrices, we can convert international portfolio data from CPIS from a residency basis to a nationality basis. In other words, both our FDI and FPI series are corrected for measurement imperfections associated with tax havens.

For both FDI and FPI, we base our econometric estimates on panel data from 2009 to 2017. Figure 1 displays these FDI and FPI data for 2017, plotted against each other on a logarithmic scale. The plot reveals some interesting facts. First, there is a great deal of variation in both FDI and FPI across countries. These two variables range from a few hundreds of thousands dollars to over a trillion dollars. This is the variation we seek to explain. Second, the two variables correlate very strongly ($\rho = 0.63$), and line up neatly on the 45° line, indicating that they are similar in size and tend to track each other

[3]We make these are available upon request as additional material

Figure 1: Foreign Portfolio and Direct Investment

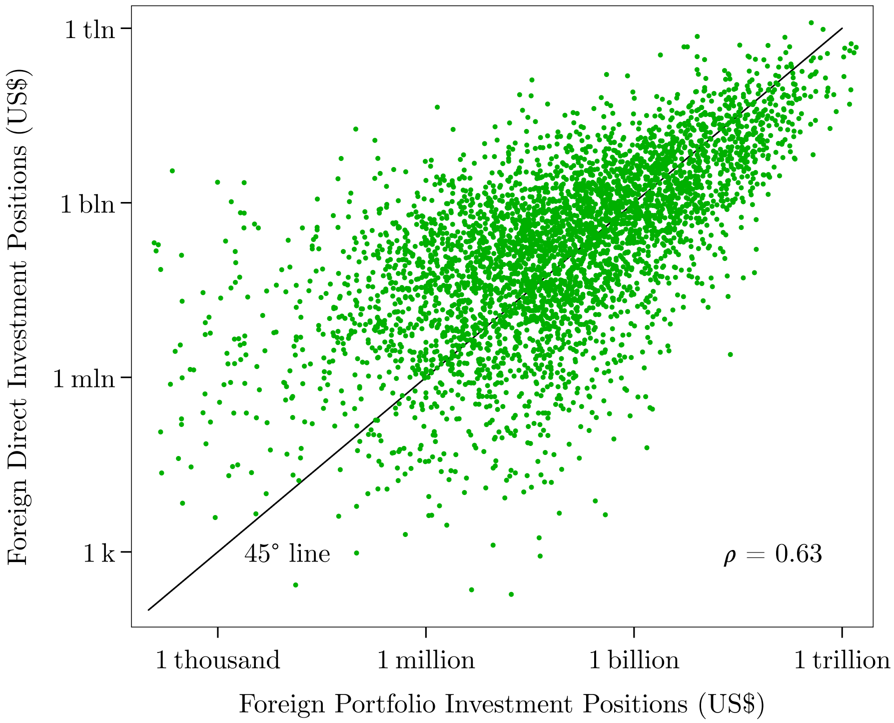

Figure Notes: the figure above plots the correlation between the restated Foreign Direct Investment data of Damgaard et al. (2019) and the restated Foreign Portfolio Investment data of Coppola et al. (2020). Each observation is a country pair and all data refers to the year 2017. The unit of measurement is US dollars at current prices. Log scale on both axes.

closely. This suggests that they might be driven by a similar set of underlying factors, an issue that our econometric analysis will clarify.

3.3 Proximate Determinants

Barriers to global capital allocation are captured in our model by the distance metric $\mathbf{d}$. We operationalize these barriers by modeling them in a parsimonious way as the result of institutional distance, cultural distance, geographic distance, and policy impediments to financial flows.

3.3.1 Institutional Distance

We construct a measure of *Institutional Distance,* defined as the sum of three dummy variables: i) The first dummy variable is equal to one if the corresponding country pair has a common currency . The presence of a common currency is a form of common institutional arrangement believed to facilitate cross-border financial flows by reducing transactions costs and eliminating exchange rate uncertainty (see Maggiori, Neiman, and Schreger, 2020). ii) The second dummy variable equals one if the legal systems of the countries in the corresponding country dyad share the same legal origin, as defined by La Porta, Lopez-de Silanes, and Shleifer (2008). Dissimilarity in legal origins is likely to make the enforcement of cross-border contracts more difficult, and therefore to reduce financial flows, particularly for FDI. iii) The third dummy reflects the existence of a Dual Taxation Treaty (DTT) as of 2012, constructed by Petkova, Stasio, and Zagler (2019) based on the tax treaty database of the International Bureau of Fiscal Documentation (IBFD). DTTs can strongly affect the ex-post return from investing abroad. Our *Institutional Distance* indicator varies between 0 and 3.

3.3.2 Cultural Distance

Our measure of *Cultural Distance* captures distance in contemporary values and beliefs, introduced by Spolaore and Wacziarg (2016). It is constructed using a set of 98 questions from the World Values Survey 1981- 2010 Integrated Questionnaire, reflecting the following question categories: a) perceptions of life; b) environment; c) work; d) family; e) politics and society; f) religion and morale; g) national identity. These questions are a subset of a broader set of 740 questions, where the subset was chosen to ensure that the questions used to compute bilateral distances remains relatively similar across pairs. For each question, the measure consists of the Euclidian distance in answers between country pairs. Distances are then averaged over questions to obtain a summary index. Averages can be computed by question category, but here we use the average over all underlying 98 questions. We re-scaled this index to span the $[0, 3]$ interval, so that the magnitude of its effect can be compared to that of *Institutional Distance.*

3.3.3 Geographic Distance

We obtained country dyad-level data on physical distance from CEPII's GeoDist dataset (Mayer and Zignago, 2011). *Geographic Distance* measures the geodesic distance between any two countries, based on a population-weighted average of the distances between individual cities. As for *Cultural Distance*, we have re-scaled this variable (whose maximum value is equal to half the earth's circumference) to the $[0, 3]$ interval, so that the magnitude of its effect can be compared to that of *Institutional Distance* and *Cultural Distance.*

3.3.4 Capital Controls

To capture capital account openness, we rely on the dataset recently developed by Fernández, Klein, Rebucci, Schindler, and Uribe (2016, henceforth FKRSU). They produce measures of inward and outward openness, called respectively KC10^{in} and KC10^{out}, based on IMF's Annual Report on Exchange Arrangements and Exchange Restrictions (AREAER). These indices are based on the presence or absence of a series of underlying legal impediments to cross-border financial flows (10 in total) coded by the authors from AREAER into dummy variables, and then summed. These two indices are available for 100 countries between 1995 and 2017.

The overlap in terms of countries between FKRSU's dataset and our own is not perfect. For countries that have missing data we predict KC10 using the KAOpen index developed by Chinn and Ito (2006), which is also based on AREAER. It is significantly less granular (it is based on three variables) and does not distinguish between inward and outward openness, but it has the advantage of being available for 182 countries from 1970 to 2018.

Both KC10 and the Chinn-Ito index are measures of *de jure* capital account openness. Unlike the previous three proximate determinants, they are time-varying variables. Since we include country of origin and destination fixed effects in our empirical model, the effect of capital controls on foreign investment will be estimated only from variation across time, within destination country. For our gravity model, we define the index *Capital Controls* as the sum between the value of KC10^{in} in the country of destination i and KC10^{out} in the origin country, except for $i = j$, where we assume no restrictions on domestic investment. Formally:

$$\text{Capital Controls}_{ij} = \begin{cases} \left(\text{KC10}_i^{\text{in}} + \text{KC10}_j^{\text{out}}\right) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \tag{3.7}$$

3.4 Instruments and Control Variables

We consider additional variables as direct determinants of capital flows, to be used either as instruments or as control variables, depending on the specification under consideration.

3.4.1 Linguistic, Ancestral and Religious Distances

The first category of variables consists of measures of historical distinctiveness that can be considered as deep determinants of contemporary institutional and cultural distances. We use measures of linguistic distance and religious distance introduced in Fearon (2003), Mecham, Fearon, and Laitin (2006) and Spolaore and Wacziarg (2016). We also use measures of ancestral distance initially developed by Spolaore and Wacziarg (2009) and later updated in Spolaore and Wacziarg (2018). As these contributions discuss in detail, linguistic distance, religious distance and ancestral distance can all be interpreted as measures of historical relatedness between populations.

Consider first *Linguistic Distance*. Different contemporary languages have descended from common ancestral languages over time. For instance, German, Italian and French all descend from a common proto-Indoeuropean language. In turn, Italian and French descend from more recent common ancestral languages (Romance languages stemming from Latin), while German does not. Thus, Italian and French are more closely related to each other than either is to German. Intuitively, this is analogous to our concept of relatedness between individuals: two siblings are more closely related to each other than they are to their first cousins, because they share more recent common ancestors (their parents) with each other, while they share more distant ancestors with their first cousins (their grandparents) and second cousins (great-grandparents). Formally, our measures of linguistic distance are computed by counting the number of different linguistic nodes separating any pair of languages, according to their classification from Ethnologue.

Religious Distance is also constructed considering number of nodes in historical trees. In this case, the trees consist of religions grouped in related historical categories. For instance, Near Eastern monotheistic religions are subdivided into Christianity, Islam and Judaism. These are further divided into finer levels of disaggregation. The number of common nodes between religions is our metric of religious proximity. Thus, Baptists are closer in religious space to Lutherans than they are to the Greek Orthodox.

Ancestral Distance, like *Linguistic Distance* and *Religious Distance*, can be interpreted as a measure of long-term historical separation times between groups. However, while linguistic distance captures how far back one must go in order to find the common ancestral language from which two modern languages descend, *Ancestral Distance* captures how far in the past one must go in order to find the common ancestral population from which two contemporary populations descend. In this respect, the measure is intuitively closer to the concept of relatedness between individuals, as applied to whole populations. *Ancestral Distance*, here, is computed using a genomic dataset on human microsatellite variation from Pemberton et al. (2013), covering 267 world populations. These populations were matched to 1,120 ethnic groups from Alesina et al. (2003).[4]

It would not be appropriate to interpret any effect of ancestral distance on any economic or societal outcomes as a direct causal effect of genetic factors. Rather, measures of distance based on microsatellite variation provide information on historical relatedness between populations. This neutral feature of *Ancestral Distance* is crucial for the interpretation of the relationship between historical relatedness and contemporary cultural distance. As populations separate over time, they are more likely to diverge in cultural traits that are transmitted randomly and with variation from one generation to the next. Thus, we should expect a positive relationship, on average, between ancestral distance and cultural or institutional distance, whereas there is no reason to expect ancestral distance to directly affect financial flows. Consequently, we can use ancestral distance as an instrument for the proximate determinants of financial flows.

3.4.2 Additional Variables

We use a variety of additional bilateral measures either as instruments or control variables. Among them are several measures of geographic distance - contiguity, access to a common sea or ocean, latitudinal distance, longitudinal distance, and whether the two countries in pair are on the same continent. Addi-

[4]Microsatellites are tracts of DNA in which specific sequences of base pairs are repeated. Microsatellites tend to mutate rapidly and randomly - that is, microsatellite variation mostly captures neutral change that is not subject to natural selection. Therefore, their variation provides no direct information about overall differences in genetic endowments.

tionally, we consider the length of the diplomatic relationship between the two countries in a pair, as a measure of the depth of their historical links. We also consider variables called *Colonial Relationship* - capturing whether two countries in a pair were ever in a colonizer-colonized relationship, and *Common Colonizer*, denoting whether the two countries in a pair ever had a common colonizer.[5]

Finally, we control for a measure of trade costs, because trade costs can induce changes in international investment. For instance, high trade costs can spur FDI in an effort to "jump" tariffs. Or, on the contrary there may be complementarities between trade in capital and trade in goods: the return to investment in a foreign country may be lower if exporting from the destination is costly, or if the investment requires paying tariffs to import capital goods into the destination country. The source of the trade cost data is the ESCAP-World Bank Trade Cost Database (2020), as initially developed in Novy (2013). This paper derives time-varying bilateral trade costs from a gravity model, which is solved analytically so that trade costs can be inferred using observed trade data. The ESCAP-World Bank Trade Cost Database[6] updates these calculations periodically, and estimates of trade costs are now available for a wide set of country pairs over the 1995-2018 period.

3.5 Coverage and Summary Statistics

After merging all the variables above, we are left with a dataset of 72 countries, $72 \times 72 = 5{,}184$ directed country pairs-observations (exclude diagonal i-to-i pairs) or 2,556 undirected country pairs, over the 2009-2017 period. The 72 countries in our dataset cover about 93% of the World GDP (based on 2017 data from the Penn World Tables, version 9.1). Directed Foreign Investment data (combining direct and portfolio investment) is available for 2,823 (or 62%) of these observations.

Table 1 displays summary statistics for the data described above.

4 Econometric Analysis

In this section, we estimate the parameter vector β, the effect of geographic, cultural and institutional distances on log foreign investment (a set of three semi-elasticities), and β_ϕ, the coefficient on capital controls. Our objective is not only to provide a quantitative assessment of the statistical impact of cross-border investment frictions, but also to retrieve structural parameters for the model that we have presented in Section 2.

[5]The data is from CEPII and can be obtained at http://www.cepii.fr/CEPII/fr/bdd_modele/presentation.asp?id=6

[6]https://www.unescap.org/resources/escap-world-bank-trade-cost-database

Table 1: Summary Statistics

Panel A: Directed, Time-Varying Variables

	Observations	Mean	StDev	Min	Max
Foreign Total Investment (US$ mln)	23,131	20,182	96,741	0	2,060,000
Foreign Direct Investment (US$ mln)	33,359	5,322	33,192	0	955,401
Foreign Portfolio Investment (US$ mln)	28,913	21,282	496,493	0	35,100,000
Capital Controls $\left(\mathrm{KC10}_i^{\mathrm{in}} + \mathrm{KC10}_j^{\mathrm{out}}\right)$	46,656	7.247	4.678	0	20
Trade Costs	44,558	0.052	0.046	0	1.322

Panel B: Distance Variables (Undirected, Cross-Sectional)

	Undirected Pairs	Mean	StDev	Min	Max
Cultural Distance	2,628	1.266	0.519	0.000	3.000
Geographic Distance	2,628	0.949	0.704	0.001	2.929
Institutional Distance	2,628	2.040	0.818	0.000	3.000

Panel C: Instrumental and Control Variables (Undirected, Cross-Sectional)

	Undirected Pairs	Mean	StDev	Min	Max
Ancestral Distance	2,557	0.021	0.017	0.000	0.077
Common Colonizer	2,628	0.029	0.168	0.000	1.000
Length of Diplomatic Tie	2,556	52.073	51.055	0.000	203.000
Religious Distance	2,418	0.788	0.210	0.000	0.999
Colonial Relationship	2,628	0.024	0.154	0.000	1.000
Common Sea	2,556	0.065	0.246	0.000	1.000
Contiguity	2,628	0.035	0.183	0.000	1.000
Latitudinal Distance	2,556	28.738	25.188	0.000	106.000
Linguistic Distance	2,418	0.936	0.191	0.000	1.000
Longitudinal Distance	2,556	61.423	52.847	0.000	276.000
Same Continent	2,628	0.339	0.443	0.000	1.000

4.1 Least Squares Analysis

We begin by performing an OLS regression of the log of foreign investment (FTI, FDI or FPI) on all our proximate measures of distance. In all regressions, we include investor country and destination country fixed effects as well as year fixed effects. Table 2 reports the estimates. Column (1) presents our estimation results for log FTI as the dependent variable. We find that *Cultural Distance, Institutional Distance* and *Geographic Distance, Capital Controls* are all statistically and economically significant predictors of FTI: the slope coefficients corresponding to these three variables are negative, sizable in magnitude (-1.167, -0.629, -1.781, -0.043) and statistically significant at the 99% confidence level. To get a notion of relative magnitudes, the coefficients can be expressed as the effect of a one standard deviation change in the independent variables in terms of a percentage change in FTI ($\%\Delta FTI = e^{\beta_x \Delta x} - 1$). We find that the largest effect is that of geographic distance (a one standard deviation increase of 0.704 units is associated with a 71.4% decrease in FTI), followed by cultural distance (a one standard deviation increase of 0.519 units is associated with a 45.4% decrease in FTI). The effect of institutional distance is also sizable (a one standard deviation increase of 0.818 units is associated with a 40.2% decrease in FTI). The standardized effect of *Capital Controls* is the smallest (a one standard deviation increase of 4.678 units is associated with a 18.2% decrease in FTI)

In Column (2) we present our estimation results for the log FDI model. We find that all four distance metrics are statistically and economically significant predictors of FDI: the standardized effects as defined above are essentially unchanged compared to the estimates for FTI. We find that *Capital Controls* is not statistically-significant, but the magnitude of its effect is reduced: a one standard deviation increase in *Capital Controls* (4.678 points) is associated with a 5.90% increase in log FDI. Column (3) considers portfolio investment as the dependent variable. We find a smaller effect of geographic distance (a standardized effect of -59.8%) and a larger effect of capital controls (with a standardized effect of -18.76%), which is now statistically significant at the 5% level: it appears that restrictions on financial flows restrict portfolio investments more than foreign direct investment. Cultural and institutional distances continue to have a meaningful effect of portfolio investment, with standardized magnitudes of -47.1% and -31.6%. For FPI, the effect of *Capital Controls* is significantly larger, -0.076, which corresponds to a standardized magnitude of -29.9%. The likely reason why this variable has a much stronger effect on FPI than on FDI is that, by construction, the variable focuses much more on portfolio flows restriction (only one of the 10 dimensions focuses on direct investment).

Finally, columns (4) through (6) repeat the analysis of the first three columns, but depart from our very parsimonious specification by adding controls for a variety of geographic distance metrics (contiguity, access to a common sea, latitudinal distance, longitudinal distance) and common history variables (linguistic distance, past colonial relationship). The coefficient estimates on cultural and institutional distance are somewhat reduced in magnitude, while the effect of geographic distance is slightly larger: for FTI, we find standardized effects of cultural distance, institutional distance, geographic distance and capital controls, respectively, of -36.3%, -37.21%, -76.7% and -18.9%. We again find that capital controls have a quantitatively larger effect on FPI than FDI. In sum, adding control variables does not fundamentally alter the inferences drawn from the more parsimonious specification.

Table 2: OLS Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	logFTI	logFDI	logFPI	logFTI	logFDI	logFPI
Cultural Distance	-1.166***	-1.217***	-1.221***	-0.927***	-0.893***	-1.081***
	(0.132)	(0.131)	(0.119)	(0.133)	(0.133)	(0.124)
Institutional Distance	-0.629***	-0.556***	-0.464***	-0.549***	-0.443***	-0.400***
	(0.065)	(0.068)	(0.062)	(0.067)	(0.071)	(0.064)
Geographic Distance	-1.781***	-1.982***	-1.294***	-1.898***	-2.287***	-1.495***
	(0.096)	(0.102)	(0.092)	(0.318)	(0.315)	(0.298)
Capital Controls	-0.043***	-0.013	-0.076***	-0.045***	-0.012	-0.079***
	(0.015)	(0.015)	(0.017)	(0.014)	(0.013)	(0.016)
Control Variables	No	No	No	Yes	Yes	Yes
Observations	21,302	23,636	21,568	19,164	21,308	18,941
R-squared	0.763	0.705	0.781	0.773	0.721	0.792

Table Notes: this table reports OLS estimates of a linear regression of the variables listed on the topmost row (*Foreign Total Investment*, *Foreign Direct Investment*, *Foreign Portfolio Investment*) on the variables in the leftmost column. Each observation is a directed country pair. All regressions include origin and destination country fixed effects as well as year fixed effects. Additional controls in columns 4-6 are *Colonial Relationship, Common Sea, Contiguity, Latitudinal Distance, Longitudinal Distance, Linguistic Distance, Same Continent* and *Trade Costs*. Standard errors (double-clustered by directed and undirected country pair) in parentheses. $^{*}p < .10$; $^{**}p < .05$; $^{***}p < .01$

4.2 Poisson Regressions

One shortcoming of the econometric model described by equation (3.6) is that, being written in logs, it can only accommodate strictly positive capital positions ($\hat{a}_{ij} > 0$). In order to incorporate country pairs with zero investment, we can re-write the regression equation (3.6) as:

$$\hat{a}_{ijt} \;=\; \exp\left(\mu_i + \gamma_j + \psi_t + \mathbf{d}'_{ij}\beta + \text{Capital Controls}_{ijt}\beta_\phi + \varepsilon_{ijt}\right) \tag{4.1}$$

thereby converting the log-linear specification into a Poisson regression. This type of model has been applied to gravity models of trade by Santos Silva and Tenreyro (2006) and Correia, Guimaraes, and Zylkin (2019). We apply the same statistical model to our model of financial positions. In order to avoid using a highly-inefficient estimator (as a consequence of the high degree of heteroskedasticity that is present in the residuals of this equation), we weight observations by the inverse of the geometric mean

Table 3: Poisson Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	FTI	FDI	FPI	FTI	FDI	FPI
Cultural Distance	-0.915***	-0.792***	-1.790***	-0.678***	-0.617***	-0.736***
	(0.141)	(0.184)	(0.278)	(0.130)	(0.142)	(0.169)
Institutional Distance	-0.129**	-0.089	-0.617***	-0.227***	-0.198***	-0.348***
	(0.064)	(0.089)	(0.104)	(0.063)	(0.074)	(0.079)
Geographic Distance	-0.924***	-1.417***	-0.357*	-1.602***	-2.087***	-0.829***
	(0.109)	(0.171)	(0.205)	(0.296)	(0.312)	(0.293)
Capital Controls	-0.025	-0.018	-0.016	-0.029**	-0.007	-0.041***
	(0.016)	(0.021)	(0.011)	(0.011)	(0.013)	(0.014)
Control Variables	No	No	No	Yes	Yes	Yes
Observations	23,131	33,359	28,913	20,883	30,223	25,786

Table Notes: the table above reports Iteratively-Reweighted Least Squares (IRLS) estimates of a Poisson regression of the variables listed on the topmost row (*Foreign Total Investment*, *Foreign Direct Investment*, *Foreign Portfolio Investment*) on the variables in the leftmost column. Each observation is a directed country pair. All regressions include origin and destination country fixed effects as well as year fixed effects. Additional controls in columns 4-6 are *Colonial Relationship, Common Sea, Contiguity, Latitudinal Distance, Longitudinal Distance, Linguistic Distance, Same Continent* and *Trade Costs*. Observations are weighted by the inverse of the geometric average of destination and origin country GDP. Standard errors (double-clustered by directed and undirected country pair) in parentheses. $^{*}p < .10$; $^{**}p < .05$; $^{***}p < .01$

of the GDPs of countries i and j (un-weighted estimates, which have larger standard errors, are shown in Appendix D). Including the zero investment pairs, the size of the sample rises substantially (by 41% for FDI and 34% for FPI, though the increase is smaller for total investment, at 8.6%).

Table 3 displays the resulting estimates. In general, we find that the standardized magnitude of Poisson estimates are slightly smaller than the corresponding OLS and IV estimates. This is especially the case in the specification without control variables for *Institutional Distance* (which is no longer statistically significant for FDI) and for *Capital Controls* (which is no longer significant for either FTI and FPI). The standardized effect of cultural distance is modestly reduced, from about 45% in columns (1)-(3) of Table 2 to about 35% in columns (1)-(3) of Table 3. Similar magnitude reductions are observed for geographic distance.

4.3 Instrumental Variable Regressions

A challenge in estimating the effect of *Cultural Distance* and *Institutional Distance* on bilateral investment positions is the potential for reverse causality: it is conceivable that two countries may converge culturally (by adopting more similar values and norms) or institutionally (by establishing a common currency or a tax treaty) as a consequence of more intense cross-border investment.[7] In that case, the OLS estimates of the gravity equation (3.6) could not be interpreted as causal. To address this issue, we turn to an IV strategy. We rely on the distinction between *proximate* and *deep* determinants of foreign financial positions. We assume that historically-determined factors affect contemporary barriers, but only influence financial flows indirectly, through their effect on contemporary cultural and institutional distance. Consistent with this exclusion restriction, we consider four instruments for cultural and institutional distance: *Ancestral Distance, Common Colonizer, Length of Diplomatic Ties*, and *Religious Distance* (other measures of historical relatedness, like *Linguistic Distance* or *Colonial Relationship*, are used as controls rather than instruments out of concern about their excludability from the second stage). When considering our four instruments, the stronger argument for excludability can be made for *Ancestral Distance*, because it is very plausible that such variable - capturing intergenerational relatedness and based on neutral genetic changes - might only impact contemporary outcomes through its historical effects on the cultural transmission of traits and beliefs, which are captured by our measure of *Cultural Distance.* A theoretical formalization of the relationship between *Ancestral Distance* and *Cultural Distance* is provided in Appendix A. A similar argument can be made for *Religious Distance*, which is also constructed using a branching tree, tracing the historical splits of different religious denominations. Thus, it is plausible that the contemporary effects of such splits on our dependent variable should operate (mainly or exclusively) through contemporary differences in values and beliefs (including, but not limited to, religious beliefs), which are measured by *Cultural Distance.* The other two instruments, *Common Colonizer* and *Length of Diplomatic Ties,* in contrast, should plausibly affect contemporary financial flows mainly by raising the likelihood of contemporary institutional cooperation, captured by *Institutional Distance.*

Our baseline empirical analysis relies on a relatively parsimonious specification where regressors consist of a set of three distance metrics, a measure of capital controls and a small set of additional control variables. A set of deep determinants of the barriers to foreign investment are excluded from the main specification and used as instruments. Of course, such exclusion restrictions can always be questioned. Thus, in the Appendix, we also consider a fully-saturated regression specification where all of the exogenous variables (controls and instruments) are entered at once. The goal is to examine the robustness of the estimated β and β_ϕ coefficients. Results are presented below in Appendix Tables B.1 (OLS) and B.2 (Poisson). Those are to be compared, respectively, to the results in Tables 2 and 3 below. We find that the effects of our four main variables of interest remain negative, statistically significant and that their

[7] For obvious reasons, no such issue arises for geographic distance)

Table 4: First-Stage Regressions

	(1)	(2)	(3)	(4)
	Cultural Distance	Institutional Distance	Cultural Distance	Institutional Distance
Ancestral Distance	6.744***	-1.121	6.295***	-3.732
	(1.367)	(2.422)	(1.316)	(2.491)
Common Colonizer	-0.456***	-0.330***	-0.429***	-0.328***
	(0.052)	(0.117)	(0.047)	(0.113)
Length of Diplomatic Tie	-0.001**	-0.002***	-0.001**	-0.002***
	(0.000)	(0.001)	(0.000)	(0.001)
Religious Distance	1.036***	0.421***	0.876***	0.072
	(0.100)	(0.163)	(0.094)	(0.160)
Control Variables	No	No	Yes	Yes
Observations	20,957	20,957	20,957	20,957
R - squared	0.677	0.440	0.705	0.485

Table Notes: the table above reports Ordinary Least Squares (OLS) estimates of a linear regression of the variables listed on the topmost row on the variables in the leftmost column. Each observation is an undirected country pair. All regressions include origin and destination country fixed effects as well as year fixed effects. All regressions control for *Geographic Distance* and *Capital Controls*. Additional controls in columns 3 and 4 are *Colonial Relationship, Common Sea, Contiguity, Latitudinal Distance, Longitudinal Distance, Linguistic Distance, Same Continent* and *Trade Costs*. Robust standard errors in parentheses. $^{*}p < .10$; $^{**}p < .05$; $^{***}p < .01$

magnitudes are generally preserved. [8]

Table 4 presents estimation results for the first-stage regressions corresponding to *Cultural Distance* and *Institutional Distance*. We present results for the parsimonious specification (columns 1 and 2), and for a specification with additional controls (columns 3 and 4).

The first stage regressions lead to interesting results. Consistent with findings in Spolaore and Wacziarg (2016), ancestral and religious distances are both positively correlated with *Cultural Distance*. Having had a common colonizer and having had lengthy diplomatic ties are both negatively correlated with *Cultural Distance* - as expected. We find weaker evidence that ancestral distance is correlated with institutional

[8] The effect of cultural distance is reduced in magnitude under both OLS and Poisson estimation, likely because the fully-saturated regression contains many measures of ancestral distance that are collinear with cultural distance.

distance, but the other variables are significant and bear the expected signs.

Results for the second stage appear in Table 5. As before, there are 6 columns, corresponding to three dependent variables (log FTI, log FDI and log FPI) and to whether we include additional controls or not. *Cultural Distance and Institutional Distance* are treated as endogenous. *Geographic Distance* and *Capital Controls* are treated as exogenous. *Ancestral Distance, Common Colonizer, Length of Diplomatic Ties* and *Religious Distance* are the instruments. Compared to the OLS results of Table 2, we find that the magnitude of the effect of the instrumented variables rises significantly. Take for instance the effect of cultural distance on log FTI (column 1). The effect of a one standard deviation increase in *Cultural Distance* was -45.4% under OLS, and it rises in magnitude to -71.65% under IV. A similar change is seen across specifications, and a similar change is seen for the effect of *Institutional Distance* as well. The rise in the effect of cultural and institutional distances comes at the expense of the effect of geographic distance - the latter has a reduced, though still sizable standardized effect of around 50%, under IV estimation. Estimates of the effect of *Capital Controls* do not differ greatly from those under OLS.

The bottom line from the IV results is that all three distance metrics continue to remain statistically and economically significant, with larger effects of cultural and institutional distances, and a smaller effect of geographic distance. These findings do not depend greatly on whether we control for additional determinants of foreign investment. Since the IV results feature usually larger magnitudes than OLS results, out of an abundance of caution we will rely on the latter for the counterfactual analysis conducted in Section 5.[9]

5 Counterfactual Analysis: Global Capital Misallocation

In this section, we calibrate the model of Section 2 using the econometric estimates of Section 4 to provide a quantitative assessment of the welfare impact of barriers to international investment. If we could hypothetically set to zero all intermediation costs associated with physical, cultural and institutional distances between countries, and let market forces reallocate capital, how would the sum and the cross-country distribution of output change? It is important to note that, in our counterfactual exercise, we are not assuming that all distances themselves would disappear. Societal differences in locations, values, beliefs, norms, policies and other traits, may continue to matter indirectly through their effects on productivities and other variables that differ across countries (empirically, such effects will continue to be captured by country fixed effects). Rather, our counterfactual assumption is that the world can now have access to new "intermediation technologies" - affecting factors such as transportation, communication, cultural translation, international cooperation, and so on - which would completely eliminate all intermediation

[9]Since our exclusion restrictions can be questioned, here we consider a fully-saturated regression specification where all of the exogenous variables (controls and instruments) are entered at once. The goal is to examine the robustness of the estimated β and β_ϕ coefficients. Results are presented below in Appendix Tables B.1 (OLS) and B.2 (Poisson). These are to be compared, respectively, to the results in Tables 2 and 3.

Table 5: Instrumental Variables Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	logFTI	logFDI	logFPI	logFTI	logFDI	logFPI
Cultural Distance	-2.429***	-2.180***	-2.022***	-2.342***	-1.825***	-1.689***
	(0.479)	(0.708)	(0.366)	(0.535)	(0.702)	(0.448)
Institutional Distance	-1.878***	-2.427**	-0.980*	-2.408***	-3.356***	-2.014***
	(0.691)	(1.213)	(0.556)	(0.722)	(1.245)	(0.653)
Geographic Distance	-1.094***	-1.156***	-0.984***	-1.680***	-1.877***	-1.216***
	(0.229)	(0.291)	(0.181)	(0.454)	(0.502)	(0.414)
Capital Controls	-0.045***	-0.013	-0.076***	-0.045***	-0.013	-0.080***
	(0.013)	(0.014)	(0.014)	(0.013)	(0.014)	(0.014)
Control Variables	No	No	No	Yes	Yes	Yes
Observations	19,246	21,468	19,167	18,989	20,987	18,848

Table Notes: the table above reports Instrumental Variable (IV) estimates of a linear regression of the variables listed on the topmost row (*Foreign Total Investment*, *Foreign Direct Investment*, *Foreign Portfolio Investment*) on the variables in the leftmost column. Instruments are: *Ancestral Distance, Common Colonizer, Length of Diplomatic Ties and Religious Distance.* Each observation is a directed country pair. All regressions include origin and destination country fixed effects as well as year fixed effects. Additional controls in columns 4-6 are *Colonial Relationship, Common Sea, Contiguity, Latitudinal Distance, Longitudinal Distance, Linguistic Distance, Same Continent* and *Trade Costs*. Standard errors (double-clustered by directed and undirected country pair) in parentheses. $^{*}p < .10$; $^{**}p < .05$; $^{***}p < .01$

costs associated with geographic, cultural and institutional differences. Thus, we will look at a world where geo-political and cultural differences between countries persist, but they no longer act as barriers to global capital allocation.

Two reasons motivate this exercise. The first is to provide a deeper understanding of the quantitative implications of the model. The second reason is related to policy. Of course, policy-makers cannot directly eliminate geographic distance. And, even if they could directly affect cultural distance, they probably should not for ethical, political, and economic reasons - including, for instance, the fact that geographic and cultural diversity across societies may have a positive impact on potential gains from trade, growth, and innovation. However, there can exist policies that reduce the *effect* of geographic, cultural and institutional distances on intermediation costs - for example, policies that facilitate travel, communication, and inter-cultural exchanges. The counterfactual exercise can be interpreted as capturing the potential benefits from such barrier-reducing policies.

5.1 Model Mapping and Calibration

Three country-level macroeconomic variables are required to take our model to the data. The first is output (y_i). We measure this as GDP in current PPP US dollars. This series is obtained from the International Monetary Fund's World Economic Outlook database (IMF-WEO). The second is labor input (ℓ_i). We measure this as total employment, which we obtain the Penn World Tables (PWT, version 9.1). From the Penn-World tables we also obtain a measure of capital-output ratio (k_i/y_i), which we combine with IMF GDP to obtain a measure of the capital stock at current prices, which we use for model validation purposes. The last data ingredient is the output-capital elasticity θ_i. We measure this as one minus the labor income share of GDP, for which we obtain country-level estimates from the International Labour Office (ILO) Department of Statistics. Finally, we calibrate $\alpha = 1/2$.[10]

Our regression estimates for β vary somewhat depending on the estimator. As noted earlier, IV estimates tend to be larger for *Cultural Distance* and *Institutional Distance*. On the basis of the range of specifications we estimated and – in order to be conservative – we calibrate the investment to-distance semi-elasticities (β) as follows: -.9 for *Cultural Distance,* -0.55 for *Institutional Distance*, -1.7 for *Geographic Distance,* -0.044 for *Capital Controls*.

The model components that remain to be identified are the matrix of portfolio shares $\mathbf{\Sigma}$, the vector of (destination country) capital stocks $\mathbf{k}$, the vector of savings $\mathbf{s}$ and total factor productivities ω_i. We show that these objects are identified given the previously-measured variables and parameters.

To begin with, notice that the matrix $\mathbf{\Sigma}$ is identified given $\mathbf{\Theta y}$ and $\mathbf{\Delta}$ (equation (2.31)). The vector of savings $\mathbf{s}$ is identified by equation (2.9). $\mathbf{k}$ is then obtained as $\mathbf{\Sigma s}$. Finally, the Cobb-Douglas production function pins down total factor productivity ω_i given k_i, ℓ_i and y_i.

5.2 Empirical Performance of the Gravity Equation

The first, most obvious question when it comes to model fit is well does the gravity equation actually fit the restated data on cross-border investment positions. After calibrating the model, we can use the gravity equation (2.21) to predict international investment by origin and destination-country.

The model-implied foreign investment positions are shown in Figure 2 against the actual data, We plot the log of Foreign Total Investment against its corresponding predicted value. It is important to note that these are *not* fitted values from regressions of Section 4, but model-implied values. The difference between the two lies in the fixed effects: while in the econometric model they are fitted, in the model they are computed as a function of model parameters and observables.

[10]The parametrization of α is inconsequential from an economic standpoint (it does not affect any of our counterfactuals): all it does is scaling the overall level of capital across all countries).

Figure 2: Empirical Performance of the Gravity Model

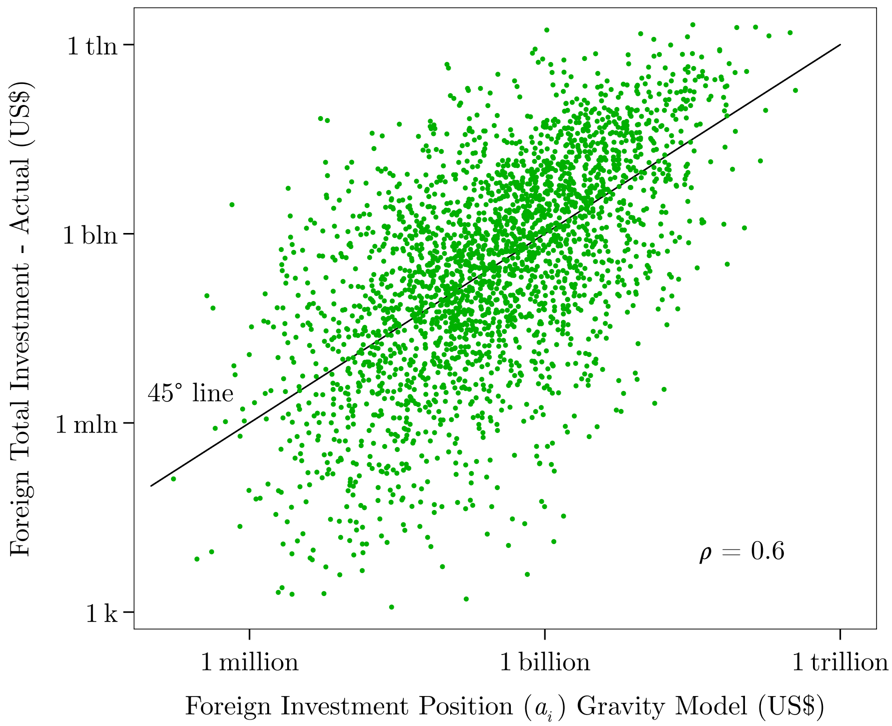

Figure Notes: the figure above plots the log of Foreign Total Investment against the log of predicted international investment from the gravity model (a_{ij}). The data is for 2017.

The gravity equation performs satisfactorily, in our view: the actual and fitted values line up neatly along the 45° line, and their correlation is 0.6, which is high considering that the equation above is not fitted using fixed effects. By comparison, the observed correlation between log FDI and log FPI (0.63). Hence, our first empirical finding is that our gravity equation predicts relatively well cross sectional variation in total investment almost as well as FPI predicts FDI (and vice-versa).

5.3 Model Fit: Country-Level Capital Stocks

We now consider another way to validate our model empirically. Notice that, in taking the model to the data, we haven't actually used any domestic capital supply data. In particular, we have not yet used capital stock data from the Penn World Tables. To further validate our model, we can then compare capital stock per employee from PWT to our model-based estimates. Once again, we take these two variables in logs, subtract the mean, and then plot the two series in a scatter plot in in Figure 3.[11]

[11]The OLG model does not differentiate between investment and capital stock, hence the level is not comparable. This is not important, since the model is isomorphic to scaling up or down all capital stocks (productivity ω_i adjusts to preserve the level of GDP).

Figure 3: Model Fit: Capital Stock per Employee

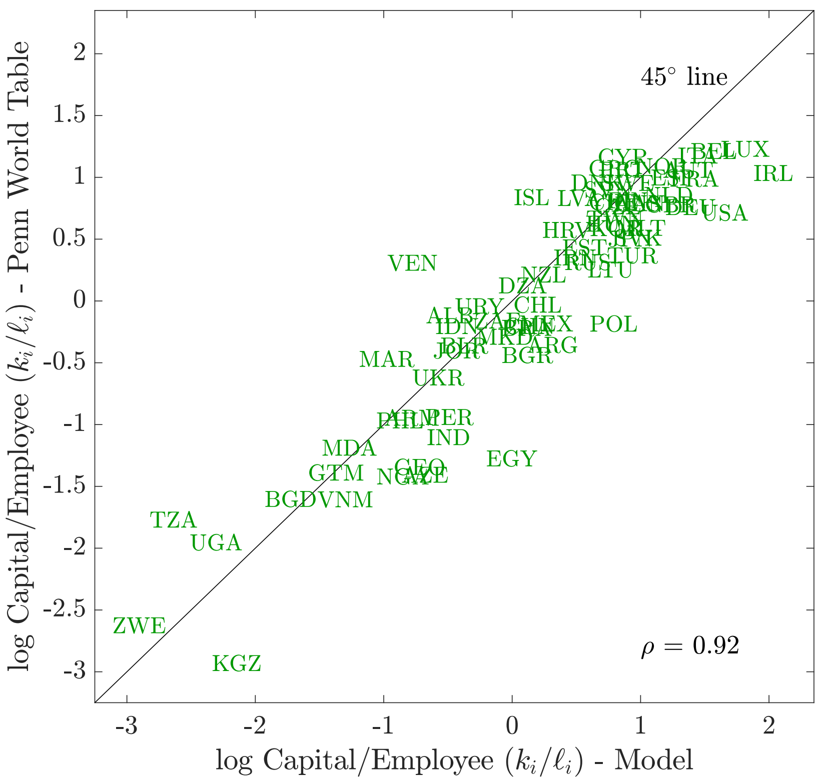

Figure Notes: the figure above plots the log of capital stock per employee (k_i/ℓ_i) from the against the corresponding model estimate. Both measures are in differences from the mean. Each observation is a country and data refers to the year 2017.

The graph shows that our model-based estimates correlate very strongly with their counterparts from Penn World Tables (the correlation is 0.9), and line up along the 45-degree line. This suggests that our model does a reasonable job capturing the cross-section of capital stocks across the 72 countries we are able to cover. Repeating the exercise without dividing by employment (ℓ_i) yields an even stronger correlation (0.97).

5.4 Rates of Return

An key endogenous variable in our model is the objective rate of return on capital (r_i). As we have seen in sub-section (2.4), the cross-country dispersion in rates of returns arises as a consequence of capital market imperfections (variation in $\mathbf{d}'\beta$ and ϕ), and can be related to the resulting GDP loss. As already shown in subsection 2.4, if we set $\beta = \mathbf{0}$ and $\beta_\phi = 0$, rates of return would be equalized and capital would be efficiently allocated across countries. It is therefore useful, in order to evaluate the model empirically, to investigate how the rates of returns produced by our model compare to other independently-computed

Figure 4: Model-Implied Rates of Return on Capital

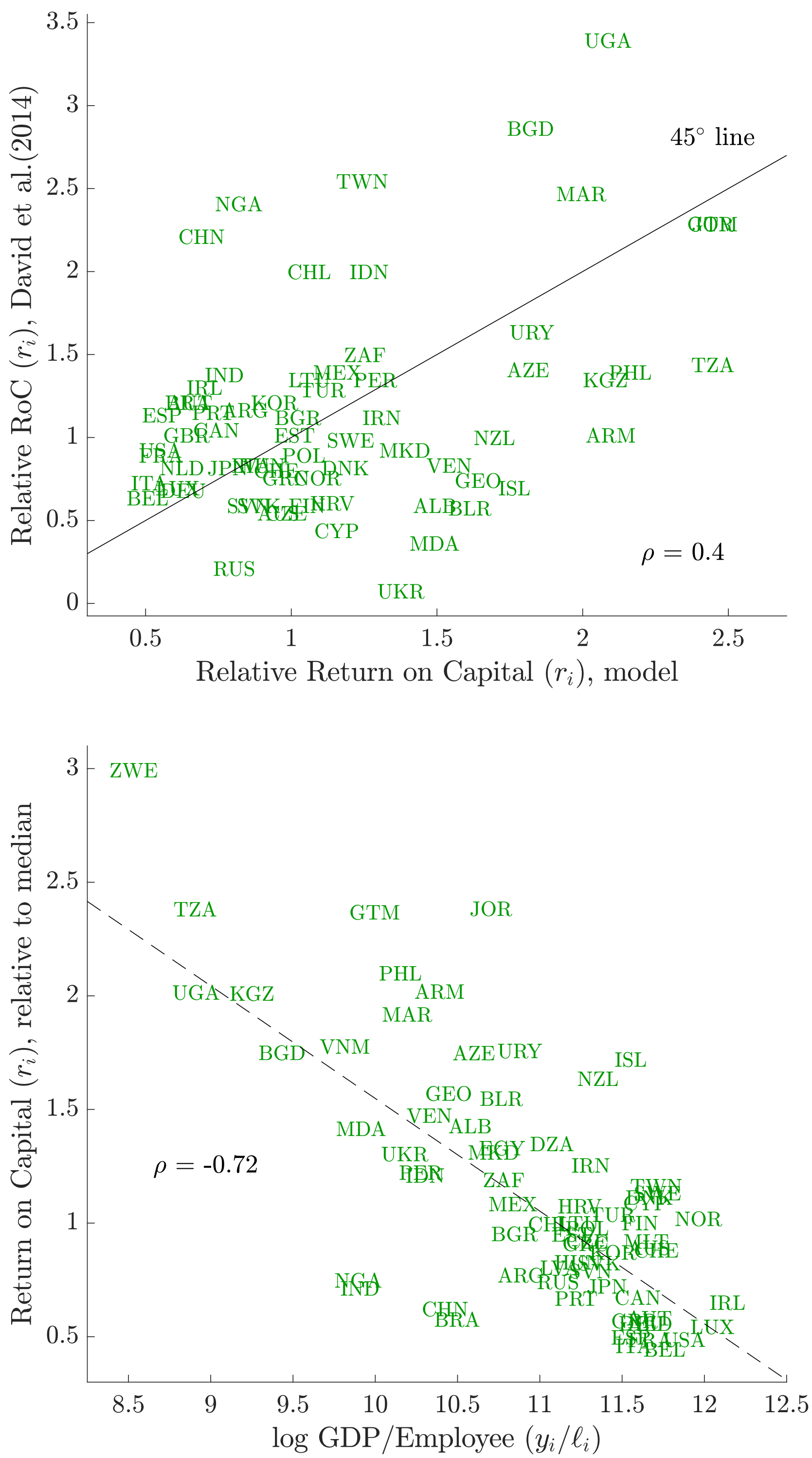

Figure Notes: the figure above plots the model-implied rate of return on capital, as a ratio to its median value, against the corresponding estimate from David, Henriksen, and Simonovska (2014, upper panel) and against the log of GDP per employee (lower panel). Each observation is a country and data refers to the year 2017.

estimates.

In a recent paper, David, Henriksen, and Simonovska (2014, henceforth DHS) produce estimates of the return on capital for 144 countries. Of these, 67 overlap with countries in our dataset. As a validation exercise, we compare the (country-level) relative rate of returns on capital from our model to DHS's estimates. By "relative", we mean that we normalize each of the two estimates by their respective median values (the levels are not comparable, because in our model there is no capital depreciation).

In Figure 4 (upper panel) we produce a scatter plot of the rates of returns estimated by DHS against those implied by our gravity model. We find that the two series are strongly correlated ($\rho = 0.4$). We argue that this correlation should be considered high because, again, the cross-country variation in rates returns generated by our model is driven entirely by our bilateral metrics of Capital Controls as well as Geographic, Cultural and Institutional Distance. If we were to set $\beta = 0$ and $\phi_{ij} = 1$ for all (i, j), the upper panel of Figure 4 would display a perfectly-vertical line. Hence, the observed correlation between our model-implied values and DHS's estimates arises exclusively from our four explanatory variables.

Another desirable property of the capital returns generated by our model is that they are consistent with a stylized fact, previously documented by DHS (and already widely acknowledged among finance professionals): rates of returns on capital correlate negatively, at the country level, with the level of economic development. That is, capital yields higher returns in emerging economies than in developed ones. We document this in the lower panel of Figure 4, where we plot the relationship between the rates of return from our model against the log of GDP per employee. The correlation between these two variables is -0.71, which is consistent with Lucas (1990)'s and Alfaro, Kalemli-Ozcan, and Volosovych (2008)'s hypothesis that too little capital is invested in emerging economies, compared to what would be allocated by an efficient international market for capital.

It is interesting to note that DHS also develop a model to explain this stylized fact. In their theoretical framework, capital yields higher returns in emerging economies due to risk and diversification (emerging assets are a worse hedge for global risk). In our framework, returns to capital are higher in emerging markets due to asset market frictions. It is not possible to judge the relative importance of these two factors based on our two models in isolation. A more general model – incorporating both risk and asset market frictions – would be needed. Also, a systematic methodology to measure asset return variances and covariances would likely be required. We see this as a promising avenue for future research.

5.5 Home Bias

Finally, we consider the out-of-sample performance of the model. As measured in the data, the matrix of investment portfolio shares $\mathbf{\Sigma}$ has a significant number of missing observations, which we can fill in using our gravity equation. Among the missing values are the entire diagonal of $\mathbf{\Sigma}$, which represents domestic investment.

Table 6: Matrix of G20 countries' Investment Shares (based on Gravity Model)

				Country of Origin - j																		
	Group	Country	Code	AUS	CAN	USA	IND	RUS	TUR	ZAF	ARG	BRA	MEX	DEU	FRA	GBR	ITA	CHN	IDN	JPN	KOR	RoW
Destination Country - i	1	Australia	AUS	71%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%
	1	Canada	CAN	1%	18%	3%	0%	0%	0%	1%	1%	1%	2%	0%	0%	2%	0%	0%	0%	0%	0%	1%
	1	United States	USA	8%	69%	92%	3%	3%	4%	11%	6%	3%	29%	3%	3%	16%	3%	1%	1%	1%	2%	7%
	2	India	IND	2%	0%	0%	67%	1%	2%	8%	0%	0%	0%	0%	0%	1%	1%	0%	4%	0%	1%	2%
	2	Russia	RUS	0%	0%	0%	1%	50%	5%	1%	0%	0%	1%	1%	2%	1%	2%	0%	1%	0%	1%	2%
	2	Turkey	TUR	0%	0%	0%	1%	3%	46%	1%	0%	0%	0%	0%	1%	1%	2%	0%	1%	0%	0%	2%
	2	South Africa	ZAF	0%	0%	0%	1%	0%	0%	39%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	3	Argentina	ARG	0%	0%	0%	0%	0%	0%	0%	44%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	3	Brazil	BRA	0%	0%	0%	0%	0%	0%	4%	23%	74%	3%	0%	1%	0%	1%	0%	0%	0%	0%	2%
	3	Mexico	MEX	0%	1%	1%	0%	1%	0%	1%	3%	4%	53%	0%	0%	0%	1%	0%	0%	0%	0%	1%
	4	Germany	DEU	0%	1%	0%	1%	4%	3%	2%	1%	0%	1%	49%	10%	6%	10%	1%	0%	1%	2%	11%
	4	France	FRA	0%	1%	0%	1%	5%	5%	2%	3%	2%	2%	7%	34%	4%	17%	0%	1%	0%	0%	9%
	4	United Kingdom	GBR	1%	4%	1%	2%	2%	2%	5%	1%	0%	1%	4%	4%	49%	3%	0%	0%	0%	0%	5%
	4	Italy	ITA	0%	1%	0%	1%	5%	8%	3%	3%	2%	1%	6%	15%	3%	28%	0%	1%	0%	0%	9%
	5	China	CHN	5%	1%	0%	9%	6%	3%	3%	0%	0%	1%	3%	1%	1%	1%	88%	9%	15%	35%	5%
	5	Indonesia	IDN	1%	0%	0%	1%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	68%	0%	0%	0%
	5	Japan	JPN	4%	1%	0%	2%	2%	1%	1%	0%	0%	1%	2%	0%	1%	1%	6%	4%	75%	22%	2%
	5	South Korea	KOR	1%	0%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	1%	4%	32%	1%
	6	Rest of the World	RoW	5%	3%	1%	7%	15%	18%	14%	13%	9%	5%	23%	29%	13%	30%	1%	7%	2%	3%	39%

Table Notes: The table above presents, for G20 countries (except Saudi Arabia, for which we do not have data) matrix of investment shares $\mathbf{\Sigma} = [\sigma_{ij}]$, based on the gravity model of Section 2. $\mathbf{\Sigma}$ lists, for every country pair (i, j), the share of capital owned by j that is invested in country i. G20 countries are officially sorted in 5 groups which are loosely based on the countries' location. We group countries in this matrix based on their official G20 group membership, and impute missing shares based on the gravity model we estimate in Section 4. All of the diagonal elements are imputed.

TABLE 7: COUNTERFACTUALS (2017)

	Observed (All Barriers)	*Zero-Gravity (No Barriers)*	*Cultural Distance*	*Institutional Distance*	*Geographic Distance*	*Capital Controls*
Welfare Statistics (Fixed Capital Stock)	(1)	(2)	(3)	(4)	(5)	(6)
World GDP (US$ trillions)	112.7	123.5	120.6	121.7	121.3	121.9
World GDP, % Difference in GDP from Zero-Gravity	-8.8%	0%	-2.4%	-1.5%	-1.8%	-1.3%
St.Dev. of $\log(k_i/\ell_i)$, % Difference from Zero-Gravity	+70.7%	0%	+20.7%	+13.9%	+39.0%	+11.7%
St.Dev. of $\log(y_i/\ell_i)$, % Difference from Zero-Gravity	+38.8%	0%	+11.6%	+6.9%	+18.1%	+5.3%
Welfare Statistics (Endogenous Capital Stock)						
World GDP (US$ trillions)	112.7	123.2	121.5	122.0	120.9	123.4
World GDP, % Difference in GDP from Zero-Gravity	-8.6%	0%	-1.4%	-1.0%	-1.9%	+0.2%
St.Dev. of $\log(k_i/\ell_i)$, % Difference from Zero-Gravity	+70.8%	0%	+17.8%	+13.5%	+49.2%	+13.1%
St.Dev. of $\log(y_i/\ell_i)$, % Difference from Zero-Gravity	+38.8%	0%	+10.1%	+7.3%	+21.9%	+6.1%

TABLE NOTES: The table above presents welfare statistics for six counterfactuals of the model described in Section 2. Each of the columns on the right side of the graph is a counterfactual, and the rows represent different welfare statistics of interest. *Observed* is the equilibrium allocation with all measured barriers (Geography, Institutions, Culture). *Zero-Gravity* is the counterfactual in which all three barriers have been removed. *Cultural Distance*, *Institutional Distance, Geographic Distance* and *Capital Controls* are four counterfactuals from which only the corresponding distortion is in place. (k_i/ℓ_i) is the capital stock per employee, while (y_i/ℓ_i) is output (GDP) per employee. Actual World PPP$ GDP (including countries not in the model) in 2017 was $121 trillion.

Based on the literature on home bias, we can predict that these diagonal elements should be large relative to non-diagonal elements. Hence, one of the strictest tests of our gravity model would be its ability to predict, out-of-sample, a significant home bias.

Table 6 displays the matrix of portfolio shares $\mathbf{\Sigma}$, based on restated IMF data, with missing values imputed using our gravity model. Because there are 72 countries in our sample, we present G20 plus the rest of the world. Two features of this matrix are striking. The first is the dominance of the United States as a receiving country, which is consistent with its role as a "financial supermarket" which was previously documented by Martin and Rey (2004). The second striking pattern is the prominence of the diagonal elements, whose average value is around 36%, and which delineate a very significant home bias.

Given that our gravity equation loads negatively on measures of cultural, institutional and geographic distance, the fact that it predicts at least some degree of home bias is not entirely surprising. What is unexpected is how large is the size of the home bias predicted by our model: hence, we argue that another strength of our structural gravity model is its ability to predict, out-of-sample, a significant domestic bias.

The magnitudes of the home bias we observe are compatible with the home bias documented by French and Poterba (1991) and (more recently) Coeurdacier and Rey (2013), albeit it should be noted that our measurements are not very comparable: we focus on a wider range of assets and our data is more recent[12]; moreover, our data comes a completely different source.

5.6 Counterfactual Analysis

We now move to the core of our contribution: we use our model to perform counterfactual analysis. Counterfactuals allow us to gain an sense of the economic impact of the barriers. In our setting, simulating a counterfactuals means changing the vector β (the semi-elasticity of investment with respect to the distance vector) and/or the matrix of capital controls (Φ), and studying how all other variables (as well as statistics of interest) change in equilibrium in response to that.

For each of the counterfactuals, we compute the corresponding World GDP. We also compute the percentage difference between the counterfactual and an undistorted (*zero-gravity*) equilibrium in terms of three statistics: World GDP, the standard deviation of the log of capital per employee and the standard deviation of log of output per employee.

Table 7 presents our counterfactual analysis. In Column (1) we present the observed, distorted equilibrium, with capital controls, geographic, institutional and cultural distance. To estimate their effect, we use the OLS β estimates. In the column (2), we present the *zero-gravity* equilibrium, from which all distortions – except those at the level of the individual investors – have been removed ($\beta = \mathbf{0},\ \Phi = \mathbf{0}$). In column (3) to (6), we consider equilibria in which the effect of *Cultural Distance*, *Institutional Distance*, *Geographic Distance* and *Capital Controls* are respectively re-introduced, in isolation (that is, all other

[12]Coeurdacier and Rey (2013) document that home bias is slowly weakening.

distortions except that indicated are removed). These four latter counterfactuals allow to gain a sense of the marginal impact of each individual distortion.

The upper panel shows results of our counterfactual, computed under the assumption that the supply of capital is inelastic (that is, the vector of savings $\mathbf{s}$ cannot react to the removal of the distortions). This set of results are meant to isolate the effect of capital misallocation, and reflect a "one-period" change in our model economy following the addition or removal of distortions. The lower panel shows results based on the assumption that the supply of capital can adjust in response to the removal of the distortions. These set of results reflect instead how the steady state equilibrium changes: they account for the fact that, when capital is reallocated away from some countries and towards some other countries, the total capital stock may need to adjust in order for steady-state conditions to be respected.

We find that barriers to the global allocation of capital have quantitatively important effects on the level of output produced globally. World GDP in the observed equilibrium of our model is measured at 112.7 US$ billion. That is 8.8% lower than in the *zero-gravity* counterfactual (column 2) under the assumption of an inelastic capital supply, and 8.6% lower when we assume that the supply of capital can respond.

We find that the four distortions (with the exception of *Capital Controls* under elastic capital supply) have quantitatively comparable effects on World GDP when considered in isolation. When all distortions except those due to *Cultural Distance* are removed, GDP is 2.4% lower than in the Zero-Gravity scenario; this figure becomes 1.4% if capital is allowed to respond. For *Institutional Distance*, the corresponding GDP losses are 1.5% (inelastic capital) and 1% (elastic capital). When *Geographic Distance* alone is considered, the GDP losses are 1.8% (inelastic capital) and 1.9% (elastic capital). *Capital Controls* have a more muted impact when considered in isolation: they reduce World GDP by 1.3% when capital supply is inelastic. They appear to have virtually no effect on world GDP when capital supply is allowed to adjust.

One common theme of these counterfactual is that GDP losses appear slightly smaller when we allow the capital stock to react. A potential explanation for this finding is that we allow countries to vary in their capital income shares (θ_i). When distortions are removed, capital generally tends to be reallocated to countries with higher capital shares, and country savings (s_i) are smaller for countries with a larger capital share. Hence, with an elastic capital supply, aggregate saving tends to be slightly lower in the zero-gravity counterfactual due to compositional effects.

While the overall effects of these four distortions on allocative efficiency and world GDP appears substantial, their effect on cross-country inequality appears even more sizable. We can gain a sense of country heterogeneity by looking at how much these distortions change the distribution of capital and output per employee.

When capital misallocation resulting from barriers to international investment are removed, we observe a significant decrease in dispersion of both capital and output per employee. When moving from the zero-gravity equilibrium to the observed (distorted) equilibrium, the standard deviation of (log) capital per

Figure 5: Distribution of Capital and Output per Employee

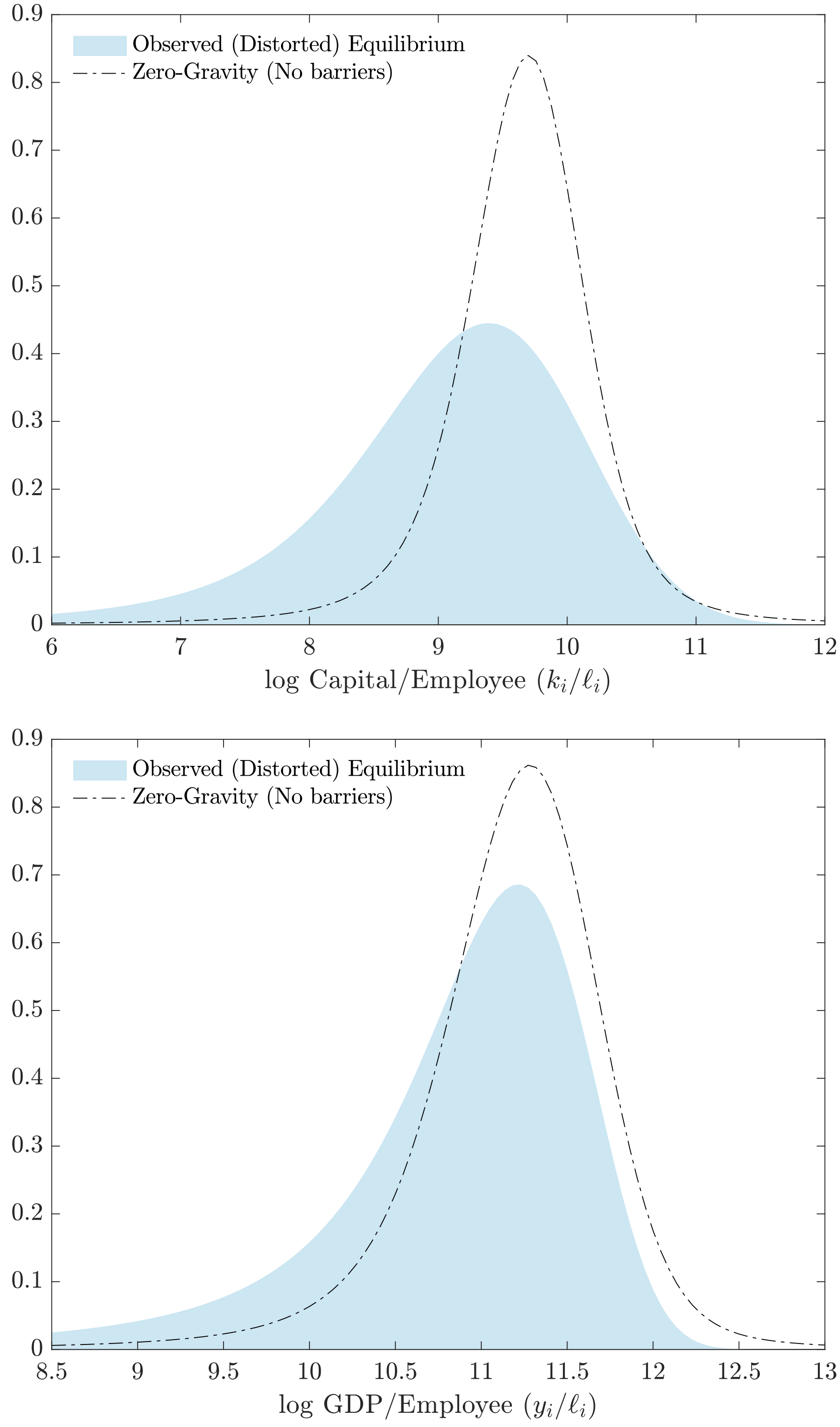

Figure Notes: the figure above fits the probability density function of a *stable* distribution (a 4-parameter family of distributions with flexible skewness and fat tails) to country-level capital stock per employee (upper panel) and GDP per employee (bottom panel). In each panel, the lighter area is the distribution in the observed, distorted equilibrium. The dotted black line is the distribution in a counterfactual scenario in which all measured distortions to capital movement (Geographic, Cultural and Institutional distance) have been removed.

employee increases by 70.7%%, while the standard deviation of log output per employee increases by 38.8%. These numbers are not sensitive to whether capital supply is assumed to be inelastic.

When *Cultural Distance* alone is considered, dispersion increases by 17.8-20.7% (depending on whether capital supply is elastic or not), with respect to the zero-gravity counterfactual, for the log of capital per employee; it increases by 10.1-11.6% for log output per employee. When *Institutional Distance* alone is re-introduced, dispersion increases by 13.5-13.9% for log capital per employee, and by 6.9-7.3% for log output per employee. When *Geographic Distance* alone is considered, dispersion increases by 39-49.2% for log capital per employee, and by 18.1-21.9% for log output per employee. Finally, when *Capital Controls* are considered in isolation, dispersion increases by 11.7-13.1% for log capital per employee, and by 5.3-6.1% for log output per employee.

Figure 5 provides a graphical representation of the effect of removing barriers to international investment on cross-country inequality. It shows how the (fitted) cross-country distribution of capital and output per employee changes in response to the removal of the barriers. For both variables, we observe a significant reduction in dispersion, but also in skewness (the left tail becomes thinner). We also can notice a general rightward shift, reflecting an increase of capital and income per employee for the median country.

The reason why we observe this reduction in inequality is that, when capital distortions are removed, capital tends to be reallocated to countries that had higher rates of returns on capital in the distorted equilibrium. As we saw previously, these tend to be countries with lower capital stock per employee and lower output per employee.

Figure 6 illustrates this effect: it is a scatter plot of the baseline level of GDP per employee (horizontal axis) against the log change in GDP per employee from moving to a zero-gravity world (vertical axis). The latter number can also be read, on the right axis scale, as the log change in capital per employee. As can be seen from the graph, there are significant "winners" and "losers" among the countries in our dataset – albeit on average most countries experience an increase in capital and output per capita. The strong negative correlation between the country-level gains and the initial level of output per employee imply that the removal of investment barriers from our gravity model leads to a substantial reduction in cross-country inequality. In other words, poorer countries (that have a lower capital/labor ratio in the baseline scenario) benefit disproportionately from capital reallocation. Some of them, such as Zimbabwe or Uganda, see capital per employee more than double, and their income per employee jump by over half.

Finally, it follows from our results in subsection 2.4 that, when we move to *Zero-Gravity*, the home bias disappears; all countries hold exactly the same portfolios.

We should clarify that, while our empirical approach allows us to consider the counterfactual effects of removing capital controls that are specific to country i and country j, we do not consider the removal of other potential obstacles to capital movements that may indirectly depend on country-specific factors such as institutional quality and the regulatory environment (see Alfaro, Kalemli-Ozcan, and Volosovych, 2008),

Figure 6: Capital Reallocation Gains by Country

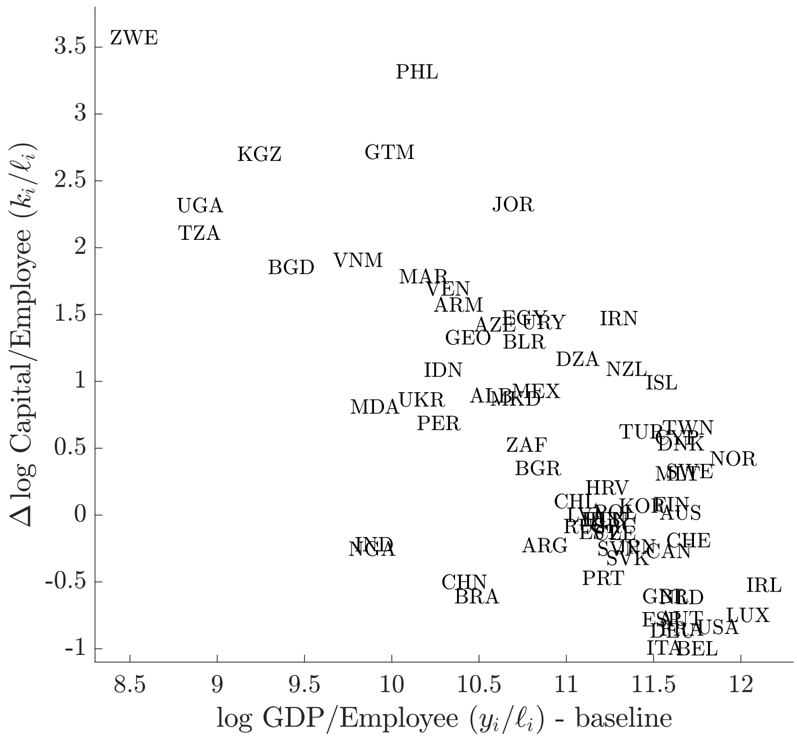

Figure Notes: the figure above displays the baseline level of log output per capita (horizontal axis), against the log change in capital stock per employee (left axis) following the removal of all measured barriers to global capital allocation. Each observation is a country, and the data is from 2017.

because those variables are captured by country fixed effects in our regressions. Finally, another factor that our counterfactual analysis cannot capture is how *within*-country capital (mis)allocation might be affected by financial integration (Gopinath, Kalemli-Özcan, Karabarbounis, and Villegas-Sanchez, 2017).

In summary, using counterfactual analysis, we find that misallocation of capital across countries – induced by capital controls as well as geographic, institutional and cultural distance – imposes quantitatively important output losses for the majority of countries, and in general for World GDP, and can potentially account for a significant share of the observed cross-country dispersion in capital/employee.

6 Conclusions

A large literature on open economy financial macroeconomics has flourished in recent years. A centerpiece of this line of research has been to study home bias in asset markets (Coeurdacier and Rey, 2013) and the special status of the US dollar among currencies (Kekre and Lenel, 2020; Gourinchas, Rey, and Sauzet, 2019), using two-country and Small Open Economy (SOE) models with asset markets. In this paper we have addressed a new question in this literature: we have studied the efficiency of capital allocation across countries.

To accomplish this, we have developed a novel multi-country overlapping-generations general equilibrium model of international capital allocation that yields a gravity equation for foreign assets demand as an equilibrium outcome. The model features atomistic agents with objective and subjective returns to investing in various markets. The returns depend crucially upon factors that affect intermediation costs, reducing the return from investing in distant markets. In turns, we interpret these costs as the result of capital account policy restrictions, as well as geographic, cultural and institutional barriers to global capital allocation.

We have estimated our gravity model empirically, using foreign direct and portfolio investment data that has been restated (Coppola et al., 2020; Damgaard et al., 2019) from a residency to a nationality basis, in order to account for the presence of offshore investment and financing vehicles. Using a variety of estimation approached (OLS, Poisson, IV), we have found that capital account policy restrictions. cultural, institutional and geographic barriers have substantial effects on the allocation of capital across different societies. The estimated effects are large in magnitude, suggesting that the removal of barriers to international capital allocation could have important effects on output, welfare and inequality across countries. Our parsimonious implementation of the model, based on these four explanatory variables, explains a significant share of the observed variation in FDI and FPI.

Our model reproduces several features of international capital international asset markets. First, it produces large and meaningful variation in rates of returns across countries; in particular, one feature that makes our model realistic is that it produces rates of returns that correlate negatively with the level of economic development. Second, our model produces, out-of-sample, a large home bias in a multi-country setting. While previous research has emphasized diversification and hedging as crucial to understanding these patterns, our analysis suggests that geopolitical factors and capital controls are also likely to play an important role in determining these patterns.

To quantify the influence of these factors on the international allocation of capital and their real impact, we have performed a number of counterfactual exercises using our model. Most importantly, we have studied how World GDP and the cross-country distribution of capital and output per worker would change if the effects of geographic, cultural and institutional distances on foreign investment were zeroed out.

Our quantitative exercise suggests that capital misallocation associated with our four barrier variables has a major impact on the distribution capital across countries, in terms of efficiency as well as inequality. World GDP is 8.6-8.8% lower than it would be if the effect of these barriers could be neutralized. The cross-country standard deviation of capital per employee is 71% higher, while the dispersion of output per employee is nearly 40% higher. Conversely, the hypothetical removal of these frictions would lead to substantial economic gains and reductions in cross-country inequality, by reallocating capital from richer countries (where the rate of return on capital is lower) to poorer countries (where the rate of return is higher). It is important to note that, in our estimation, some countries "lose out" (output and capital per employee drop), following the removal of international capital distortions.

Our study contributes to the literature on open economy financial macroeconomics, by making theoretical as well as empirical progress in modeling international asset markets in a multi-country, general-equilibrium setting. It also connects to the macroeconomics literature on resource misallocation, by studying the real effects of international asset market frictions.

We conclude by suggesting directions for future research. Our work departs in a significant manner from the existing small open economy macro-financial literature in that it focuses on geo-political and policy frictions, as opposed to diversification and hedging, as potential drivers of international capital allocation. Hence, a potential goal of future research could be to produce a unified theory of international investment, which incorporates asset market imperfections as well as currency and risk considerations in a multi-country environment.

A second, interesting potential direction for future research is to incorporate, in our model, a richer specification of the goods sector, which we deliberately chose to keep simple. While recent work (Arkolakis et al., 2018; Ramondo and Rodríguez-Clare, 2013; Tintelnot, 2016) has made theoretical progress in combining trade and global value chains (which can be viewed as a narrower aspect of foreign direct investment), a fully-integrated model for international trade and investment would allow to study the combined effect of asset market imperfections and goods markets distortions, accounting for possible ways in which the two interact with each other.

Finally, another direction for future research that we suggest is to enrich the model along the dimension of consumer/investor heterogeneity. In our model, all agents have the same saving rate: this limits our ability to study the implications of financial globalization for *within*-country inequality (we can only study cross-country inequality). An interesting question that could be investigated with a richer household sector is therefore how financial globalization might affect *within*-country income inequality.

In his seminal 1990 paper, Robert E. Lucas asked: "Why doesn't capital flow from rich to poor countries?". In this paper, we have provided new insights on this question: we have found that the complex network of cross-country investment positions that we see in the data is at least partly shaped by geo-political barriers, and this has major efficiency and distributional real effects, including hindering the flow of capital to least-developed countries.

References

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003): "Fractionalization," *Journal of Economic growth*, 8, 155–194.

Alfaro, L., S. Kalemli-Ozcan, and V. Volosovych (2008): "Why doesn't capital flow from rich to poor countries? An empirical investigation," *The review of economics and statistics*, 90, 347–368.

Anderson, J. E. and E. Van Wincoop (2003): "Gravity with gravitas: A solution to the border puzzle," *American economic review*, 93, 170–192.

Arkolakis, C., N. Ramondo, A. Rodríguez-Clare, and S. Yeaple (2018): "Innovation and production in the global economy," *American Economic Review*, 108, 2128–73.

Baqaee, D. R. and E. Farhi (2020): "Productivity and misallocation in general equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.

Blonigen, B. A. and J. Piger (2014): "Determinants of foreign direct investment," *Canadian Journal of Economics/Revue canadienne d'économique*, 47, 775–812.

Bove, V. and G. Gokmen (2018): "Genetic distance, trade, and the diffusion of development," *Journal of Applied Econometrics*, 33, 617–623.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2019): "Migrants, ancestors, and foreign investments," *The Review of Economic Studies*, 86, 1448–1486.

Chinn, M. D. and H. Ito (2006): "What matters for financial development? Capital controls, institutions, and interactions," *Journal of development economics*, 81, 163–192.

Coeurdacier, N. and H. Rey (2013): "Home bias in open economy financial macroeconomics," *Journal of Economic Literature*, 51, 63–115.

Coppola, A., M. Maggiori, B. Neiman, and J. Schreger (2020): "Redrawing the map of global capital flows: The role of cross-border financing and tax havens," *Available at SSRN 3525169*.

Correia, S., P. Guimaraes, and T. Zylkin (2019): "PPMLHDFE: Stata module for Poisson pseudo-likelihood regression with multiple levels of fixed effects," .

Damgaard, J., T. Elkjaer, and N. Johannesen (2019): "What Is Real and What Is Not in the Global FDI Network?" *International Monetary Fund Working Papers*.

David, J. M., E. Henriksen, and I. Simonovska (2014): "The risky capital of emerging markets," Tech. rep., National Bureau of Economic Research.

David, J. M. and V. Venkateswaran (2019): "The Sources of Capital Misallocation," *American Economic Review*, 109, 2531–2567.

De Ménil, G. (1999): "Real capital market integration in the EU: How far has it gone? What will the effect of the euro be?" *Economic Policy*, 14, 166–201.

Di Giovanni, J. (2005): "What drives capital flows? The case of cross-border M&A activity and financial deepening," *Journal of International Economics*, 1, 127–149.

Eaton, J. and S. Kortum (2002): "Technology, geography, and trade," *Econometrica*, 70, 1741–1779.

Eichengreen, B. and P. Luengnaruemitchai (2008): "Bond markets as conduits for capital flows: how does Asia compare?" in *International Financial Issues in the Pacific Rim: Global Imbalances, Financial Liberalization, and Exchange Rate Policy (NBER-EASE Volume 17)*, University of Chicago Press, 267–313.

Fearon, J. (2003): "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth*, 8, 195–222.

Felbermayr, G. J. and F. Toubal (2010): "Cultural proximity and trade," *European Economic Review*, 54, 279–293.

Fensore, I., S. Legge, and L. Schmid (2017): "Human barriers to international trade," *University of St. Gallen Discussion Paper.*

Fernández, A., M. W. Klein, A. Rebucci, M. Schindler, and M. Uribe (2016): "Capital control measures: A new dataset," *IMF Economic Review*, 64, 548–574.

French, K. R. and J. M. Poterba (1991): "Investor Diversification and International Equity Markets," *The American Economic Review*, 81, 222–226.

Ghosh, S. and H. Wolf (1999): "The geography of capital flows," *Capital Inflows to Emerging Markets, University of Chicago Press, Chicago.*

Gopinath, G., Ş. Kalemli-Özcan, L. Karabarbounis, and C. Villegas-Sanchez (2017): "Capital allocation and productivity in South Europe," *The Quarterly Journal of Economics*, 132, 1915–1967.

Gourinchas, P.-O., H. Rey, and M. Sauzet (2019): "The international monetary and financial system," *Annual Review of Economics*, 11, 859–893.

Guiso, L., P. Sapienza, and L. Zingales (2009): "Cultural biases in economic exchange," *Quarterly Journal of Economics*, 124, 1095–1131.

Hall, R. E. and C. I. Jones (1999): "Why do some countries produce so much more output per worker than others?" *The quarterly journal of economics*, 114, 83–116.

Head, K. and J. Ries (2008): "FDI as an outcome of the market for corporate control: Theory and evidence," *Journal of International Economics*, 74, 2–20.

Helpman, E., M. Melitz, and Y. Rubinstein (2008): "Estimating trade flows: Trading partners and trading volumes," *The quarterly journal of economics*, 123, 441–487.

Hopenhayn, H. A. (2014): "Firms, misallocation, and aggregate productivity: A review," *Annu. Rev. Econ.*, 6, 735–770.

Kekre, R. and M. Lenel (2020): "The Dollar and the Global Price of Risk," *University of Chicago Booth School of Business, working paper.*

La Porta, R., F. Lopez-de Silanes, and A. Shleifer (2008): "The Economic Consequences of Legal Origins," *Journal of Economic Literature*, 285–332.

Lane, P. R. and G. M. Milesi-Ferretti (2008): "International investment patterns," *The Review of Economics and Statistics*, 90, 538–549.

——— (2018): "The external wealth of nations revisited: international financial integration in the aftermath of the global financial crisis," *IMF Economic Review*, 66, 189–222.

Leblang, D. (2010): "Familiarity breeds investment: Diaspora networks and international investment," *American political science review*, 104, 584–600.

LUCAS, R. E. (1990): "Why doesn't capital flow from rich to poor countries?" *The American Economic Review*, 80, 92–96.

MAGGIORI, M., B. NEIMAN, AND J. SCHREGER (2020): "International Currencies and Capital Allocation," *Journal of Political Economy (forthcoming).*

MARTIN, P. AND H. REY (2004): "Financial super-markets: size matters for asset trade," *Journal of international Economics*, 64, 335–361.

MAYER, T. AND S. ZIGNAGO (2011): "Notes on CEPII's distances measures: The GeoDist database," .

MCFADDEN, D. (1973): "Conditional logit analysis of qualitative choice behavior," *Frontiers in Econometrics*, 105–142.

MCGRATTAN, E. R. AND J. A. SCHMITZ JR (1999): "Explaining cross-country income differences," *Handbook of macroeconomics*, 1, 669–737.

MECHAM, R. Q., J. FEARON, AND D. LAITIN (2006): "Religious classification and data on shares of major world religions," *unpublished, Stanford University*, 1, 18.

NOVY, D. (2013): "Gravity redux: measuring international trade costs with panel data," *Economic inquiry*, 51, 101–121.

PEMBERTON, T. J., M. DEGIORGIO, AND N. A. ROSENBERG (2013): "Population Structure in a Comprehensive Genomic Data Set on Human Microsatellite Variation," *G3: Genes, Genomes, Genetics*, 3, 891–907.

PETKOVA, K., A. STASIO, AND M. ZAGLER (2019): "On the relevance of double tax treaties," *International Tax and Public Finance*, 1–31.

PORTES, R. AND H. REY (2005): "The determinants of cross-border equity flows," *Journal of international Economics*, 65, 269–296.

RAMONDO, N. AND A. RODRÍGUEZ-CLARE (2013): "Trade, Multinational Production, and the Gains from Openness," *Journal of Political Economy*, 121, 273–322.

ROSE, A. K. AND M. M. SPIEGEL (2009): "Noneconomic engagement and international exchange: The case of environmental treaties," *Journal of Money, Credit and Banking*, 41, 337–363.

SANTOS SILVA, J. M. AND S. TENREYRO (2006): "The log of gravity," *The Review of Economics and statistics*, 88, 641–658.

SPOLAORE, E. AND R. WACZIARG (2009): "The Diffusion of Development," *The Quarterly journal of economics*, 124, 469–529.

——— (2012): *Long-Term Barriers to the International Diffusion of Innovations*, University of Chicago Press, chap. Chapter 1, 11–46.

——— (2016): "Ancestry, Language and Culture," in *The Palgrave handbook of economics and language*, Springer, 174–211.

——— (2018): "Ancestry and Development: New Evidence," *Journal of Applied Econometrics*, 33, 748–762.

TINTELNOT, F. (2016): "Global Production with Export Platforms*," *The Quarterly Journal of Economics*, 132, 157–209.

Online Appendix

Barriers to Global Capital Allocation

A A Model of Ancestral and Cultural Distance

As documented in our empirical analysis, we observe a strong positive relation between *Ancestral Distance* and *Cultural Distance*, In this Appendix, we present a simple analytical framework that formally illustrates how such relation would naturally emerge in a setting where 1) ancestral distance measures the time since two populations have been separated (that is, they are no longer the same population), and 2) cultural change takes place as the result of random shocks. The framework builds on Spolaore and Wacziarg (2009, 2012) and Becker, Enke and Falk (2020).

Assume that at time 0 there exists only one population, with cultural traits denoted by a real number C^0. At time 1, the population splits into $P^1 > 1$ populations. Each new population $i = 1, 2, ..., P^1$ inherits the cultural traits of its ancestral population plus a shock ϵ_i^1 , so that

$$C_i^1 = C^0 + \epsilon_i^1 \tag{A.1}$$

We assume that the shocks are non-degenerate integrable random variables, independent and identically distributed across time and space. This assumption is a useful benchmark simplification, consistent with the view of cultural change as mainly due to random drift. This assumption is sufficient to obtain our main result (for discussions of more general assumptions, see Spolaore and Wacziarg, 2009 and Becker, Enke and Falk, 2020).

At time 2, populations split into $P^2 > P^1$ populations. Again, each new population $i = 1, 2, ..., P^2$ inherits the culture of its parental population the previous period plus a random shock ϵ_i^2. Let $a_{mt}(i)$ denote the ancestral population, living at time m, from which a population $i$living at time t descended., where $a_{tt}(i) = i$. Therefore, at time 2 we have:

$$C_i^2 = C^0 + \epsilon^1_{a_{12}(i)} + \epsilon_i^2 \tag{A.2}$$

In general, at time t, the cultural traits of each population $i = 1, 2, ..., P^t$ are equal to the sum of all previous shocks experienced by population i's ancestral populations plus the new shock:

$$C_i^t = C^0 + \sum_{m=1}^{t} \epsilon^m_{a_{mt}(i)} \tag{A.3}$$

Let $d^t_C(i, j) \equiv |C_i^t - C_j^t|$ denote the cultural distance between population $i$and population j at time t. Let $d^t_A(i, j)|$ denote the ancestral distance between between population $i$and population j, which is defined as the number of periods $N(i,j)$ in which population $i$and population j have different ancestors - that is, the number of periods when $a_m(i) \neq a_m(j)$. Assuming that the shocks are independent and identically distributed random variables, it follows that two populations at a larger ancestral distance from each other can also be expected to be at a higher cultural distance from each other. That is

$$d_A^t(i,j) > d_A^t(k,l) \Longleftrightarrow E[d_C^t(i,j)] > E[d_C^t(k,l)] \tag{A.4}$$

This relation is formally analogous to Proposition 1 in Becker, Enke and Falk (2020, online appendix, Section 3), and a formal derivation can be obtained along similar lines. Specifically, let n_{ij} be the number of periods up to time T when population $i$and population j were separated, and n_{kl}be the number of periods up to time T when population $k$and population l. Then, we can re-write the above preposition as:

$$n_{ij} > n_{kl} \Longleftrightarrow E[|C_i^T - C_j^T|] > E[C_k^T - C_l^T|] \tag{A.5}$$

By definition:

$$C_i^T - C_j^T = \sum_{a_{mT}(i) \neq a_{mT}(j)}^{T} (\epsilon_{a_{mt}(i)}^m - \epsilon_{a_{mt}(i)}^m) \tag{A.6}$$

where the sum of shocks is taken for all periods $m = 1, 2, ...T$ where the two populations do not share a common ancestral population. By defining $\eta_1, ...\eta_T, \nu_1, ...\nu_T$ as i.i.d. random variables having the same distribution as the shocks above, implying:

$$E[|C_i^T - C_j^T|] = E[|\sum_{q=1}^{n_{ij}} (\eta_q - \nu_q)|] \tag{A.7}$$

By the same token, we have:

$$E[|C_k^T - C_l^T|] = E[|\sum_{q=1}^{n_{kl}} (\eta_q - \nu_q)|] \tag{A.8}$$

Thus, our claim follows if we can show that:

$$n_{ij} > n_{kl} \Longleftrightarrow E[|\sum_{q=1}^{n_{ij}} (\eta_q - \nu_q)|] > E[|\sum_{q=1}^{n_{kl}} (\eta_q - \nu_q)|] \tag{A.9}$$

The right-hand side of this equivalence is formally identical to equation (1) in Becker, Enke and Falk (2020), online appendix, Section 3, page 10, and can be derived in the same way, using their Lemma 1 (the details are available upon request).

B Fully-Saturated Regression Specification

Our empirical analysis relies on a relatively parsimonious specification where a set of three distance metrics, a measure of capital movement policy restrictions and a small set of controls are entered in the specification. We also consider a set of deep determinants of the barriers to foreign investment that are excluded from the main specification. Since these exclusion restrictions can be questioned, here we consider a fully-saturated regression specification where all of the exogenous variables (controls and instruments) are entered at once. The goal is to examine the robustness of the estimated β and β_ϕ coefficients. Results are presented below in Appendix Tables B.1 (OLS) and B.2 (Poisson). These are to be compared, respectively, to the results in Tables 2 and 3.

Table B.1: Fully-Saturated Regression Specification

	(1)	(2)	(3)
	logFTI	logFDI	logFPI
Cultural Distance	-0.556***	-0.319**	-0.859***
	(0.133)	(0.145)	(0.128)
Institutional Distance	-0.493***	-0.373***	-0.360***
	(0.065)	(0.070)	(0.062)
Geographic Distance	-1.615***	-1.898***	-1.297***
	(0.292)	(0.296)	(0.278)
Capital Controls	-0.046***	-0.015	-0.079***
	(0.013)	(0.013)	(0.014)
Ancestral Distance	-21.445***	-26.567***	-9.422
	(7.249)	(6.751)	(6.426)
Common Colonizer	1.755***	1.710***	1.460***
	(0.395)	(0.323)	(0.409)
Length of Diplomatic Tie	0.006***	0.007***	0.005***
	(0.001)	(0.001)	(0.001)
Religious Distance	-1.116***	-1.198***	-0.269
	(0.418)	(0.420)	(0.403)
Colonial Relationship	1.124***	1.432***	0.837***
	(0.244)	(0.236)	(0.214)
Common Sea	-0.015	0.073	0.094
	(0.150)	(0.167)	(0.144)
Contiguity	0.803***	1.068***	0.646***
	(0.205)	(0.225)	(0.188)
Latitudinal Distance	0.009**	0.010**	0.015***
	(0.004)	(0.005)	(0.004)
Longitudinal Distance	0.006**	0.007**	0.008***
	(0.003)	(0.003)	(0.002)
Same Continent	0.426**	0.158	0.884***
	(0.181)	(0.179)	(0.189)
Linguistic Distance	-1.447***	-1.776***	-0.631
	(0.434)	(0.404)	(0.428)
Trade Costs	-7.250***	-4.821***	-8.233***
	(1.841)	(1.688)	(1.790)
Observations	18,989	20,987	18,848
R-squared	0.781	0.732	0.795

Table B.2: Fully-Saturated Poisson Regression Specification

	(1)	(2)	(3)
	FTI	FDI	FPI
Cultural Distance	-0.457***	-0.337**	-0.611***
	(0.126)	(0.168)	(0.127)
Institutional Distance	-0.195***	-0.169**	-0.274***
	(0.057)	(0.070)	(0.066)
Geographic Distance	-1.676***	-1.996***	-1.010***
	(0.338)	(0.407)	(0.267)
Capital Controls	-0.028**	-0.007	-0.041***
	(0.012)	(0.013)	(0.014)
Ancestral Distance	20.417***	4.430	28.418***
	(7.824)	(10.661)	(6.479)
Common Colonizer	1.284**	0.865	2.237***
	(0.582)	(0.548)	(0.646)
Length of Diplomatic Tie	0.002	0.003	0.001
	(0.002)	(0.003)	(0.001)
Religious Distance	-2.831***	-2.520***	-2.984***
	(0.410)	(0.552)	(0.437)
Colonial Relationship	0.496***	0.415**	0.635***
	(0.139)	(0.205)	(0.149)
Common Sea	0.183	0.327*	-0.079
	(0.116)	(0.181)	(0.105)
Contiguity	0.389***	0.517***	0.329***
	(0.093)	(0.118)	(0.075)
Latitudinal Distance	0.009	0.014	0.004
	(0.006)	(0.009)	(0.004)
Longitudinal Distance	0.005**	0.004	0.004**
	(0.002)	(0.003)	(0.002)
Same Continent	-0.240	-0.310	-0.047
	(0.286)	(0.458)	(0.186)
Linguistic Distance	0.100	-1.118***	0.858**
	(0.362)	(0.411)	(0.408)
Trade Costs	-2.479	4.344	-9.633***
	(4.623)	(6.213)	(2.751)
Observations	20,662	29,615	25,434
R-squared	0.783	0.732	0.803

C Counterfactual Analysis using Alternative Estimates

The following tables replicates Table 7, using alternative estimates instead of the baseline OLS estimates for the investment-distance semi-elasticities (β). Table C.2 uses IV estimates, while Table C.1 uses Poisson regression estimates.

Table C.1: Counterfactuals using Poisson regression Estimates (2017)

	Observed (All Barriers)	*Zero-Gravity (No Barriers)*	*Cultural Distance*	*Institutional Distance*	*Geographic Distance*	*Capital Controls*
Welfare Statistics (Fixed Capital Stock)	(1)	(2)	(3)	(4)	(5)	(6)
World GDP (US$ trillions)	112.7	120.7	119.0	120.2	118.5	120.0
World GDP, % Difference in GDP from Zero-Gravity	-6.7%	0%	-1.5%	-0.4%	-1.8%	-0.6%
St.Dev. of $\log(k_i/\ell_i)$, % Difference from Zero-Gravity	+76.2%	0%	+18.5%	+9.6%	+42.3%	+8.3%
St.Dev. of $\log(y_i/\ell_i)$, % Difference from Zero-Gravity	+43.6%	0%	+10.9%	+6.4%	+18.4%	+3.8%
Welfare Statistics (Endogenous Capital Stock)						
World GDP (US$ trillions)	112.7	120.4	119.8	120.3	117.9	121.0
World GDP, % Difference in GDP from Zero-Gravity	-6.4%	0%	-0.6%	-0.1%	-2.1%	+0.4%
St.Dev. of $\log(k_i/\ell_i)$, % Difference from Zero-Gravity	+76.3%	0%	+15.2%	+9.4%	+48.4%	+9.1%
St.Dev. of $\log(y_i/\ell_i)$, % Difference from Zero-Gravity	+43.6%	0%	+9.0%	+6.3%	+19.9%	+4.1%

Table C.2: Counterfactuals using IV Estimates (2017)

	Observed (All Barriers)	*Zero-Gravity (No Barriers)*	*Cultural Distance*	*Institutional Distance*	*Geographic Distance*	*Capital Controls*
Welfare Statistics (Fixed Capital Stock)	(1)	(2)	(3)	(4)	(5)	(6)
World GDP (US$ trillions)	112.7	122.9	114.3	115.1	122.2	121.3
World GDP, % Difference in GDP from Zero-Gravity	-8.3%	0%	-7.0%	-6.4%	-0.6%	-1.3%
St.Dev. of $\log(k_i/\ell_i)$, % Difference from Zero-Gravity	+22.0%	0%	+56.2%	+71.6%	+17.5%	+12.1%
St.Dev. of $\log(y_i/\ell_i)$, % Difference from Zero-Gravity	+1.4%	0%	+31.6%	+36.9%	+8.9%	+6.3%
Welfare Statistics (Endogenous Capital Stock)						
World GDP (US$ trillions)	112.7	121.4	109.5	110.9	120.2	121.7
World GDP, % Difference in GDP from Zero-Gravity	-7.2%	0%	-9.8%	-8.7%	-1.0%	+0.2%
St.Dev. of $\log(k_i/\ell_i)$, % Difference from Zero-Gravity	+22.3%	0%	+94.0%	+97.1%	+20.9%	+12.9%
St.Dev. of $\log(y_i/\ell_i)$, % Difference from Zero-Gravity	+1.5%	0%	+53.6%	+50.1%	+10.1%	+6.8%

D Unweighted Poisson Regressions

In this Appendix we replicate Table 3 using an unweighted Poisson regression.

Table D.1: Poisson Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	FTI	FDI	FPI	FTI	FDI	FPI
Cultural Distance	-0.272*	-0.448**	-0.589	-0.257*	-0.484***	-0.159
	(0.160)	(0.196)	(0.372)	(0.142)	(0.176)	(0.157)
Institutional Distance	-0.294***	-0.204***	-0.714***	-0.318***	-0.169**	-0.425***
	(0.048)	(0.062)	(0.129)	(0.069)	(0.082)	(0.071)
Geographic Distance	-0.773***	-0.998***	-0.361**	-1.193***	-1.718***	-0.706**
	(0.078)	(0.103)	(0.161)	(0.278)	(0.327)	(0.292)
Capital Controls	-0.022**	-0.008	-0.199***	-0.023***	-0.008	-0.031**
	(0.009)	(0.012)	(0.060)	(0.009)	(0.012)	(0.013)
Controls	No	No	No	Yes	Yes	Yes
Observations	23,131	33,359	28,913	21,167	30,947	25,485