

Pellegrino, Bruno; Spolaore, Enrico; Wacziarg, Romain

**Working Paper**

## Barriers to Global Capital Allocation

CESifo Working Paper, No. 9086

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Pellegrino, Bruno; Spolaore, Enrico; Wacziarg, Romain (2021) : Barriers to Global Capital Allocation, CESifo Working Paper, No. 9086, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/235456>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Barriers to Global Capital Allocation

*Bruno Pellegrino, Enrico Spolaore, Romain Wacziarg*

## **Impressum:**

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email [office@cesifo.de](mailto:office@cesifo.de)

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: <https://www.cesifo.org/en/wp>

# Barriers to Global Capital Allocation

## Abstract

We quantify the impact of barriers to international investment, using a novel multi-country dynamic general equilibrium model with heterogeneous investors and imperfect capital mobility. Our model yields a gravity equation for bilateral foreign asset positions. We estimate this gravity equation using recently developed foreign investment data that have been restated to account for offshore investment and financing vehicles. We show that a parsimonious implementation of the model with four barriers (geographic distance, cultural distance, foreign investment taxation, and political risk) accounts for a large share of the observed variation in bilateral foreign investment positions. Our model predicts (out of sample) a significant home bias, higher rates of return on capital in emerging markets, as well as “upstream” capital flows. In our benchmark calibration, we estimate that the capital misallocation induced by these barriers reduces World GDP by 7%, 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 80% higher than it would be in a world without barriers to international investment, while the dispersion in output per employee is 42% higher.

JEL-Codes: E220, E440, F200, F300, F400, G150, O400.

Keywords: capital allocation, capital flows, foreign investment, culture, geography, gravity, international macroeconomics, international finance, misallocation, open economy.

*Bruno Pellegrino*

*University of Maryland, Robert H. Smith School of Business*

*USA – College Park, MD 20742*

*bpellegr@umd.edu*

*Enrico Spolaore*

*Tufts University, Department of Economics*

*USA – Medford, MA, 02155*

*enrico.spolaore@tufts.edu*

*Romain Wacziarg*

*UCLA Anderson School of Management*

*USA – Los Angeles CA 90095*

*wacziarg@ucla.edu*

May 2021

We thank Julian Di Giovanni and Dmitry Mukhin for their excellent discussions of this paper at ASSA 2021 and the Yale Junior Finance conference. We also thank Ariel Burstein, Henry Friedman, Jean Imbs, Sebnem Kalemli-Özcan, Pete Kyle, Andrei Levchenko, Max Maksimovic, Alessandro Rebucci, as well as seminar participants at CSEF, Tufts, the University of Maryland, UCLA, UT Austin, and the University of Verona for helpful comments. We thank Pablo Roa Prieto and Michele Schincaglia for outstanding research assistance.

# 1 Introduction

International capital flows have increased greatly in recent decades (Lane and Milesi-Ferretti, 2018). However, economists have been asking why we do not observe much larger flows from capital-abundant to capital-scarce economies (Lucas, 1990; Alfaro, Kalemli-Ozcan, and Volosovych, 2008), while portfolios continue to be disproportionately allocated towards domestic assets (“home bias” in international financial markets - French and Poterba, 1991; Lau et al., 2010; Coeurdacier and Rey, 2013). Are there significant barriers to global capital allocation, and do they matter for income and welfare?

To address such questions, this paper makes three new contributions. First, we provide a novel general-equilibrium model of international investment that can be taken to the data. Second, we use the model to estimate what factors prevent capital from flowing freely from one country to another, using (a) updated measures of international investment that take into account offshore investment and financing vehicles, (b) novel measures of cultural distance, capturing a broad set of differences in values and beliefs across countries, and (c) updated measures of foreign investment taxation and political risk. Third, we use our empirical estimates to quantify the impact of barriers on the efficient allocation of capital across countries, as well as cross-country inequality.

In our model, heterogeneous investors choose to allocate capital among different destinations with varying returns. Cross-border investment is affected by expropriation risk and taxation, as well as by intermediation costs, which we parametrize in terms of measures of physical and cultural distances between societies. 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 the investment frictions and other fundamentals. To calibrate the model, we estimate the gravity equation using recently-developed foreign investment data that have been restated by Damgaard, Elkjaer, and Johannesen (2019) and Coppola, Maggiori, Neiman, and Schreger (2020) from a residency to a nationality basis. These databases account for the presence of offshore investment and financing vehicles located in tax havens such as Bermuda and the Cayman Islands.<sup>1</sup> Intermediation costs and other obstacles to capital flows are captured using a parsimonious specification, with four types of barriers: geographic distance, cultural distance, foreign investment taxes, and political risk.

We obtain three sets of empirical results. First, our measures of barriers exert substantial effects on international financial positions, controlling for origin and destination fixed effects. The effects are similar for different subcategories of foreign investment (equity vs. debt, foreign direct investment vs. foreign portfolio investment). They are robust to using different specifications, and remain quantitatively large irrespective of the estimation method (OLS regressions, Poisson regression, and IV regressions). Our IV approach, used to ensure that our estimates are not amplified by reverse causality, is based on

---

<sup>1</sup>The first paper restates IMF and OECD data on foreign direct investment, while the latter restates data on portfolio investment.

the assumption that deeply-rooted factors, such as measures of ancestral and religious distance between populations, have an impact on contemporary measures of cultural 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 estimated barriers introduce significant capital misallocation across countries. Compared to a situation without barriers, World GDP is 7% lower. Interestingly, the effect of cultural distance is of a similar magnitude as the effect of geographic distance. Barriers to global capital allocation also contribute significantly to cross-country inequality. We find that the standard deviation of log capital per employee is 80% higher than it would be in a world without barriers to international financial flows, and the dispersion in output per employee is 42% higher.

Our core contribution is to bring together, in a structural general equilibrium framework, the theoretical and empirical literatures on gravity in international finance. Two path-breaking theoretical contributions in this area were Martin and Rey (2004) and Okawa and Van Wincoop (2012). The first provided a two-country, two-period model of international investment capturing a number of features of empirical gravity relations. The latter provided the first fully specified theoretical gravity model of foreign investment (broadly-defined) in a two-period multi-country setting. Our model differs from these contributions along several dimensions. We provide a multi-country, multi-period setting, and integrate assets markets with the real economy. Portfolio diversification at the country level arises, in our setting, from investor heterogeneity rather than risk aversion. This approach delivers a tractable, exact form for the gravity equation which does not rely on approximation methods.

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 as well: our theory is built to describe cross-border investment in a 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 specified gravity equation, where bilateral investment flows are proportional to the GDP of investor and destination countries. Other related models are those by Sellin and Werner (1993) and Jin (2012). Unlike previous contributions, our framework explicitly features geographic and cultural barriers to the international allocation of capital, and delivers an explicit gravity relationship that we take to the data, allowing us to perform counterfactuals.

Empirically, our paper contributes to a large literature that estimates the determinants of interna-

tional 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). Building on 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).

Building on this large empirical literature on the determinants of international flows, we take a further step by explicitly connecting theory and empirics in a unified framework, allowing us to quantify the extent of misallocation induced by barriers to international investment. Thus, our work also contributes to the literature on 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). This paper also contributes to related research on the effects of financial integration and globalization (Gourinchas and Jeanne, 2006, 2013; Caselli and Feyrer, 2007; Mendoza et al., 2009; Colacito and Croce, 2010; Hoxha et al., 2013; Broner and Ventura, 2016). Finally, it connects to the trade literature on multinational enterprises (Arkolakis et al., 2018; Ramondo and Rodríguez-Clare, 2013; Tintelnot, 2016), which uses gravity models to study multi-national production, a specific aspect of foreign direct investment. While we share the structural modeling approach of these contributions, our focus is instead on foreign investment in a broader sense and on capital misallocation. Since 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 origins of cross-country income differences (Hall and Jones, 1999; McGrattan and Schmitz Jr, 1999 - among many others).

## 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 heterogeneous 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  $i \in \{1, 2, \dots, n\}$ . Each country has fixed labor supply  $\ell_i$  and a representative firm (also called  $i$ ) that acts competitively and produces a perfectly-substitutable and tradable good using a Cobb-Douglas production function:

$$y_{it} = \omega_{it} k_{it}^{\theta_i} \ell_{it}^{1-\theta_i} \quad (2.1)$$

where  $y_{it}$  is the level of output and  $k_{it}$  and  $\ell_{it}$  denote capital and labor input.  $\omega_i$  is country  $i$ 's total factor productivity. The production function parameter  $\theta_i$ , 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:

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

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 next period's production. In what follows, we 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:

$$r_{it} = \theta_i \frac{y_{it}}{k_{it}} \quad w_{it} = (1 - \theta_i) \frac{y_{it}}{\ell_{it}} \quad (2.3)$$

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  $\kappa$ . In short, plants are a discretization of country  $i$ 's capital stock. This implies that the number of active plants in each country is  $k_i/\kappa$ , and that there are total  $K/\kappa$

plants distributed over the  $n$  countries, where:

$$K \stackrel{\text{def}}{=} \sum_i k_i \quad (2.4)$$

## 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  $\ell_i$  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:

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

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:

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

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

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:

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

We look for a steady-state equilibrium, with constant consumption, output, capital and saving ( $\mathbf{c}_t, \mathbf{y}_t, \mathbf{k}_t, \mathbf{s}_t$ ). Thus, we will 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  $\alpha$  of their labor earnings:

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

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

$$\sum_{i=1}^n k_{it+1} = \sum_{j=1}^n s_{jt} \quad (2.10)$$

Define  $a_{ij}$  as the assets purchased in country  $i$  by investors from country  $j$ . Therefore, 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  $j$

$$k_i = \sum_{j=1}^n a_{ij}; \quad (2.11)$$

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

$$s_j = \sum_{i=1}^n a_{ij} \quad (2.12)$$

### 2.3 Asset Allocation

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

$$\sigma_{ij} \stackrel{\text{def}}{=} \frac{a_{ij}}{s_j} \quad (2.13)$$

In matrix form, the following equation describes the flow of capital across countries:

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

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: an 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  $R(x, z)$ :

$$R(x, z) \stackrel{\text{def}}{=} (1 - \tau_{ij}) \cdot r_i \cdot e^{-\lambda(x, z)} \quad (2.15)$$

where  $r_i(x)$  is the return on capital in country  $i$  where plant  $x$  is located.  $\tau_{ij}$  is a tax imposed by country  $i$  to investors from country  $j$ . We intend  $\tau_{ij}$  to represent actual taxes levied by the destination country government, as well as broader government-imposed costs, such as the risk of expropriation.<sup>2</sup>  $\lambda(x, z)$  is

---

<sup>2</sup>It is possible to microfound the  $\tau_{ij}$  as the probability of  $j$  investors being expropriated. For this, the utility function of the investor needs to be made stochastic. Specifically, investors need to hold risk-neutral Epstein-Zin preferences with unitary intertemporal elasticity:

$$U(z) = (1 - \alpha) \log c_t(z) + \alpha \log \mathbf{E} c_{t+1}(z) \quad (2.16)$$

The resulting model will then be isomorphic to one where  $\tau_{ij}$  is simply the fraction of capital income expropriated.

the intermediation cost incurred by investor  $z$  to invest in plant  $x$ . We intend this term to capture the costs associated with monitoring, enforcing and acquiring information about cross-border investments.

We assume that  $\lambda(x, z)$  can be decomposed linearly into three components: 1) a 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 random idiosyncratic component that varies at the investor-plant pair level  $(x, z)$ ; 3) an investor  $(j)$  country-level rebate:

$$\lambda(x, z) = \mathbf{d}'_{ij} |\beta| - \xi(x, z) - G_j \quad (2.17)$$

The components of the vector

$$\mathbf{d}_{ij} = \left[ d_{ij}^1 \quad d_{ij}^2 \quad \cdots \quad d_{ij}^D \right]' \quad (2.18)$$

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 distance.

$\xi(x, z)$  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  $G_j$  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,  $\tau_{ij}$  does not (directly) affect the aggregate resource constraint: it distorts asset allocation, but it does not actually destroy capital. 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 cost shocks over different locations.<sup>3</sup> Hence, each agent  $z$  invests their savings  $s_j$  in the plant that offers the highest subjective return:

$$x^*(z) = \operatorname{argmax}_x \left[ \log(1 - \tau_{ij(x,z)}) + \log r_{i(x)} + \mathbf{d}'_{ij(x,z)} \beta + \xi(x, z) \right] \quad (2.19)$$

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  $x$ , located in country  $i$ , is:

$$\sigma_j(x) = \frac{(1 - \tau_{ij}) \cdot r_i \cdot \exp(\mathbf{d}'_{ij} \beta)}{\sum_{x'=1}^{K/\kappa} (1 - \tau_{i(x')j}) \cdot r_{i(x')} \cdot \exp(\mathbf{d}'_{i(x')j} \beta)} \quad (2.20)$$

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 heterogeneity, at the *aggregate* level it acts as a

---

<sup>3</sup>We allow for a consideration of risk in an extension of the model, see Section 2.7.

motive for diversification. This is 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 can only observe country-level investment positions).

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

$$\sigma_{ij} \stackrel{\text{def}}{=} \sum_{x \in i} \sigma_j(x) = \frac{k_i}{\kappa} \cdot \sigma_j(x) \quad \forall x \in j \in \{1, 2, \dots, n\} \quad (2.21)$$

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

$$\sigma_{ij} = \frac{(1 - \tau_{ij}) \cdot r_i k_i \cdot \exp(\mathbf{d}'_{ij} \beta)}{\sum_{c=1}^n (1 - \tau_{cj}) \cdot r_c k_c \cdot \exp(\mathbf{d}'_{cj} \beta)} \quad (2.22)$$

which mimics Eaton and Kortum (2002)'s equation for international trade shares. This expression does not depend on plant size ( $\kappa$ ). Hence, in practice, we can ignore the fact that  $k_i$  may not be divisible by  $\kappa$ , because we can make plants arbitrarily small ( $\kappa \rightarrow 0$ ) without affecting country portfolios.

## 2.4 Efficient Allocation of Capital

Let  $Y$  be World GDP:

$$Y \stackrel{\text{def}}{=} \sum_{i=1}^n Y_i \quad (2.23)$$

and let us call a vector  $\mathbf{k} = (k_1, k_2, \dots, k_n)'$ , 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  $K \stackrel{\text{def}}{=} \sum_{i=1}^n k_i$ .

We can now show that equilibrium in input and asset markets, combined with the absence of taxes on international investment ( $\tau_{ij} = 0$  for all  $i, j$  and  $\mathbf{d}'_{ij} \beta$  is invariant across  $i, j$ ), 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  $\sum_i \Delta k_i = 0$ :

$$\Delta Y \approx \sum_{i=1}^n r_i \Delta k_i \quad (2.24)$$

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  $r$ :

$$k_i = r^{-\frac{1}{1-\theta_i}} (\theta_i \omega_i)^{\frac{1}{1-\theta_i}} \ell_i \quad (2.25)$$

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  $i, j$ ), if there are no investment taxes or risk of expropriation ( $\tau_{ij} = 0$  for all  $i, j$ ) and investors are optimizing, then equation (2.22) implies that all countries have their capital invested in identical destination country portfolios:

$$\sigma_{ij} = \frac{r_i k_i}{\sum_{c=1}^n r_c k_c} = \frac{k_i}{K} \quad \forall (i, j) \quad (2.26)$$

The right hand side of the equation is a consequence of the fact that the share of capital invested in country  $i$  ( $\sigma_{ij}$ ) is independent of the origin country  $j$ . Because the right side of equation (2.26) 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 investment taxes ( $\tau_{ij} = 0$  for all  $i, j$ ).<sup>4</sup>

We have also shown that, if asset markets are in equilibrium and undistorted, all origin countries  $j$  hold identical portfolios of foreign assets ( $\sigma_{ij}$  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  $\mathbf{k}$ <sup>5</sup>:

$$\Delta Y \approx r \sum_{i=1}^n \Delta k_i - \frac{1}{2} \sum_{i=1}^n (1 - \theta_i) \frac{r}{k_i} (\Delta k_i)^2 \quad (2.27)$$

In order to focus on 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

<sup>4</sup>This is not a re-statement of the First Welfare Theorem, because it is a statement about GDP, not welfare.

<sup>5</sup>This expression is derived using the fact that, in equilibrium, the rate of return  $r$  is equal to the marginal product of capital

rearrange the second-order term as:

$$\frac{\Delta Y}{Y} \approx -\frac{1}{2} \sum_{i=1}^n (1 - \theta_i) \frac{rk_i}{Y} (\Delta \log k_i)^2 \quad (2.28)$$

We then use the following facts:

$$\Delta \log r_i = -(1 - \theta_i) \Delta \log k_i \quad (2.29)$$

$$r_i k_i = \theta_i y_i \quad (2.30)$$

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

$$\frac{\Delta Y}{Y} \approx -\frac{1}{2} \sum_{i=1}^n \frac{\theta_i}{1 - \theta_i} \cdot \frac{y_i}{Y} \cdot (\Delta \log r_i)^2 \quad (2.31)$$

This expression can be seen as a weighted measure of dispersion of the rate of return  $r_i$  across countries. This formula allows us to approximate World GDP loss based only on observed rates of return. A simple way to operationalize this measure is to compute half the weighted variance of  $r_i$ , where  $\left(\frac{\theta_i}{1 - \theta_i} \cdot \frac{y_i}{Y}\right)$  is the weight given to country  $i$ .<sup>6</sup>

## 2.5 Gravity

Since, in equilibrium, the capital share of income is equal to GDP times  $\theta_i$ , Equation (2.22) can be rearranged as follows:

$$\sigma_{ij} = \frac{(1 - \tau_{ij}) \cdot \theta_i y_i \cdot \exp(\mathbf{d}'_{ij} \beta)}{\sum_{c=1}^n (1 - \tau_{cj}) \cdot \theta_c y_c \cdot \exp(\mathbf{d}'_{cj} \beta)} \quad (2.32)$$

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  $G_j$ :

$$G_j \stackrel{\text{def}}{=} \sum_{c=1}^n (1 - \tau_{cj}) \theta_c y_c \cdot \exp(\mathbf{d}'_{cj} \beta) \quad (2.33)$$

Multiplying both sides by  $s_j$  and using the fact that  $s_j = \alpha(1 - \theta_j) y_j$ , equation (2.32) can be rearranged as a gravity equation:

$$\boxed{\text{Gravity}} : a_{ij} = \theta_i (1 - \theta_j) (1 - \tau_{ij}) \cdot \frac{\alpha}{G_j} \cdot \frac{y_i \cdot y_j}{\exp|\mathbf{d}'_{ij} \beta|} \quad (2.34)$$

---

<sup>6</sup>This computation assumes that the efficient rate  $r$  is approximately equal to the weighted mean of  $r_i$ .

## 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  $\Sigma$  is a function of the vector of country output ( $\mathbf{y}$ ) and the matrices of the distortions  $\tau_{ij}$  and the distances  $\mathbf{d}_{ij}$  – that is:

$$\Sigma = \Sigma(\mathbf{y}; \Theta, \Delta, \mathbf{T}) \quad (2.35)$$

$$\text{with } \Delta \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{d}'\beta_{11}, & \mathbf{d}'\beta_{12}, & \dots & \mathbf{d}'\beta_{1n} \\ \mathbf{d}'\beta_{21}, & \mathbf{d}'\beta_{22}, & \dots & \mathbf{d}'\beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{d}'\beta_{n1}, & \mathbf{d}'\beta_{n2}, & \dots & \mathbf{d}'\beta_{nn} \end{bmatrix}; \quad \Theta \stackrel{\text{def}}{=} \begin{bmatrix} \theta_1, & 0, & \dots & 0 \\ 0 & \theta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_n \end{bmatrix} \quad (2.36)$$

$$\mathbf{T} \stackrel{\text{def}}{=} \begin{bmatrix} \tau_{11}, & \tau_{12}, & \dots & \tau_{1n}, \\ \tau_{21}, & \tau_{22}, & \dots & \tau_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{n1}, & \tau_{n2}, & \dots & \tau_{nn} \end{bmatrix}; \quad (2.37)$$

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

$$\mathbf{k} = \Sigma(\mathbf{y}(\mathbf{k}); \Theta, \Delta, \mathbf{T}) \cdot \alpha(\mathbf{I} - \Theta)\mathbf{y}(\mathbf{k}) \quad (2.38)$$

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 (2.38). Notice that there is a trivial equilibrium at  $\mathbf{k} = \mathbf{0}$ . When we solve equation (2.38) 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:

$$c_j = \mathbf{r}'\mathbf{a}_j + (1 - \theta_j)w_j\ell_j \quad (2.39)$$

where

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} & a_{2j} & \dots & a_{nj} \end{bmatrix}' \quad (2.40)$$

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

$$\underbrace{c_j - (y_j - s_j)}_{\text{Net Imports}} = \mathbf{r}'\mathbf{a}_j - \alpha w_j\ell_j - r_j k_j + s_j = \underbrace{\mathbf{r}'\mathbf{a}_j - r_j k_j}_{\text{Net Foreign Income}} \quad (2.41)$$

That is, all consumption in excess of production (net of savings) is financed by a positive net foreign capital

income. Conversely, a negative balance in net foreign income has to be balanced by a trade surplus.

## 2.7 Extensions

Our model can be extended to accommodate additional barriers and frictions to global capital allocation. In Section 6, we present three extensions: we model capital controls, the presence of currency hedging costs, and risky asset returns in destination countries. In that section, we also explore the empirical implications of allowing for these additional frictions.

## 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:

$$\hat{a}_{ij} = a_{ij} \cdot \exp(\varepsilon_{ij}) \quad \text{with} \quad \mathbf{E}(\varepsilon_{ij} | i, j, \mathbf{d}_{ij}, \tau_{ij}) = 0 \quad (3.1)$$

we can then re-write equation (2.34) as the following fixed effects linear regression model for the log of foreign investment:

$$\log \hat{a}_{ij} = \mu_i + \eta_j + \log(1 - \tau_{ij}) + \mathbf{d}'_{ij}\beta + \varepsilon_{ij} \quad (3.2)$$

where  $\mu_i$  is a country of origin fixed effect and  $\eta_j$  is a country of origin fixed effect. Because our foreign investment data has a panel structure, we expand the gravity equation above to include the time dimension:

$$\log \hat{a}_{ijt} = \mu_{it} + \eta_{jt} + \log(1 - \tau_{ij}) + \mathbf{d}'_{ij}\beta + \varepsilon_{ijt} \quad (3.3)$$

This is our main econometric specification. The dependent variable is measured using data on *Foreign Equity Investment*, *Foreign Debt Investment*, and the sum of the two (*Total Foreign Assets*).<sup>7</sup> To capture  $\mathbf{d}$ , we propose a parsimonious specification based on two measures of distance, hypothesized to capture information asymmetries in international investment: *Geographic Distance* and *Cultural Distance*. To capture  $\tau_{ij}$ , we use bilateral data on investment taxes and risk of expropriation.  $\mu_{it}$  and  $\eta_{jt}$  are year-specific destination and origin country fixed effects.

Because the vector of distances  $\mathbf{d}$  varies at the level of the *undirected* country pair, in our regression analysis we compute standard errors clustered by undirected country pair. Additional bilateral variables

---

<sup>7</sup>In the Appendix, we also consider the determinants of global asset holdings that distinguishes between *Foreign Direct Investment* and *Foreign Portfolio Investment*, as is often done in the literature. We prefer to focus on the debt / equity distinction in the main analysis because the distinction between FDI and equity FPI is somewhat arbitrary. See for instance Blanchard and Acalin (2016) and Alfaro and Chauvin (2020).

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 account 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 tax-efficient manner. 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. This is the source of our FDI data.

Regarding portfolio investment, our main source is data from Coppola, Maggiori, Neiman, and Schreger (2020). They use data from IMF’s Coordinated Portfolio Investment Survey (CPIS), and restate them to account for the presence of shell companies in tax havens - often used to issue securities. To do so, they use *reallocation matrices* based on fund holdings data from Morningstar. Using these reallocation matrices, they converted international portfolio data from CPIS from a residency basis to a nationality basis. Their Foreign Portfolio Investment (FPI) data is further broken down between debt and equity. This is the source of our FPI data.<sup>8</sup>

To obtain a measure of *Total Foreign Assets* (or *Foreign Total Investment*, FTI), we sum the FPI and FDI series (both are in current international US Dollars). Further, we create a series of *Foreign Equity Investment* by adding up FDI and the equity portion of FPI, and a series for *Foreign Debt Investment* by isolating the debt portion of the FPI series.

For both *Foreign Debt Investment* and *Foreign Equity Investment*, we base our econometric estimates on panel data from 2009 to 2017 (the equity series starts in 2013). Figure 1 displays the corresponding 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 foreign debt and equity investment across countries. These two variables range from a few hundred thousand dollars to over a trillion dollars. This is the variation we seek to explain. Second, the two variables correlate very strongly ( $\rho = 0.68$ ), and line up neatly on the 45° line, indicating that they are similar in size and tend to track each other closely. This suggests

---

<sup>8</sup>Coppola et al. (2020) combine all European Monetary Union countries into a single entity. We use unrestated FPI data from CPIS to break down aggregate EMU figures into investment data corresponding to each individual member country.

FIGURE 1: FOREIGN EQUITY AND DEBT INVESTMENT (2017)

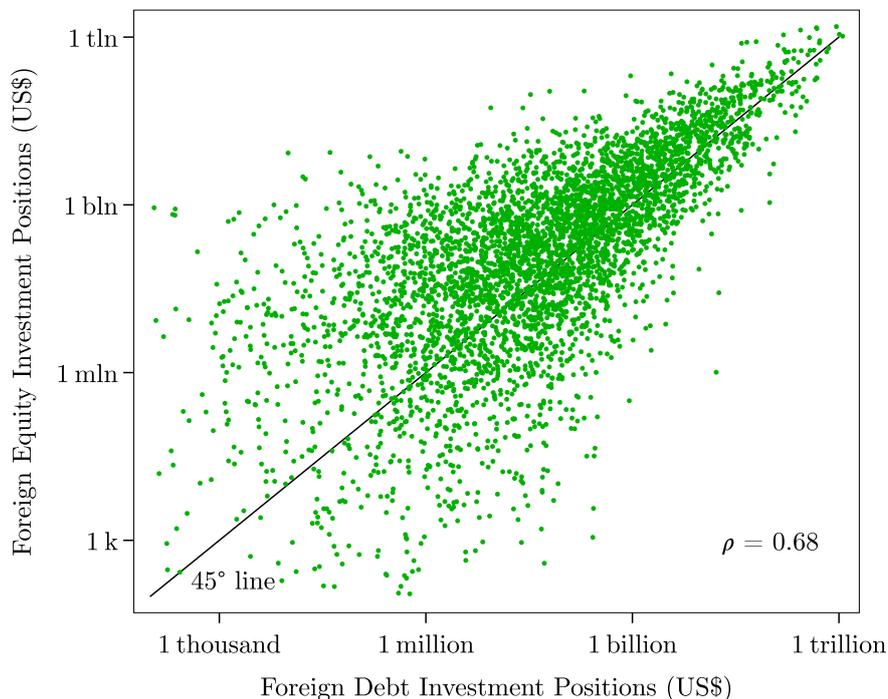


FIGURE NOTES: This figure plots restated Foreign Equity Investment against restated Foreign Debt Investment. Each observation is a country pair and all data refer to the year 2017. The unit of measurement is US dollars at current prices. Log scale on both axes.

that they might be driven by a similar set of underlying factors, an issue that our econometric analysis will clarify. Similar observations hold for the distinction between FDI and FPI, which are considered as alternative dependent variables in the Appendix.

### 3.3 Main Explanatory Variables

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 cultural distance and geographic distance. Other impediments to global capital flows are considered in the counterfactual analysis (bilateral taxes on foreign investment, political risk) as well as in the Appendix (currency risk, capital controls).

#### 3.3.1 Cultural and Geographic Distance

Our vector of distances  $\mathbf{d}_{ij}$  includes two distances: *Cultural Distance* and *Geographic 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 set of 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 *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 the two effects can be compared.

### 3.3.2 Foreign Investment Taxes and Political Risk

Our bilateral measure of taxation,  $\tau_{ij}$  captures two factors: actual taxation on capital income paid to foreign entities, as well as political risk (i.e. probability of expropriation). Formally:

$$(1 - \tau_{ij}) = \left(1 - \tau_{ij}^k\right) (1 - \pi_i) \quad \text{for } i \neq j \quad (3.4)$$

where  $\tau_{ij}^k$  is foreign investment taxation and  $\pi_{ij}$  is the rate of expropriation. We measure these two factors separately.

Previous studies have used Dual Taxation Treaties (DTTs), also referred to as Dual Taxation Agreements (DTAs), to predict cross-border investment positions and flows (Di Giovanni, 2005). Tax treaties are expected to increase investment flows between contracting parties. For this study, we were able to obtain a more direct measure of the bilateral tax rate  $\tau_{ij}^k$ . We measure it using withholding rates on interests and dividend income disbursed to non-resident individuals, which we obtain from Tax Research Platform of the International Bureau of Fiscal Documentation (IBFD). This platform estimates bilateral withholding rates between countries taking into consideration double taxation agreements that are in force between countries.

The IBFD platform reports “domestic rates” (withholding rates that apply to interests and dividends not covered by a tax treaty) as well as beneficiary country-specific “Treaty Rates”. DTTs are generally designed to impose a cap on the tax rates applied to interests and dividends paid out to non-resident entities. It is therefore possible for domestic rates to be *lower* than treaty rates (in other words, treaty rate caps may not necessarily bind). In such cases, the (lower) domestic withholding rate generally applies. For this reason, when one domestic rate is reported and one treaty rate is reported for investors from some

country  $j$ , we take the lower of the two as our estimate for the tax rate applied to country  $j$  investors.

One complication is that IBFD might report multiple domestic rates, as well as multiple treaty rates for the same residency country and income type. Italy, for example, imposes a domestic withholding rate on interest income (disbursed to non-residents) that may be 0%, 12.5% or 26% (depending on the source of the income and other specificities). At the same time, it imposes a maximum treaty rate of 4%, 5%, 10% or 15% on interest income disbursed to Chilean investors as a consequence of the Italy-Chile tax treaty.

To compute a single estimate the applicable tax rate to each country (for either dividends or interests), we consider every possible combination/pair of domestic and treaty rates (0% and 4%, 0% and 5%, and so on and so forth). Within each pair, we take the lower between the domestic and the treaty rate, and then take the median across all pairs. Finally, we average the withholding rate for interest income with that for dividend income. This yields our final estimate of  $\tau_{ij}^k$ .

To measure the rate of expropriation  $\pi_{ij}$  we use empirical estimates from Alfaro, Kalemli-Ozcan, and Volosovych (2008, henceforth AKV), who measure the sensitivity of foreign investment inflows (in millions of US\$) per capita to a composite measure produced by the International Country Risk Group (ICRG). This measure, which we obtain for every year in our dataset, ranges from zero (extreme political risk) to ten (virtually no political risk).<sup>9</sup>

To infer from their estimates a measure of political risk that we can input into our model, we use the fact that, in our model, for a small country  $i$ :

$$\frac{\partial \log \sum_j a_{ij}}{\partial \log (1 - \pi_i)} \approx 1 \quad (3.5)$$

From the AKV's regression and summary statistics tables, we can compute:

$$\frac{\partial \log \sum_j a_{ij}}{\partial \text{PR}_i} = \underbrace{\left( \frac{\partial \sum_j a_{ij}}{\partial \text{PR}_i} \cdot \frac{1}{\text{Population}_i} \right)}_{\beta_{\text{AKV}}} / \frac{\sum_j a_{ij}}{\text{Population}_i} \quad (3.6)$$

where  $\text{PR}_i$  is ICRG's measure of political risk and the term in parentheses ( $\beta_{\text{AKV}}$ ) is the regression coefficient estimated by AKV, and the denominator on the right (foreign investment inflows per capita) can be obtained from AKV's summary statistics table. Combining the two equations above, and assuming that  $\pi_{ij} = 0$  when  $\text{PR}_i = 10$ , we then have the following estimate for the expropriation rate ( $1 - \pi_{ij}$ ):

$$(1 - \pi_i) = \exp \left[ \beta_{\text{AKV}} (\text{PR}_i - 10) \frac{\text{Population}_i}{\sum_j a_{ij}} \right] \quad (3.7)$$

AKV perform instrumental variable regressions using two different datasets in their analysis (IMF

---

<sup>9</sup>The political risk index is missing for a handful of countries, for whom we input a political risk score of 5.

and KLSV). For each of the two datasets, we use the  $\beta_{AKV}$  estimate that controls for the initial level of GDP per capita. This leads to two different estimates for  $\pi_i$ . To calibrate our model, we take the simple average of the two.

Finally, we exclude from our computations domestic taxation (which, if applied uniformly to all foreign income, does not affect capital allocation) and we assume that domestic investors are not subject to expropriation risk, i.e. we set  $\tau_{ij} = 0$  for  $i = j$ .

### 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 are deep determinants of contemporary 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).<sup>10</sup> 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

---

<sup>10</sup>The analogy is not perfect because individuals have two parents, while languages typically evolve sequentially from “ancestor” languages. For example, the ancestors of the Italian language, according to Ethnologue are, in order: Indo-European, Italic, Romance, Italo-Western, and Italo-Dalmatian.

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).

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. Their variation provides no direct information about overall differences in genetic endowments, so it would not be appropriate to interpret the effects of ancestral distance on 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 (Appendix A provides a formal treatment of this idea). Hence, 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. Additionally, 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.<sup>11</sup>

We obtain the control variable *Currency Peg* from the dataset of exchange rate arrangements con-

---

<sup>11</sup>The data are from CEPII and can be obtained at [http://www.cepii.fr/CEPII/fr/bdd\\_modele/presentation.asp?id=6](http://www.cepii.fr/CEPII/fr/bdd_modele/presentation.asp?id=6)

TABLE 1: SUMMARY STATISTICS

**Panel A: Directed, Time-Varying Variables**

	Observations	Mean	StDev	Min	Max
Foreign Assets (US\$ mln, 2013-2017)	13,551	34,093	590,000	0	35,100,000
Foreign Equity Assets (US\$ mln, 2013-2017)	13,598	19,356	322,000	0	20,400,000
Foreign Debt Assets (US\$ mln, 2009-2017)	29,807	12,732	264,000	0	21,000,000
Trade Costs (2009-2017)	43,352	0.051	0.046	0	1.322

**Panel B: Directed, Cross-Sectional Variables**

	Directed Pairs	Mean	StDev	Min	Max
Currency Hedging Costs ( $\gamma_{ij}$ )	4,900	0.000	0.069	-0.209	0.209
Foreign Investment Tax ( $\tau_{ij}$ )	4,900	0.100	0.068	0.000	0.523

**Panel C: Undirected, Time-Varying Variables**

	Undirected Pairs	Mean	StDev	Min	Max
Currency Peg (2009-2017)	22,365	0.383	0.486	0.000	1.000
Investment Treaty (2009-2017)	22,365	0.488	0.500	0.000	1.000

**Panel D: Undirected, Cross-Sectional Variables**

	Undirected Pairs	Mean	StDev	Min	Max
Cultural Distance	2,485	4.196	1.746	0.000	10.000
Geographic Distance	2,485	3.069	2.339	0.004	9.764
Ancestral Distance	2,346	0.021	0.016	0.000	0.074
Religious Distance	2,281	0.786	0.209	0.000	0.999
Border Contiguity	2,485	0.036	0.187	0.000	1.000
Colonial Relationship	2,485	0.025	0.156	0.000	1.000
Common Colonizer	2,485	0.030	0.171	0.000	1.000
Common Legal Origin	2,485	0.357	0.479	0.000	1.000
Common Sea	2,415	0.110	0.313	0.000	1.000
Customs Union	2,485	0.173	0.379	0.000	1.000
Economic Integration Agreement	2,485	0.272	0.445	0.000	1.000
Free Trade Area	2,485	0.377	0.485	0.000	1.000
Latitudinal Distance	2,415	28.857	25.699	0.000	106.000
Linguistic Distance	2,281	0.935	0.193	0.000	1.000
Longitudinal Distance	2,415	58.708	51.576	0.000	276.000

structured by Ilzetki, Reinhart, and Rogoff (2019). The dataset coverage ends in 2016. We extend the data to 2017 by carrying over the last year of data. The variable *Currency Peg* captures the presence of a currency peg as well as the presence of a common currency. We also obtained, from the World Bank’s International Center for the Settlement of Investment Disputes (ICSID), data on the presence of bilateral investment treaties, which we code as the dummy variable *Investment Treaty*.<sup>12</sup>

To control for trade policy, we obtain data on regional trade agreements (RTAs) and their member countries from the WTO websites. We construct bilateral dummy variables representing joint memberships in *Customs Union*, *Free Trade Agreements*, and *Economic Integration Agreements* as of 2017. 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<sup>13</sup> 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 70 countries, i.e.  $70 \times 70 = 4,900$  directed country pairs-observations (including diagonal *i-to-i* pairs) or 2,415 undirected country pairs, over the 2009-2017 period. The 70 countries in our dataset cover about 92% of the World GDP (based on 2017 data from the Penn World Tables, version 9.1). Table 1 displays summary statistics for the data described above.

## 4 Econometric Analysis

In this section, we estimate the parameter vector  $\beta$ , the effect of geographic and cultural distances on log foreign investment (two semi-elasticities). 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 of Section 2, in order to conduct counterfactual analysis.

---

<sup>12</sup><https://icsid.worldbank.org/resources/publications/investment-treaty-series>

<sup>13</sup><https://www.unescap.org/resources/escap-world-bank-trade-cost-database>

TABLE 2: OLS REGRESSIONS

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dep.variable in logs:</i>	Assets	Equity	Debt	Assets	Equity	Debt
Cultural Distance	-0.419** (0.038)	-0.461** (0.047)	-0.301** (0.033)	-0.333** (0.041)	-0.395** (0.050)	-0.230** (0.033)
Geographic Distance	-0.526** (0.031)	-0.577** (0.038)	-0.348** (0.025)	-0.398** (0.076)	-0.526** (0.094)	-0.288** (0.058)
Control Variables	No	No	No	Yes	Yes	Yes
Observations	12,542	12,014	22,386	11,157	10,746	19,899
<i>R</i> -squared	0.771	0.716	0.831	0.803	0.759	0.850
Within <i>R</i> -squared	0.214	0.196	0.137	0.327	0.311	0.250

TABLE NOTES: This table reports OLS estimates of a linear regression of the log of the variable listed on the top row (*Foreign Assets, Foreign Equity Assets, Foreign Debt Assets*) on the variables in the leftmost column, using data from 2013-2017 for columns 1, 2, 4 and 5, and 2009-2017 for columns 3 and 6. Each observation is a directed country pair. All regressions include origin country  $\times$  time ( $i, t$ ) fixed effects and destination country  $\times$  time ( $j, t$ ) fixed effects. Additional controls in columns 4-6 are *Border Contiguity, Common Sea, Latitudinal Distance, Longitudinal Distance, Colonial Relationship, Common Colonizer, Common Legal Origin, Linguistic Distance, Currency Peg, Customs Union, Economic Integration Agreement, Free-Trade Agreement, Foreign Investment Tax, Investment Treaty* and *Trade Costs*. Standard errors (clustered by undirected country pair) in parentheses. \* $p < .05$ ; \*\* $p < .01$

#### 4.1 Least Squares Analysis

We begin by performing an OLS regression of the log of foreign investment (debt, equity or total) on the two main measures of distance. Table 2 reports the estimates. Column (1) presents estimation results with for the log of total assets (i.e. Foreign Total Investment or FTI), as the dependent variable. We find that *Cultural Distance* and *Geographic Distance* are both statistically and economically significant predictors of FTI: the slope coefficients corresponding to these two variables are negative, sizable in magnitude (-0.419 and -0.526 respectively) 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 large effects of both barriers: a one standard deviation increase of geographic distance (2.339 units) is associated with a 70.8% decrease in FTI, while a one standard deviation increase in cultural distance (1.746 units) is associated with a 51.9% decrease in FTI.

In Column (2) we present estimation results using log foreign equity investment as the dependent variable. We find again that both barriers are statistically and economically significant: the standardized effects as defined above are slightly larger than those for log FTI. Column (3) considers log foreign debt investment as the dependent variable. We find smaller effects of geographic distance (a standardized effect of -55.7%) and cultural distance (with a standardized effect of -40.9%), both statistically significant at the 5% level.

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 measures of geographic distance (contiguity, access to a common sea, latitudinal distance, longitudinal distance), common history variables (linguistic distance, past colonial relationship, common colonizer, common legal origins), as well as variables possibly capturing bilateral facilitators of capital exchange (currency peg, customs union, economic integration agreement, free-trade agreement, foreign investment tax, investment treaty and trade costs). The coefficient estimates on cultural and geographic distances are somewhat reduced in magnitude: for FTI, we find standardized effects of cultural distance and geographic distance equal, respectively, to -44.1% and -60.6%. We again find that these two barriers have quantitatively larger effects on foreign equity investments than foreign debt investment. In sum, adding control variables does not fundamentally alter the inferences drawn from the more parsimonious specification.

## 4.2 Pseudo-Poisson Regressions

One shortcoming of the econometric model described by equation (3.3) 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 (3.3) as:

$$\hat{a}_{ijt} = \exp [\mu_{it} + \eta_{jt} + \log(1 - \tau_{ij}) + \mathbf{d}'_{ij}\beta + \varepsilon_{ijt}] \quad (4.1)$$

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 weigh observations by the inverse of the geometric mean of the GDPs of countries  $i$  and  $j$  (un-weighted estimates, which have larger standard errors, are shown in Appendix G). Including the zero investment pairs, the size of the sample rises substantially compared to that in Table 2 (by over 30% for foreign debt investment and about 13% for equity, though the increase is smaller for total investment, at about 8%).

Table 3 displays the resulting estimates. In general, we find that the standardized magnitude of Poisson estimates are slightly smaller than the corresponding OLS estimates, and the standardized effect of cultural

TABLE 3: PSEUDO-POISSON REGRESSIONS

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dependent variable:</i>	Assets	Equity	Debt	Assets	Equity	Debt
Cultural Distance	-0.420** (0.069)	-0.346** (0.054)	-0.480** (0.101)	-0.181** (0.047)	-0.151** (0.055)	-0.287** (0.058)
Geographic Distance	-0.265** (0.046)	-0.262** (0.043)	-0.384** (0.066)	-0.299** (0.296)	-0.365** (0.312)	-0.196 <sup>†</sup> (0.293)
Control Variables	No	No	No	Yes	Yes	Yes
Observations	13,547	13,594	29,737	12,076	12,118	26,110

TABLE NOTES: This table reports Iteratively-Reweighted Least Squares (IRLS) estimates of a Pseudo-Poisson regression of the variables listed on the topmost row (*Foreign Assets, Foreign Equity Assets, Foreign Debt Assets*) on the variables in the leftmost column. Each observation is a directed country pair. All regressions include origin country  $\times$  time ( $i, t$ ) fixed effects and destination country  $\times$  time ( $j, t$ ) fixed effects. Additional controls in columns 4-6 are *Border Contiguity, Common Sea, Latitudinal Distance, Longitudinal Distance, Colonial Relationship, Common Colonizer, Common Legal Origin, Linguistic Distance, Currency Peg, Customs Union, Economic Integration Agreement, Free-Trade Agreement, Foreign Investment Tax, Investment Treaty* and *Trade Costs*. Observations are weighted by the inverse of the geometric average of destination and origin country GDP. Standard errors (clustered by undirected country pair) in parentheses. <sup>†</sup> $p < .10$ ; \* $p < .05$ ; \*\* $p < .01$

distance is now sometimes larger than that of geographic distance. For instance, in the specification of column 1, the standardized effect of cultural distance is to reduce total foreign assets by 52.0% while that of geographic distance is -46.2%. In contrast to the results that use only the intensive margin, we no longer find smaller effects for debt compared to equity foreign investment. Broadly speaking, a consideration of the extensive margin does not greatly affect our basic finding that both geographic and cultural barriers exert quantitatively meaningful and statistically significant negative effects of foreign asset holdings - and that their respective effects are commensurate with each other.

### 4.3 Instrumental Variables Regressions

A challenge in estimating the effect of cultural 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) as a consequence of more intense cross-border investment.<sup>14</sup> In that case,

<sup>14</sup>For obvious reasons, reverse causality is not an issue for geographic distance.

TABLE 4: FIRST-STAGE REGRESSIONS

	(1)	(2)
	Cultural Distance	Cultural Distance
Ancestral Distance	39.132** (4.534)	25.771** (3.538)
Religious Distance	3.843** (0.294)	2.962** (0.269)
Control Variables	No	Yes
Observations	11,022	11,022
$R$ -squared	0.653	0.737
Within $R$ -squared	0.227	0.412
Cragg-McDonald Wald $F$ statistic	1067.358	793.211
Kleibergen-Paap Wald $F$ statistic	100.663	93.913
Stock & Yogo Critical Value ( $r=10\%$ )	19.93	19.93

TABLE NOTES: This table 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. These correspond to the first stage of the IV regressions (1) and (4) presented in Table 5. Each observation is an undirected country pair. All regressions include origin country  $\times$  time ( $i, t$ ) fixed effects and destination country  $\times$  time ( $j, t$ ) fixed effects. All regressions control for *Geographic Distance*. Additional controls in columns 3 and 4 are *Border Contiguity*, *Common Sea*, *Latitudinal Distance*, *Longitudinal Distance*, *Colonial Relationship*, *Common Colonizer*, *Common Legal Origin*, *Linguistic Distance*, *Currency Peg*, *Customs Union*, *Economic Integration Agreement*, *Free-Trade Agreement*, *Foreign Investment Tax*, *Investment Treaty* and *Trade Costs*. Robust standard errors in parentheses. \* $p < .05$ ; \*\* $p < .01$

the OLS estimates of the gravity equation (3.3) 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 only influence financial flows indirectly, through their effect on contemporary cultural distance. Consistent with this exclusion restriction, we consider two instruments for cultural distance: *Ancestral Distance* 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). A stronger argument for excludability can be made for *Ancestral Distance*, because it is very plausible that such a 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, captured by our measure of *Cultural Distance*.<sup>15</sup> 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), measured by *Cultural Distance*.

Table 4 presents estimation results for the first-stage regressions. We present results for the parsimonious specification (column 1), and for the specification with additional controls (column 2). First stage regressions lead to interesting results. Consistent with findings in Spolaore and Wacziarg (2016), ancestral and religious distances are both positively and significantly correlated with cultural distance: the instruments are strongly predictive of the endogenous variable in the first stage, as shown by the two first stage F-statistics presented on Table 4.

Results for the second stage appear in Table 5. As before, there are 6 columns, corresponding to three dependent variables (log total foreign assets, log foreign equity investment and log debt investment) and to whether we include additional controls or not. *Cultural Distance* is treated as endogenous. Compared to the OLS results of Table 2, we find that the magnitude of the effect of cultural distance rises significantly.<sup>16</sup> 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 -51.9% under OLS, and it rises in magnitude to -79.7% under IV. Similar differences are seen across specifications. On the other hand, across specifications the standardized magnitude of the effect of geographic distance is roughly unchanged compared to OLS (in column 1, it is -67.2% versus -70.8% under OLS, for instance).

The bottom line from the IV results is that both distance metrics continue to remain statistically and economically significant, with a larger effects of cultural distance compared to OLS. These findings do not depend greatly on whether we control for additional determinants of foreign investment, and is robust for log total foreign assets, log foreign debt assets and log foreign equity assets. 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.

---

<sup>15</sup>A theoretical formalization of the relationship between *Ancestral Distance* and *Cultural Distance* is provided in Appendix A.

<sup>16</sup>Finding IV estimates on the instrumented variables that are larger in magnitude than OLS estimates is quite common in the literature, even in cases (like ours) where we expect reverse causality to bias OLS estimates away from zero. A common explanation is that IV estimation helps address attenuation bias coming from measurement error, if error in measurement of the instrumental variables is uncorrelated with error in measurement of the instrumented (endogenous) regressor.

TABLE 5: INSTRUMENTAL VARIABLES REGRESSIONS

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dep.variable in logs:</i>	Assets	Equity	Debt	Assets	Equity	Debt
Cultural Distance	-0.912** (0.104)	-1.018** (0.128)	-0.536** (0.077)	-0.711** (0.122)	-0.793** (0.151)	-0.338** (0.087)
Geographic Distance	-0.477** (0.034)	-0.532** (0.042)	-0.350** (0.027)	-0.411** (0.084)	-0.537** (0.098)	-0.302** (0.060)
Control Variables	No	No	No	Yes	Yes	Yes
Observations	11,022	10,611	19,704	11,022	10,611	19,704
Within $R$ -squared	0.177	0.167	0.140	0.305	0.295	0.245
Kleibergen-Paap $\chi$ P-value	0.000	0.000	0.000	0.000	0.000	0.000
Sargan-Hansen $J$ P-value	0.548	0.654	0.088	0.304	0.276	0.786

TABLE NOTES: This table reports Instrumental Variable (IV) estimates of a linear regression of the log of the variable listed on the top row (*Foreign Assets*, *Foreign Equity Assets*, *Foreign Debt Assets*) on the variables in the leftmost column. *Cultural Distance* is the endogenous regressor and the excluded instruments are *Ancestral Distance* and *Religious Distance*. Each observation is a directed country pair. All regressions include origin country  $\times$  time ( $i, t$ ) fixed effects and destination country  $\times$  time ( $j, t$ ) fixed effects. The additional controls in columns 4-6 are *Border Contiguity*, *Common Sea*, *Latitudinal Distance*, *Longitudinal Distance*, *Colonial Relationship*, *Common Colonizer*, *Common Legal Origin*, *Linguistic Distance*, *Currency Peg*, *Customs Union*, *Economic Integration Agreement*, *Free-Trade Agreement*, *Foreign Investment Tax*, *Investment Treaty* and *Trade Costs*. Standard errors (clustered by undirected country pair) in parentheses.  $*p < .05$ ;  $**p < .01$

## 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 geographic and cultural 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, and other cultural 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 costs associated with geographic and cultural distances. Thus, we will look at a world where geopolitical 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 and cultural 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 ( $\ell_i$ ). We measure this as total employment, which we obtain from 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, used for model validation purposes. The last data ingredient is the output-capital elasticity  $\theta_i$ . We measure this as one minus the labor income share of GDP, for which we obtain country-level estimates from the International Labor Office (ILO), Department of Statistics. Finally, we calibrate  $\alpha = 1/2$ .<sup>17</sup>

Our regression estimates for  $\beta$  vary somewhat depending on the estimator. As noted earlier, IV estimates tend to be larger for *Cultural 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 ( $\beta$ ) using the estimates of column 4 of Table 2 (which includes the full set of controls): -0.333 for *Cultural Distance*, and -0.398 for *Geographic Distance*. We also use data on  $\tau_{ij}$ , composed of both investment taxes and risk of expropriation, as described in Section 3.3.2.

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  $\omega_i$ . These objects are identified given the previously-measured variables and parameters: the matrix  $\mathbf{\Sigma}$  is identified given  $\mathbf{\Theta y}$  and  $\mathbf{\Delta}$  (equation (2.32)). The vector of savings  $\mathbf{s}$  is identified by equation (2.9).  $\mathbf{k}$  is then

---

<sup>17</sup>The parametrization of  $\alpha$  is inconsequential from an economic standpoint (it does not affect any of our counterfactuals): all it does is to scale the overall level of capital across all countries).

FIGURE 2: EMPIRICAL PERFORMANCE OF THE GRAVITY MODEL

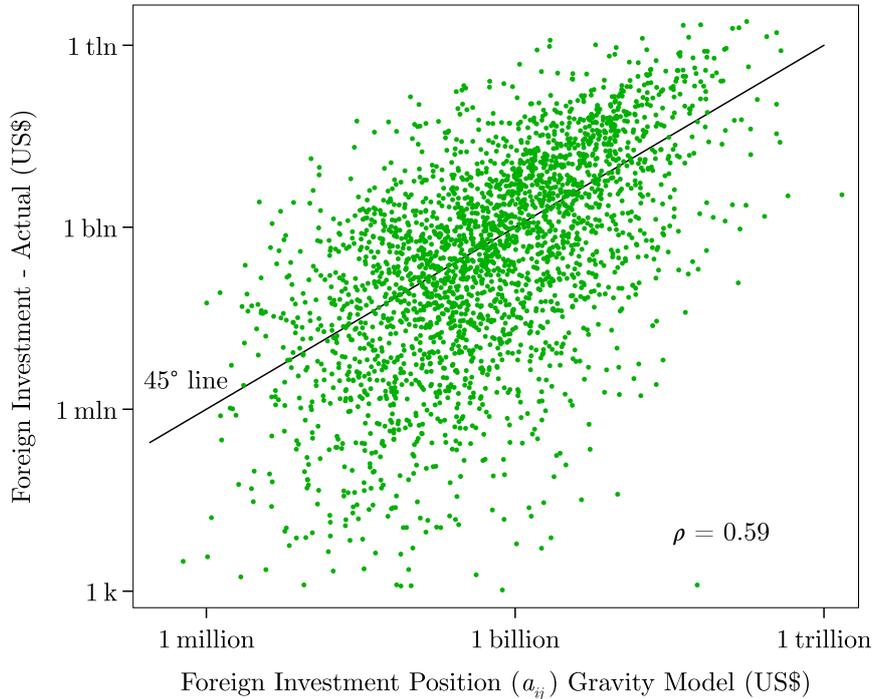


FIGURE NOTES: This figure plots the log of Foreign Total Investment against the log of predicted international investment from the gravity model ( $a_{ij}$ ). The data are for 2017.

obtained as  $\Sigma s$ . Finally, the Cobb-Douglas production function pins down total factor productivity  $\omega_i$  given  $k_i$ ,  $\ell_i$  and  $y_i$ .

## 5.2 Empirical Performance of the Gravity Equation

The first question is how well the gravity equation actually fits the restated data on cross-border investment positions. After calibrating the model, we can use the gravity equation (2.22) to predict international investment by origin and destination-country.

The model-implied foreign investment positions are shown in Figure 2 against the actual data on Foreign Total Investment. Model-implied values are *not* fitted values from the regressions in Section 4. 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.

The gravity equation performs satisfactorily: 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. Hence, we conclude that the calibrated gravity equation predicts cross-sectional variation in total

FIGURE 3: MODEL FIT: CAPITAL STOCK PER EMPLOYEE

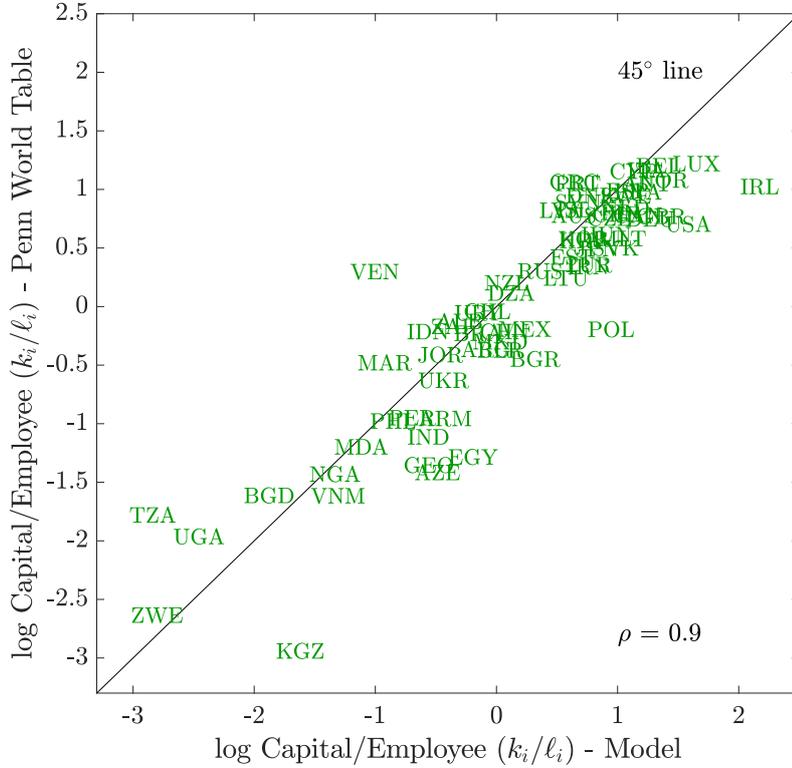


FIGURE NOTES: This figure plots the log of capital stock per employee ( $k_i/l_i$ ) from the against the corresponding model estimate. Both measures are in differences from the mean. Each observation is a country and data refer to the year 2017.

investment quite well.

### 5.3 Model Fit: Country-Level Capital Stocks

We now consider another way to validate our model empirically. When taking the model to the data, we have not used any domestic capital data. To further validate our model, we can compare capital stock per employee from the Penn World Tables to our model-based estimates of the same variable. We take these two variables in logs, subtract the mean, and then plot the two series in a scatter plot in in Figure 3.<sup>18</sup>

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 tightly around the 45-degree line. Thus, our model

<sup>18</sup>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  $\omega_i$  adjusts to preserve the level of GDP).

performs well in terms of predicting the cross-section of capital stocks across the 70 countries we are able to cover. Repeating the exercise without dividing by employment ( $\ell_i$ ) yields a 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 described 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  $\tau_{ij}$ ), and can be related to the resulting GDP loss. As already shown in subsection 2.4, if we set  $\beta = \mathbf{0}$  and  $\mathbf{T} = \mathbf{0}$ , rates of return would be equalized and capital would be efficiently allocated across countries. It is therefore useful, in order to evaluate the empirical performance of the model, to investigate how the rates of returns produced by our model compare to those independently estimated by other researchers.

In a recent paper, David, Henriksen, and Simonovska (2014, henceforth DHS) produce estimates of the return on capital for 144 countries.<sup>19</sup> Of these, 65 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 ( $\rho = 0.57$ ). 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 bilateral metrics of investment taxes as well as geographic and cultural distances. If we were to set  $\beta = 0$  and  $\tau_{ij} = 1$  for all  $(i, j)$ , the upper panel of Figure 4 would display a 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: rates of returns on capital correlate negatively, at the country level, with the level of economic development. 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 -0.76. This is consistent with the observation that rates of return are much higher in developing countries. Therefore, if differentials in such rates were the only driver of international investment, one should expect large flows of capital from richer to poorer country. This raises the important question, originally emphasized by Lucas (1990) and later studied in

---

<sup>19</sup>DHS's measurement of returns to aggregate capital builds on those of (Caselli and Feyrer, 2007) and Gomme, Ravikumar, and Rupert (2011). In their seminal paper, Caselli and Feyrer found that, after accounting for differences in the the relative price of investment goods and after adjusting capital shares for land and other natural resources, capital returns were approximately equal across countries in the year they considered (1996). DHS have confirmed such finding for that specific year, but found large return differentials over longer periods. Similar differences between returns in developed and developing markets have been found by Daly (2010) for the period 1981-2008.

FIGURE 4: MODEL-IMPLIED RATES OF RETURN ON CAPITAL

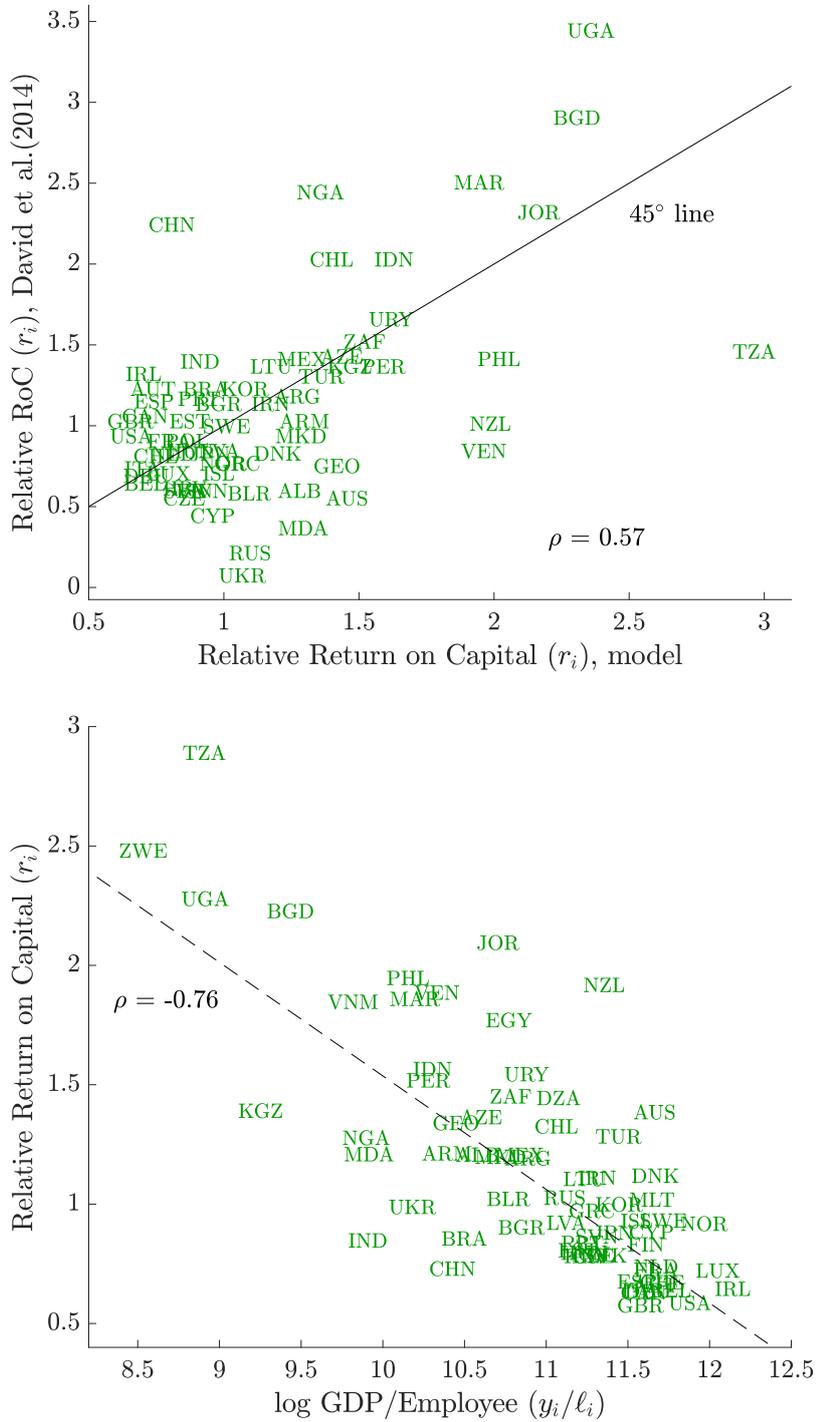


FIGURE NOTES: This figure plots the model-implied rate of return on capital, as a ratio to its median value, against the corresponding estimate from David, Henriksen, and Simonovska (2014, upper panel) and against the log of GDP per employee (lower panel). Each observation is a country and the data are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

depth by Alfaro, Kalemli-Ozcan, and Volosovych (2008), of why we do not observe such flows in the data. The effect of the barriers to global capital allocation, which are at the center of our empirical analysis, can help explain the absence of large movements of capital towards developing countries, thus shedding light on Lucas’s question.<sup>20</sup>

## 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  $\Sigma$  has a significant number of missing observations, which can be filled in using our gravity equation. Among the missing values are the entire diagonal of  $\Sigma$ , which represents domestic investment. 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  $\Sigma$ , based on restated IMF data, with missing values imputed using our gravity model. To avoid having to display 70 countries, the table includes only the G20 plus the rest of the world. Two features of this matrix are of note. The first is the dominance of the United States as a receiving country, which is consistent with its role as a “financial supermarket”, previously documented by Martin and Rey (2004). The second striking pattern is the prominence of the diagonal elements, whose average value is over 50%, pointing to very significant home bias.

Given that our gravity equation loads negatively on measures of cultural and geographic distance, the fact that it predicts at least some degree of home bias is not entirely surprising. What is unexpected is the large size of the home bias predicted by our model: hence, 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 calculate are comparable to those documented by French and Poterba (1991) and (more recently) Coeurdacier and Rey (2013), albeit it should be noted that the measurement approaches are not directly comparable: we focus on a wider range of assets and our data are more recent.<sup>21</sup> Moreover, our data come from a different source.

Another stylized fact that our model is able to account for is that home bias correlates positively with rates of return. This fact was robustly documented by Lau, Ng, and Zhang (2010) who used, as a measure of home bias for country  $i$ , the share of domestic fund holdings in  $i$ ’s stock market capitalization divided by their  $i$ ’s world-market capitalization weight, expressed in logs.

---

<sup>20</sup>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. This is a promising avenue for future research.

<sup>21</sup>Coeurdacier and Rey (2013) document that home bias is slowly weakening.

TABLE 6: MATRIX OF G20 COUNTRIES INVESTMENT SHARES (BASED ON GRAVITY MODEL)

Group	Country	Code	Country of Origin - $j$																RoW		
			AUS	CAN	USA	IND	RUS	TUR	ZAF	ARG	BRA	MEX	DEU	FRA	GBR	ITA	CHN	IDN		JPN	KOR
1	Australia	AUS	44%	0%	0%	0%	0%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
1	Canada	CAN	1%	9%	5%	0%	0%	0%	1%	2%	1%	3%	1%	2%	1%	0%	0%	0%	0%	0%	1%
1	United States	USA	12%	71%	82%	1%	3%	4%	6%	15%	9%	36%	6%	15%	8%	1%	1%	3%	2%	9%	9%
2	India	IND	3%	0%	0%	70%	5%	10%	14%	1%	1%	1%	1%	1%	3%	2%	10%	2%	5%	6%	6%
2	Russia	RUS	1%	1%	0%	2%	46%	4%	2%	1%	1%	1%	3%	2%	3%	1%	1%	1%	2%	4%	4%
2	Turkey	TUR	0%	0%	0%	2%	2%	41%	2%	1%	1%	0%	1%	1%	2%	0%	1%	0%	0%	3%	3%
2	South Africa	ZAF	1%	0%	0%	0%	0%	0%	27%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%
3	Argentina	ARG	0%	0%	0%	0%	0%	0%	1%	30%	3%	1%	0%	0%	0%	0%	0%	0%	0%	1%	1%
3	Brazil	BRA	1%	1%	1%	0%	0%	1%	5%	17%	61%	3%	0%	0%	1%	0%	0%	0%	0%	2%	2%
3	Mexico	MEX	1%	3%	3%	0%	1%	0%	1%	6%	5%	44%	1%	1%	1%	0%	0%	0%	0%	2%	2%
4	Germany	DEU	1%	2%	1%	1%	4%	3%	2%	2%	1%	1%	37%	11%	11%	0%	0%	1%	1%	8%	8%
4	France	FRA	1%	1%	1%	0%	2%	2%	2%	2%	2%	1%	6%	28%	8%	0%	0%	0%	0%	5%	5%
4	United Kingdom	GBR	2%	4%	2%	1%	2%	2%	3%	2%	1%	1%	9%	11%	26%	0%	0%	1%	1%	6%	6%
4	Italy	ITA	1%	1%	1%	1%	2%	4%	4%	3%	2%	1%	6%	7%	6%	0%	0%	0%	1%	6%	6%
5	China	CHN	12%	1%	1%	9%	10%	4%	5%	1%	1%	1%	3%	2%	3%	89%	8%	20%	37%	7%	7%
5	Indonesia	IDN	3%	0%	0%	2%	0%	1%	2%	0%	0%	0%	0%	0%	0%	0%	69%	1%	1%	1%	1%
5	Japan	JPN	7%	1%	1%	1%	2%	1%	1%	0%	0%	1%	2%	1%	1%	3%	2%	62%	16%	1%	1%
5	South Korea	KOR	2%	0%	0%	1%	1%	1%	1%	0%	0%	0%	1%	0%	1%	2%	1%	5%	27%	1%	1%
6	Rest of the World	RoW	9%	6%	4%	9%	20%	23%	22%	18%	11%	6%	24%	26%	24%	2%	6%	3%	6%	38%	38%

TABLE NOTES: This table presents, for G20 countries (except Saudi Arabia, for which we do not have data) matrix of investment shares  $\Sigma = [\sigma_{ij}]$ , based on the gravity model of Section 2.  $\Sigma$  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 country locations. 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.

FIGURE 5: RATES OF RETURN AND HOME BIAS

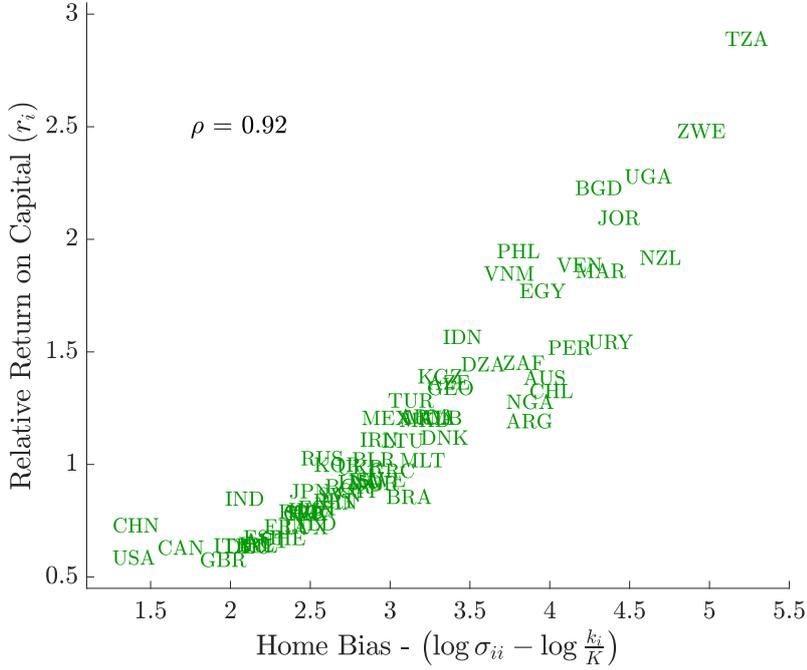


FIGURE NOTES: This figure plots the model-implied rate of return on capital, as a ratio to its median value, against a measure of home bias - the ratio of the shares of country  $i$  investment in country  $i$  portfolio and in the World Portfolio ( $\log \sigma_{ii} - \log \frac{k_i}{K}$ ). Each observation is a country and data refer to the year 2017.

To show that our general equilibrium model is capable of reproducing this correlation, we compute our own model-consistent version of this measure:

$$\text{Home Bias}_i \stackrel{\text{def}}{=} \log \sigma_{ii} - \log \frac{k_i}{K} \quad (5.1)$$

and plot it against the model-implied return to capital ( $r_i$ ) in Figure 5. As visible from the graph, the two correlate strongly and positively.

## 5.6 Counterfactual Analysis

We now move to the core of this section: the counterfactual analysis. Counterfactuals allow us to gain a sense of the economic impact of the barriers. In our setting, simulating a counterfactuals means changing the vector  $\beta$  (the semi-elasticity of investment with respect to the distance vector) and/or the matrix of investment taxes ( $\mathbf{T}$ ), and studying how all other variables (as well as statistics of interest) change in equilibrium as a result. 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 the main results from the counterfactual analysis. In column (1), we present the observed, distorted equilibrium, with investment taxes, geographic distance and cultural distance as observed. In column (2), we present the *zero-gravity* equilibrium, from which all distortions – except those at the level of the individual investors – have been removed ( $\beta = \mathbf{0}$ ,  $\mathbf{T} = \mathbf{0}$ ). In column (3) to (5), we consider equilibria in which the effect of *Cultural Distance*, *Geographic Distance* and investment taxes are respectively re-introduced in isolation (that is, all other distortions except the one indicated are removed). These three 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 is meant to isolate the effect of capital misallocation, and reflects 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 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 satisfied.

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 111.3 US\$ billion. That is 6.9% lower than in the *zero-gravity* counterfactual (column 2) under the assumption of an inelastic capital supply, and 7% lower when we assume that the supply of capital can respond.

We also find that the three distortions have quantitatively meaningful effects on World GDP when considered in isolation. When all distortions except those due to *Cultural Distance* are removed, GDP is 3.5% lower than in the Zero-Gravity scenario; this figure becomes 2.6% if capital is allowed to respond. When *Geographic Distance* alone is considered, the GDP losses are 1.5% (inelastic capital) and 2% (elastic capital). Investment taxes have a more muted impact when considered in isolation: they reduce World GDP by 0.5% when capital supply is inelastic.<sup>22</sup>

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 ( $\theta_i$ ). 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.

---

<sup>22</sup>When we allow capital supply to adjust, we find that the removal of this friction in isolation leads to a slight *increase* in world GDP. This is explained by a reallocation of economic activity towards countries with lower savings rates.

TABLE 7: COUNTERFACTUALS (2017)

	Observed (All Barriers)	Zero-Gravity (No Barriers)	Cultural Distance	Geographic Distance	Political Risk & Taxes
	(1)	(2)	(3)	(4)	(5)
<b>Welfare Statistics (Fixed Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	119.6	115.5	117.8	119.0
World GDP, % Difference in GDP from Zero-Gravity	-6.9%	0%	-3.5%	-1.5%	-0.5%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+78.6%	0%	+36.5%	+35.8%	+14.9%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+44.0%	0%	+18.8%	+15.1%	+8.9%
<b>Welfare Statistics (Endogenous Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	119.6	116.5	117.2	120.4
World GDP, % Difference in GDP from Zero-Gravity	-7.0%	0%	-2.6%	-2.0%	+0.7%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+78.5%	0%	+31.5%	+39.2%	+15.2%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+44.0%	0%	+16.0%	+15.3%	+8.9%

TABLE NOTES: This table presents welfare statistics for four counterfactuals of the model described in Section 2. Each of columns (2)-(5) is a counterfactual, and the rows represent different welfare statistics of interest. *Observed* is the equilibrium allocation with all measured barriers. *Zero-Gravity* is the counterfactual in which all three barriers (*Cultural Distance*, *Geographic Distance* and *Political Risk & Taxes*) have been removed. Columns (3)-(5) illustrate three counterfactuals from which only the corresponding distortion is in place. ( $k_i/\ell_i$ ) is the capital stock per employee, while ( $y_i/\ell_i$ ) is output (GDP) per employee. Actual World PPP\$ GDP (including countries not in the model) in 2017 was \$121 trillion.

FIGURE 6: DISTRIBUTION OF CAPITAL AND OUTPUT PER EMPLOYEE

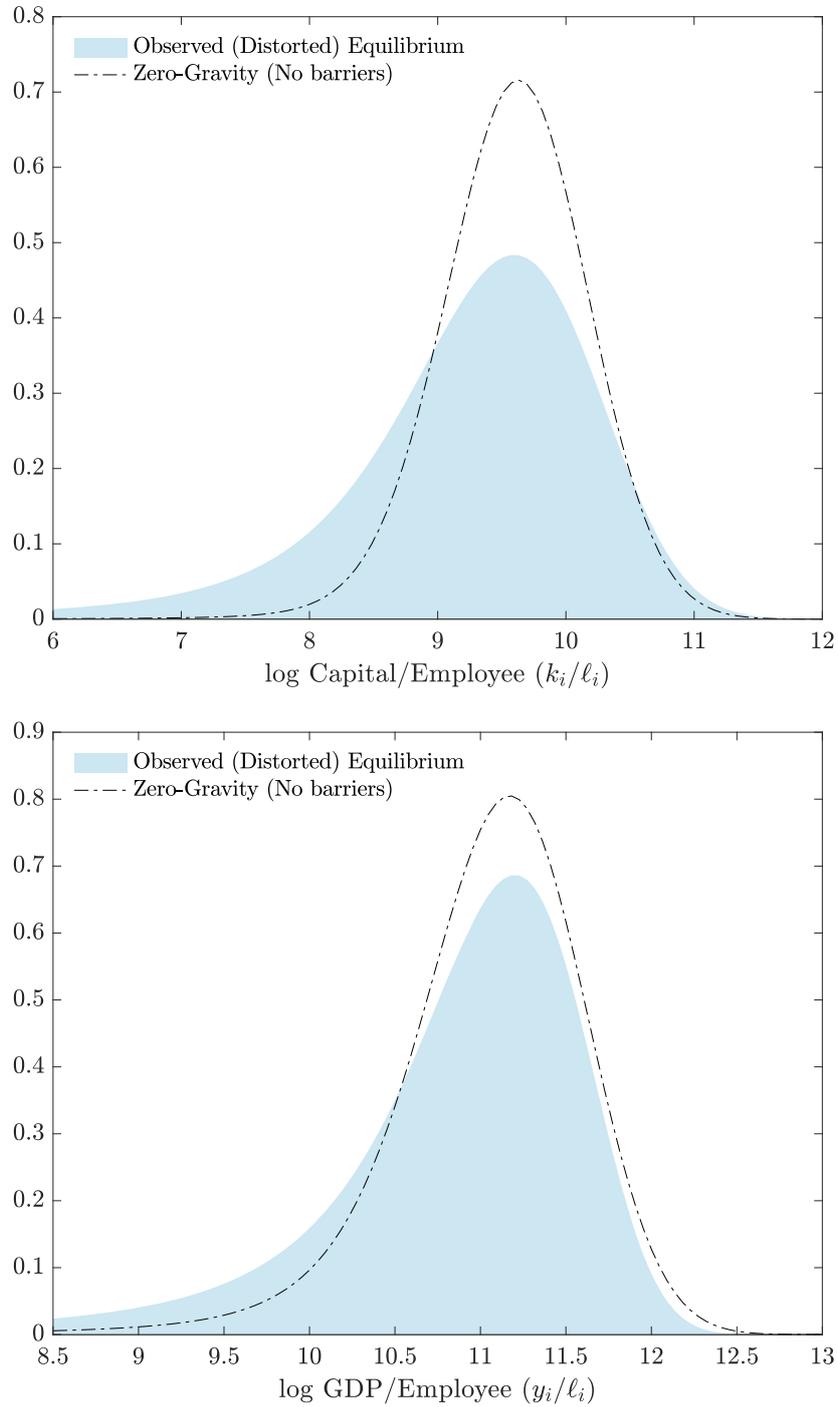


FIGURE NOTES: This figure fits the probability density function of a *stable* distribution (a 4-parameter family of distributions with flexible skewness and fat tails) to country-level capital stock per employee (upper panel) and GDP per employee (bottom panel). In each panel, the lighter area is the 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 and cultural distances, as well as investment taxes) have been removed.

While the overall effects of these three distortions on allocative efficiency and World GDP appear substantial, their effect on cross-country inequality is 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 employee increases by 78.6%%, while the standard deviation of log output per employee increases by 44%. These numbers are not sensitive to whether capital supply is assumed to be inelastic.

When *Cultural Distance* alone is maintained, dispersion in log capital per employee is 31.5%-36.5% higher than in the zero-gravity benchmark (depending on whether capital supply is elastic or not). The dispersion of log output per employee is 16%-18.8% higher. For *Geographic Distance*, the magnitudes are commensurate with these numbers. Finally, we find that maintaining only investment taxes, dispersion in log capital per employee is about 15% higher compared to the zero-gravity benchmark, while dispersion in log output per employee is 8.9% higher. In other words, cultural distance appears to have the largest marginal effect, followed by geographic distance. Investment taxes have smaller, but still sizable effects on cross-country inequality.

Figure 6 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 per employee 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.

What lies beyond this reduction in inequality under zero-gravity? 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 discussed previously, these tend to be countries with lower capital stock per employee and lower output per employee. Figure 7 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 implies that the removal of barriers leads to a substantial reduction in cross-country inequality. In other words, poorer countries 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, we consider a comparison of net foreign asset positions (defined here as the difference between



FIGURE 8: NET FOREIGN ASSETS AND DEVELOPMENT

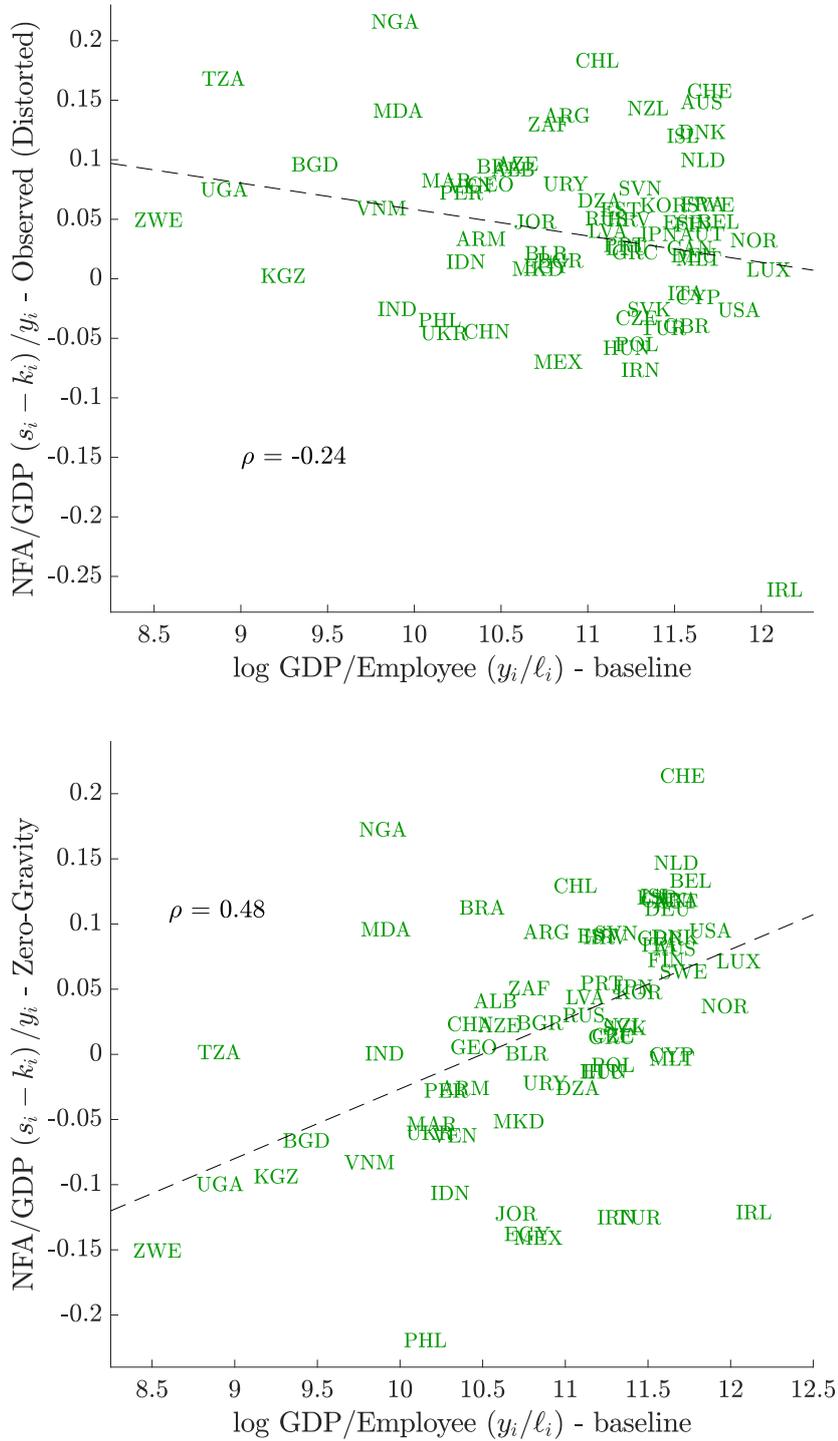


FIGURE NOTES: the figure above plots the model-implied Net Foreign Assets ( $s_i - k_i$ ) as a share of GDP ( $y_i$ ), against the log of GDP per employee. The top panel plots NFA/GDP in the observed distorted equilibrium, while the bottom panel plots NFA/GDP in the Zero-Gravity counterfactual, in which all barriers are removed.

## 6 Robustness Checks and Extensions

### 6.1 Coefficients Stability

How stable are the coefficient estimates on *Cultural Distance* and *Geographic Distance* over time? Appendix B, Figure B.1 plots coefficient estimates from a variation of our baseline regression specification (Table 2, column 1). The dependent variable is still the log of foreign total investment, but the right hand side variables (*Cultural Distance*, *Geographic Distance*) are interacted with year fixed effects to produce time-varying coefficients. The 95% confidence interval is plotted together with the calibrated coefficients (dotted line). The dotted line always falls within the confidence interval, and close to its center for both variables. This time-stability of the main regression estimates of interest provides evidence that our choice of calibrated effects of cultural and geographic distance is well-founded.

### 6.2 Alternative Breakdown of Foreign Investment Statistics

In our main estimation, we broke down FTI into debt and equity components. Here we consider instead another conventional breakdown of capital flows: between Foreign Direct Investment (FDI) and Foreign Portfolio Investment (FPI). Appendix C, Table C.1 presents the results, using the same specification as that of Table 2. We find that cultural and geographic distances exert negative, statistically significant and economically meaningful negative effects on FDI and FPI, whether one does not include additional controls (columns 2 and 3) or whether one includes them (columns 5 and 6).

### 6.3 Restated vs. Un-restated Data

In our main estimation exercise, we use foreign investment data that are restated to account for the effect of tax havens. Appendix D, Table D.1 replicates the regressions of Table 2 using non-restated (residency-based) data on foreign total investment, foreign debt investment and foreign equity investment. The sample involves a substantially larger number of observations, especially when no control variables are added (columns 1-3). Nonetheless, the standardized magnitudes of the estimates are very close to those obtained from Table 2. Across the six specifications, with non-restated data the effect of cultural distance is slightly smaller, while that of geographic distance is slightly larger than when using data that are restated on a nationality basis.

### 6.4 Capital Controls

One type of barrier that we have deliberately omitted from our model is capital account policy restrictions. We did so because our model is not designed to address questions of macro-prudential policy, i.e. short-term considerations about macroeconomic stability (we focus instead on a long-run steady-state).

Nonetheless, capital controls are enacted in order to affect capital flows, and thus we worry whether their effect may interact with that of our variables in a way that might change our results in a meaningful way.

A simple way to theoretically model the effect of capital controls is to amend the equation describing the subjective  $(x, z)$  return  $R$  as follows:

$$R(x, z) \stackrel{\text{def}}{=} \phi_{ij} \cdot (1 - \tau_{ij}) \cdot r_i \cdot e^{-\lambda(x, z)} \quad (6.1)$$

$\phi_{ij}$  is defined over the interval  $[0, 1]$ : it captures the degree of capital account openness (the lack of capital controls) facing  $j$ -investors seeking to invest in country  $i$ .  $\phi_{ij} = 1$  implies that investment from  $j$  to  $i$  is unrestricted. For domestic investors ( $i = j$ )  $\phi_{ij}$  is always 1 by definition. This leads to the following modified equation for the portfolio shares:

$$\sigma_{ij} = \frac{\phi_{ij} \cdot (1 - \tau_{ij}) \cdot r_i k_i \cdot \exp(\mathbf{d}'_{ij} \beta)}{\sum_{c=1}^n \phi_{cj} \cdot (1 - \tau_{cj}) \cdot r_c k_c \cdot \exp(\mathbf{d}'_{cj} \beta)} \quad (6.2)$$

Turning to the empirical implementation, we measure the degree of *de jure* capital account openness between country  $i$  and country  $j$  using data from Jahan and Wang (2016), which is based on qualitative information from the IMF's Annual Report on Exchange Arrangements and Exchange Restrictions (AREAER). Their dataset consists in a set of dummy variables that encode the presence of inflow or outflow capital account restrictions on specific types of investments. We use data for the most recent year in their dataset, 2013.

For each country in their dataset, we use the following set of ten dummy variables. The first two dummies represent, respectively, restrictions on inflowing and outflowing direct investment. The next set of four dummies represent restrictions on equity portfolio investment: two represent restrictions on (respectively) the sale and purchase of domestic equity by non-residents, while the other two represent restrictions on (respectively) the sale and purchase of foreign equity by residents. The third and last set of four dummies covers restrictions on debt portfolio investment: two reflect, respectively, restrictions on sale and purchase of domestic debt securities by non-residents; the other two represent restrictions on sale and purchase of foreign debt securities by residents.

To estimate  $\phi_{ij}$ , we consider the five inflow restrictions dummies for country  $i$  (1 for FDI, 2 for portfolio equity and 2 for debt) as well as the five outflow restrictions dummies for country  $i$  (1 for FDI, 2 for portfolio equity and 2 for debt). We model each restriction as a (shadow) tax on foreign investment. To be conservative in our analysis, we assume that these taxes take on large values. In particular, we assume that both dummies corresponding to FDI investment restrictions are equivalent to a tax of 50%, while each of the four dummy variables corresponding to FPI restrictions is equivalent to a tax of 25%. We first compute the total tax (compounding both  $i$  inflow restrictions and  $j$  outflow restrictions) at the asset class level (FDI, equity portfolio and debt), and then take the simple average across the three asset

classes. Formally:

$$\phi_{ij} = \frac{1}{3} \left[ (1 - 50\%)^{N_{\text{FDI}}} + (1 - 25\%)^{N_{\text{Equity FPI}}} + (1 - 25\%)^{N_{\text{Debt}}} \right] \quad (6.3)$$

where  $N$  indicates the number of dummy variables in AREAER for each asset class/country pair (two, four and four respectively).

In Appendix E, Table E.1 we repeat our counterfactual analysis for the extended model with capital controls. To make the exercise symmetric to the baseline one, capital controls remain in place in each of the five scenarios being studied – that is, we do not remove the effect of capital controls in any of the estimated counterfactuals, not even when we move to “zero-gravity”. The percentage world GDP losses is unchanged with respect to our baseline exercise when the capital stock is fixed (6.9%). When the world capital stock is endogenous the world GDP loss 6.6% – very close to our baseline estimate of 7%. The marginal effects of the individual frictions are also very similar to the one found in the baseline exercise.

## 6.5 Currency Hedging Costs and Risk Premia

Another aspect of international investment that we have left out of the model is currency risk. In our basic model, there is no explicit notion of money. However, there is a tractable way to incorporate currency risk in our framework. We start from the observation that the vast majority of international investors hedge currency risk. Sialm and Zhu (2020) find that over 90% of US-based international fixed income funds hedge currency risk with derivatives. A similar stylized fact holds for equity investments. According to the EU-EFIGE survey (a survey of 15,000 manufacturing firms from the EU and the UK), about two-thirds of the firms engaging in foreign direct investment are hedged against currency risk, either through derivatives or because the foreign subsidiaries invoices in the same currencies as their parent company. This percentage rises to 85% when responses are weighted by firm employment size.

Based on these facts, a parsimonious way to incorporate currencies in our theory is to model the currency hedging cost directly. An agent investing from country  $j$  to country  $i$  that hedges with forward contracts will exchange  $j$  currency for  $i$  currency at a spot exchange rate, and will then repatriate their investment return at the forward rate. This implies that the investor is subjected to a multiplicative cost (or gain) equal to the forward premium on the  $j/i$  exchange rate.

Thus, a simple way to introduce this hedging cost in our model (without modeling currency risk explicitly) is to add an additional component in the international investment fee  $\lambda(x, z)$  paid to the intermediating agent based in  $j$ 's domestic country:

$$\lambda(x, z) = \mathbf{d}'_{ij} |\beta| + \gamma_{ij} - \xi(x, z) - G_j \quad (6.4)$$

Here  $\gamma_{ij}$  represents the currency hedging cost, which we estimate using exchange rate forward premia.

Let  $\bar{r}_{ij}$  be the country  $i$  objective return, hedged in country  $j$  currency:

$$\bar{r}_{ij} \stackrel{\text{def}}{=} r_i \cdot e^{-\gamma_{ij}} \quad (6.5)$$

We then have the following amended gravity equation for the  $(i, j)$  portfolio share:

$$\sigma_{ij} = \frac{(1 - \tau_{ij}) \cdot \bar{r}_{ij} k_i \cdot \exp(\mathbf{d}'_{ij} \beta)}{\sum_{c=1}^n (1 - \tau_{cj}) \cdot \bar{r}_{cj} k_c \cdot \exp(\mathbf{d}'_{cj} \beta)} \quad (6.6)$$

An intuitive aspect of re-defining country returns this way is that  $\bar{r}$  can also be thought of as a risk-adjusted return. This is because, if *covered* interest parity holds, the forward premium is given by the risk-free return differential between in country  $i$  and  $j$ . This is in turn related, by the fundamental exchange rate valuation equation (Campbell and Clarida, 1987; Froot and Ramadorai, 2005), to the risk premium on  $i$ 's currency from the point of view of a  $j$  investor. In other words, the cost of hedging a high-yielding currency is equal to the forgone currency risk premium, and this allows us to interpret  $\bar{r}_{ij}$  as the foreign investment return, adjusted for currency risk.

In order to incorporate currency hedging costs in our model, we use forward premia on exchange rates. We download, from Bloomberg, spot as well as 1-year forward rates against the US dollar for the currencies of the countries covered in our foreign investment data. We use the earliest available data for the year 2017. We then compute cross-rates for all available currency pairs. For currency pairs where data is unavailable or unreliable, we estimate forward premia using the covered interest rate parity. We use 1-year interbank rate differentials where available, and central bank rates differentials as a last resort. Interest rate data (for January 2017) is sourced from TradingEconomics.com. Finally, we fill the data for the remaining currency pairs (which are missing because the forward premia were estimated using different methodologies) using a linear regression of log forward premia on  $i$  and  $j$  fixed effects. For pegged currency pairs, we assume that the currency hedging cost is zero.

In Appendix E, Table E.2 we repeat our counterfactual analysis for the extended model with currency hedging costs. As in the previous robustness exercise with capital controls, currency hedging costs remain in place throughout the four scenarios. The world GDP loss that we find according to this extended model is larger, but only slightly: -7.2% when the capital stock is fixed and -7.5% when the capital stock is endogenous. As for the capital controls extension, the marginal effect of the three set of barriers remains very close to the baseline level.

## 6.6 Risky Country Returns

A third factor ignored thus far is the intrinsic riskiness of the investment (both domestic and foreign) undertaken. Our third extension aims to tractably introduce a notion of risk in our otherwise non-

stochastic model. To do that, we introduce risk in the conversion of final good at time  $t$  into productive capital at time  $t + 1$ . In particular, each unit of final good saved by  $z$  becomes  $\exp \zeta(x, z)$  units of productive capital in the next period. The shock  $\zeta$ , unlike  $\xi$ , is not known in advance to the investors, and is assumed to be conditionally Gaussian, with mean  $-v_i^2/2$  and standard deviation  $v_i$ , so that there is no change in the average to the capital supply in the destination country. Moreover, because investors are atomistic, shocks average out and there is no aggregate uncertainty. Next period consumption nonetheless stochastic from the point of view of investor  $z$ , whose lifetime utility is now amended to be:

$$U(z) = (1 - \alpha) \cdot \log c_t(z) + \alpha \cdot \mathbf{E}[\log c_{t+1}(z)] \quad (6.7)$$

Because agent  $z$  holds log sub-utility, they choose the plant with the highest expected log return, which can be written as:

$$\mathbf{E}[\log R_{t+1}(z)] \stackrel{\text{def}}{=} (1 - \tau_{ij}) \cdot r_i \cdot \exp \left[ -\lambda(x, z) - \frac{1}{2} v_i^2 \right] \quad (6.8)$$

We can then re-write  $z$ 's portfolio optimization problem as:

$$x^*(z) = \underset{x}{\operatorname{argmax}} \left[ \log(1 - \tau_{ij(x,z)}) + \log r_{i(x)} - \frac{v_{i(x)}^2}{2} + \mathbf{d}'_{ij(x,z)} \beta + \xi(x, z) \right] \quad (6.9)$$

This leads to the following amended expression for the portfolio shares:

$$\sigma_{ij} = \frac{(1 - \tau_{ij}) \cdot \theta_i y_i \cdot \exp \left( \mathbf{d}'_{ij} \beta - \frac{1}{2} v_i^2 \right)}{\sum_{c=1}^n (1 - \tau_{cj}) \cdot \theta_c y_c \cdot \exp \left( \mathbf{d}'_{cj} \beta - \frac{1}{2} v_c^2 \right)} \quad (6.10)$$

This treatment of investment risk is simplified, in that the consumer's discount factor is non-stochastic, there is no aggregate risk, and hedging motives are assumed away. However, an advantage of this simplification is that it only requires us to find an estimate of  $v_i$ . To estimate the country-level variance of the asset returns ( $v_i^2$ ), we download country equity volatility indices from FRED (the original source of the data is Bloomberg). For the few (emerging) countries for which this is unavailable, we use the CBOE Emerging Markets ETF Volatility Index as a proxy.

In Appendix E, Table E.3 we repeat our counterfactual analysis for the extended model with risky investment. As for the previous robustness exercises with capital controls and currency hedging costs, the effect of risk on portfolio allocations remains in place in all four scenarios being considered. Our results are virtually unchanged when we account for risky returns.

## 6.7 Sensitivity Analysis on Coefficient Estimates

It is reasonable to ask how the results of our counterfactual analysis would change if we were to utilize our Pseudo-Poisson estimates, or our Instrumental Variables (IV) estimates, to calibrate  $\beta$  - the semi-elasticity of foreign investment with respect to cultural and geographic distance.

We address this question in Appendix F, Tables F.1-F.2. There we present the analysis of Table 7, using these alternative estimates for  $\beta$ . We find that the GDP loss induced by capital misallocation would be 4.6% if we were to use the Pseudo-Poisson coefficients, and 9.2% if we were to use the IV estimates. The variable whose marginal effect is most sensitive to this choice is *Cultural Distance*. We continue to find that the removal of barriers would result in significant reductions in world inequality under both Poisson and IV estimates.

## 7 Conclusion

In this paper, we focused on how barriers to international financial flows affect the efficiency of capital allocation across countries. To study such barriers, we 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 on factors that hinder the flows of information across societies and affect intermediation costs, reducing the return from investing in distant markets. These costs result from geographic and cultural barriers to capital allocation.

We estimated our gravity model empirically, using foreign investment data that have been restated from a residency to a nationality basis, in order to account for the presence of offshore investment and financing vehicles (Coppola et al., 2020; Damgaard et al., 2019). Using a variety of estimation approaches (OLS, Poisson, IV), we found that geographic and cultural 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 cultural and geographic barriers, foreign investment taxes and political risk - explains a significant share of the observed variation in foreign investment.

Our model reproduces several features of international asset markets. First, it produces large and meaningful variation in rates of returns across countries. These rates of returns correlate negatively with the level of economic development, a realistic feature of the model. 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 geographic and cultural factors are also likely to play an important role.

To quantify the influence of these factors on the international allocation of capital and their real

impact, we performed a number of counterfactual exercises using our model. We studied how World GDP and the cross-country distribution of capital and output per worker would change if the effects of geographic and cultural distances on foreign investment were zeroed out. This quantitative exercise suggests that capital misallocation associated with barriers to the global allocation of capital has a sizable impact on the distribution of capital across countries, in terms of efficiency as well as inequality. World GDP is 7% lower than it would be if the effect of these barriers could be neutralized. The cross-country standard deviation of capital per employee is 80% higher, while the dispersion of output per employee is 42% higher. Conversely, the hypothetical removal of 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). At the same time, 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 end by suggesting directions for future research. Our work departs significantly from the existing small open economy macro-financial literature: it focuses on geo-cultural and policy frictions, as opposed to diversification and hedging, as potential drivers of international capital allocation. 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 a richer specification of the goods sector, which we deliberately chose to keep simple in our model. A fully-integrated model for international trade and investment would make it possible to study the combined effect of asset market imperfections and of goods markets distortions, accounting for possible ways in which the two interact with each other.

Finally, another direction for future research would be to enrich the model along the dimension of consumer/investor heterogeneity. In our framework, all agents have the same saving rate. An interesting question, which could be investigated with a richer household sector, is how financial globalization might affect *within*-country income inequality.

In 1990, Robert Lucas asked: “Why doesn’t capital flow from rich to poor countries?”. This paper sheds new light on this question. Geographic, cultural and policy barriers are important determinants of cross-country investment positions. They have major effects on efficiency and distribution, including hindering the flow of capital from richer to poorer societies.

## References

- ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT, AND R. WACZIARG (2003): “Fractionalization,” *Journal of Economic Growth*, 8, 155–194.
- ALFARO, L. AND J. CHAUVIN (2020): *Encyclopedia of International Economics and Global Trade, Vol. 1: Foreign Direct Investment and the Multinational Enterprise*, World Scientific, chap. Foreign Direct Investment, Finance and Economic Development, 231–258.
- 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.
- BAQAEI, D. R. AND E. FARHI (2020): “Productivity and misallocation in general equilibrium,” *The Quarterly Journal of Economics*, 135, 105–163.
- BLANCHARD, O. AND J. ACALIN (2016): “What Does Measured FDI Actually Measure?” *Peterson Institute for International Economics Policy Brief*.
- 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.
- BRONER, F. AND J. VENTURA (2016): “Rethinking the Effects of Financial Globalization,” *The Quarterly Journal of Economics*, 131, 1497–1542.
- BURCHARDI, K. B., T. CHANEY, AND T. A. HASSAN (2019): “Migrants, ancestors, and foreign investments,” *The Review of Economic Studies*, 86, 1448–1486.
- CAMPBELL, J. Y. AND R. H. CLARIDA (1987): “The dollar and real interest rates,” in *Carnegie-Rochester Conference Series on Public Policy*, Elsevier, vol. 27, 103–139.
- CASELLI, F. AND J. FEYRER (2007): “The Marginal Product of Capital,” *The Quarterly Journal of Economics*, 122, 535–568.
- COEURDACIER, N. AND H. REY (2013): “Home bias in open economy financial macroeconomics,” *Journal of Economic Literature*, 51, 63–115.
- COLACITO, R. AND M. M. CROCE (2010): “The short and long run benefits of financial integration,” *American Economic Review*, 100, 527–31.
- 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. JOHANNESSEN (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*.
- FRENCH, K. R. AND J. M. POTERBA (1991): “Investor Diversification and International Equity Markets,” *American Economic Review*, 81, 222–226.
- FROOT, K. A. AND T. RAMADORAI (2005): “Currency returns, intrinsic value, and institutional-investor flows,” *The Journal of Finance*, 60, 1535–1566.
- GHOSH, S. AND H. WOLF (1999): “The geography of capital flows,” *Capital Inflows to Emerging Markets*, University of Chicago Press, Chicago.
- GOMME, P., B. RAVIKUMAR, AND P. RUPERT (2011): “The return to capital and the business cycle,” *Review of Economic Dynamics*, 14, 262–278.
- 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. AND O. JEANNE (2006): “The Elusive Gains from International Financial Integration,” *The Review of Economic Studies*, 73, 715–741.

- (2013): “Capital flows to developing countries: The allocation puzzle,” *The Review of Economic Studies*, 80, 1484–1515.
- GUIO, L., P. SAPIENZA, AND L. ZINGALES (2009): “Cultural biases in economic exchange,” *The 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,” *Annual Review of Economics*, 6, 735–770.
- HOXHA, I., S. KALEMLI-OZCAN, AND D. VOLLRATH (2013): “How big are the gains from international financial integration?” *Journal of Development Economics*, 103, 90–98.
- ILZETZKI, E., C. M. REINHART, AND K. S. ROGOFF (2019): “Exchange arrangements entering the twenty-first century: Which anchor will hold?” *The Quarterly Journal of Economics*, 134, 599–646.
- JAHAN, M. S. AND D. WANG (2016): “Capital account openness in low-income developing countries: Evidence from a new database,” *International Monetary Fund Working Papers*.
- JIN, K. (2012): “Industrial Structure and Capital Flows,” *American Economic Review*, 102, 2111–2146.
- 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.
- LAU, S. T., L. NG, AND B. ZHANG (2010): “The world price of home bias,” *Journal of Financial Economics*, 97, 191–217.
- 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?” *American Economic Review*, 80, 92–96.
- 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.
- MENDOZA, E. G., V. QUADRINI, AND J.-V. RIOS-RULL (2009): “Financial integration, financial development, and global imbalances,” *Journal of Political economy*, 117, 371–416.
- NOVY, D. (2013): “Gravity redux: measuring international trade costs with panel data,” *Economic Inquiry*, 51, 101–121.
- OKAWA, Y. AND E. VAN WINCOOP (2012): “Gravity in international finance,” *Journal of International Economics*, 87, 205–215.
- 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.
- 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.
- SELLIN, P. AND I. M. WERNER (1993): “International investment barriers in general equilibrium,” *Journal of International Economics*, 34, 137–151.
- SIALM, C. AND Q. ZHU (2020): “Currency Management by International Fixed Income Mutual Funds,” *Available at SSRN*.
- 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.

# Barriers to Global Capital Allocation - ONLINE APPENDIX

Bruno Pellegrino, Enrico Spolaore, Romain Wacziarg

## 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* (Table 4). 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  $i = 1, 2, \dots, P^1$  inherits the cultural traits of its ancestral population plus a shock  $\epsilon_i^1$ , so that

$$C_i^1 = C^0 + \epsilon_i^1 \quad (\text{A.1})$$

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  $i = 1, 2, \dots, P^2$  inherits the culture of its parental population the previous period plus a random shock  $\epsilon_i^2$ . 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:

$$C_i^2 = C^0 + \epsilon_{a_{12}(i)}^1 + \epsilon_i^2 \quad (\text{A.2})$$

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

$$C_i^t = C^0 + \sum_{m=1}^t \epsilon_{a_{mt}(i)}^m \quad (\text{A.3})$$

Let  $d_C^t(i, j) \equiv |C_i^t - C_j^t|$  denote the cultural distance between population  $i$  and population  $j$  at time  $t$ . Let  $d_A^t(i, j)$  denote the ancestral distance 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

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

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 proposition as:

$$n_{ij} > n_{kl} \iff E[|C_i^T - C_j^T|] > E[|C_k^T - C_l^T|] \quad (\text{A.5})$$

By definition:

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

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

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

By the same token, we have:

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

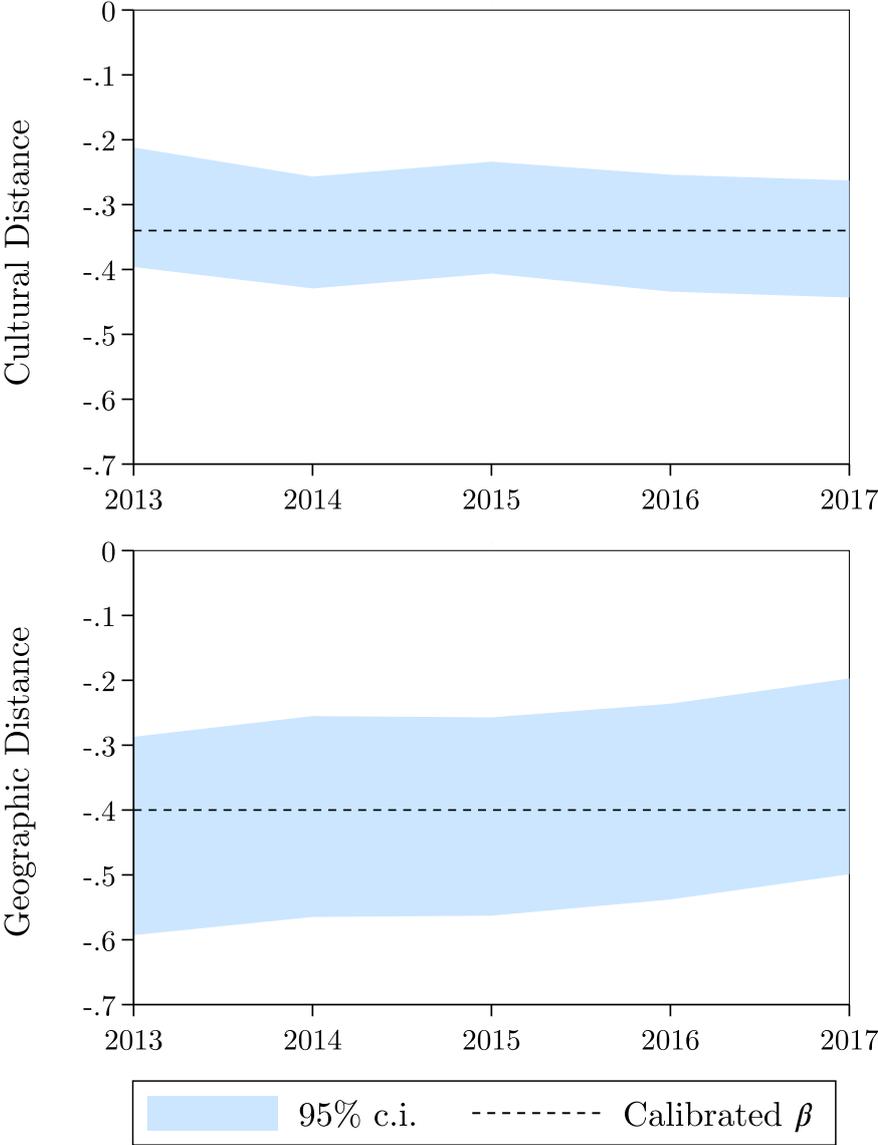
Thus, our claim follows if we can show that:

$$n_{ij} > n_{kl} \iff E\left[\left|\sum_{q=1}^{n_{ij}} (\eta_q - \nu_q)\right|\right] > E\left[\left|\sum_{q=1}^{n_{kl}} (\eta_q - \nu_q)\right|\right] \quad (\text{A.9})$$

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 Regression Coefficients Stability

FIGURE B.1: COEFFICIENTS STABILITY OVER TIME



## C Robustness Check: alternative Breakdown of Foreign Investment

TABLE C.1: OLS REGRESSIONS USING FDI/FPI BREAKDOWN INSTEAD OF EQUITY/DEBT

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dep.variable in logs:</i>	Assets	FDI	FPI	Assets	FDI	FPI
Cultural Distance	-0.419** (0.038)	-0.456** (0.037)	-0.295** (0.034)	-0.333** (0.041)	-0.304** (0.044)	-0.216** (0.035)
Geographic Distance	-0.526** (0.031)	-0.589** (0.035)	-0.383** (0.027)	-0.398** (0.076)	-0.579** (0.098)	-0.299** (0.059)
Control Variables	No	No	No	Yes	Yes	Yes
Observations	12,542	13,230	23,499	11,157	11,964	20,988
<i>R</i> -squared	0.771	0.710	0.826	0.803	0.748	0.847
Within <i>R</i> -squared	0.214	0.222	0.135	0.327	0.301	0.256

## D Robustness Check: Residency-Based Foreign Investment Data

TABLE D.1: OLS REGRESSIONS USING UN-RESTATED DATA

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dep.variable in logs:</i>	Assets	Equity	Debt	Assets	Equity	Debt
Cultural Distance	-0.397** (0.038)	-0.432** (0.042)	-0.424** (0.036)	-0.283** (0.042)	-0.299** (0.046)	-0.309** (0.037)
Geographic Distance	-0.616** (0.031)	-0.620** (0.032)	-0.408** (0.026)	-0.597** (0.089)	-0.604** (0.091)	-0.460** (0.077)
Control Variables	No	No	No	Yes	Yes	Yes
Observations	20,254	19,687	18,068	11,157	11,964	20,988
<i>R</i> -squared	0.761	0.728	0.777	0.803	0.748	0.847
Within <i>R</i> -squared	0.218	0.202	0.182	0.327	0.301	0.256

## E Counterfactual Analysis using Model Extensions

The following tables replicate Table 7 for the three model extensions presented in Section 6: Capital Controls, Currency Hedging Costs, and Risky Asset Returns.

TABLE E.1: COUNTERFACTUALS WITH CAPITAL CONTROLS (2017)

	Observed (All Barriers)	Zero-Gravity (No Barriers)	Cultural Distance	Geographic Distance	Political Risk & Taxes
	(1)	(2)	(3)	(4)	(5)
<b>Welfare Statistics (Fixed Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	119.6	114.5	116.6	118.2
World GDP, % Difference in GDP from Zero-Gravity	-6.9%	0%	-4.2%	-2.5%	-1.1%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+53.8%	0%	+27.5%	+24.1%	+13.4%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+36.3%	0%	+17.3%	+13.2%	+8.3%
<b>Welfare Statistics (Endogenous Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	119.2	115.5	116.9	119.6
World GDP, % Difference in GDP from Zero-Gravity	-6.6%	0%	-3.1%	-2.0%	+0.3%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+52.0%	0%	+25.5%	+28.3%	+13.2%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+35.2%	0%	+16.4%	+14.5%	+8.2%

TABLE E.2: COUNTERFACTUALS WITH CURRENCY HEDGING COSTS (2017)

	Observed (All Barriers)	Zero-Gravity (No Barriers)	Cultural Distance	Geographic Distance	Political Risk & Taxes
	(1)	(2)	(3)	(4)	(5)
<b>Welfare Statistics (Fixed Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	120.0	115.7	118.0	119.1
World GDP, % Difference in GDP from Zero-Gravity	-7.2%	0%	-3.6%	-1.6%	-0.7%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+78.2%	0%	+36.2%	+36.4%	+15.8%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+43.5%	0%	+18.5%	+15.2%	+8.9%
<b>Welfare Statistics (Endogenous Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	120.3	117.0	117.8	120.8
World GDP, % Difference in GDP from Zero-Gravity	-7.5%	0%	-2.8%	-2.1%	+0.4%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+78.2%	0%	+31.3%	+40.3%	+16.0%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+43.5%	0%	+15.7%	+15.5%	+8.9%

TABLE E.3: COUNTERFACTUALS WITH RISKY ASSET RETURNS (2017)

	Observed (All Barriers)	Zero-Gravity (No Barriers)	Cultural Distance	Geographic Distance	Political Risk & Taxes
	(1)	(2)	(3)	(4)	(5)
<b>Welfare Statistics (Fixed Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	119.6	115.4	117.8	119.0
World GDP, % Difference in GDP from Zero-Gravity	-6.9%	0%	-3.5%	-1.5%	-0.6%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+78.3%	0%	+36.4%	+35.8%	+15.0%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+44.0%	0%	+18.8%	+15.1%	+8.9%
<b>Welfare Statistics (Endogenous Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	119.7	116.5	117.3	120.4
World GDP, % Difference in GDP from Zero-Gravity	-7.0%	0%	-2.7%	-2.0%	+0.6%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+78.3%	0%	+31.5%	+39.3%	+15.2%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+44.0%	0%	+16.0%	+15.3%	+8.9%

## F Counterfactual Analysis with Alternate Coefficient Estimates

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

TABLE F.1: COUNTERFACTUALS USING POISSON REGRESSION ESTIMATES (2017)

	Observed (All Barriers)	Zero-Gravity (No Barriers)	Cultural Distance	Geographic Distance	Political Risk & Taxes
	(1)	(2)	(3)	(4)	(5)
<b>Welfare Statistics (Fixed Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	116.6	115.4	115.0	116.1
World GDP, % Difference in GDP from Zero-Gravity	-4.6%	0%	-1.1%	-1.4%	-0.5%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+75.6%	0%	+21.0%	+35.1%	+19.5%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+40.3%	0%	+11.8%	+13.8%	+10.8%
<b>Welfare Statistics (Endogenous Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	116.5	116.2	114.3	117.1
World GDP, % Difference in GDP from Zero-Gravity	-4.5%	0%	-0.3%	-1.9%	+0.5%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+75.6%	0%	+17.8%	+34.8%	+19.5%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+40.3%	0%	+9.8%	+12.5%	+10.6%

TABLE F.2: COUNTERFACTUALS USING IV ESTIMATES (2017)

	Observed (All Barriers)	Zero-Gravity (No Barriers)	Cultural Distance	Geographic Distance	Political Risk & Taxes
	(1)	(2)	(3)	(4)	(5)
<b>Welfare Statistics (Fixed Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	122.6	112.6	120.6	121.9
World GDP, % Difference in GDP from Zero-Gravity	-9.2%	0%	-8.2%	-1.6%	-0.6%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+53.7%	0%	+53.6%	+33.3%	+10.5%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+26.9%	0%	+26.1%	+15.3%	+6.7%
<b>Welfare Statistics (Endogenous Capital Stock)</b>					
World GDP (US\$ trillions)	111.3	121.6	109.7	119.5	122.7
World GDP, % Difference in GDP from Zero-Gravity	-8.4%	0%	-9.7%	-1.7%	+0.9%
St.Dev. of $\log(k_i/\ell_i)$ , % Difference from Zero-Gravity	+54.0%	0%	+82.7%	+41.9%	+11.4%
St.Dev. of $\log(y_i/\ell_i)$ , % Difference from Zero-Gravity	+27.1%	0%	+43.4%	+18.6%	+7.2%

## G Unweighted Poisson Regressions

In this Appendix, we replicate Table 3 without applying weights to the observations.

TABLE G.1: POISSON REGRESSIONS

	(1)	(2)	(3)	(4)	(5)	(6)
	Assets	Equity	Debt	Assets	Equity	Debt
Cultural Distance	-0.263** (0.086)	-0.229** (0.063)	-0.293* (0.131)	-0.121* (0.051)	-0.098 <sup>†</sup> (0.060)	-0.198** (0.046)
Geographic Distance	-0.285** (0.042)	-0.251** (0.032)	-0.387** (0.062)	-0.365** (0.080)	-0.444** (0.080)	-0.177 (0.108)
Controls	No	No	No	Yes	Yes	Yes
Observations	13,547	13,594	29,737	12,076	12,118	26,110