EUROPEAN CENTRAL BANK

EUROSYSTEM

Working Paper Series

Nicola Rebmann, Konstantin M. Wacker

FDI, gravity, and aggregation: revisiting the distance elasticity with sector-level FDI data

No 3245

ABSTRACT

This paper re-examines foreign direct investment motives in the 'FDI gravity' model (Kleinert and Toubal, 2010), focusing on the role of distance. More precisely, we investigate whether aggregate and pooled gravity models for FDI obscure relevant heterogeneities across sectors. This is possible through the novel MREID dataset, which provides us with FDI data at the 2-digit NAICS level for 184 countries over the period 2010 to 2020. Our results reveal that aggregate and pooled models mask significant sector heterogeneities in two aspects: (i) in the importance of horizontal versus vertical FDI motives, and (ii) in the distance elasticity. Distance is negatively correlated with FDI on an aggregate level, which is robust to multiple econometric specifications, but exhibits significant sector heterogeneity. Our analysis suggests the presence of complex sector-specific components that cannot easily be explained with standard economic rationales.

JEL classification: F21, F23, C33

Keywords: Gravity model, foreign direct investment, sector heterogeneity

Non-technical summary

Foreign direct investment (FDI) is arguably the most important type of cross-country capital flows and mostly reflects investment of multinational firms. It is widely perceived to promote economic growth, employment and productivity in host countries, which is why many countries actively aim to attract FDI. Understanding what drives FDI is hence critical to facilitate such policy efforts. Furthermore, for central banks, understanding the investment motives of multinational firms is crucial for assessing international financial linkages, particularly because FDI can reflect external dependencies, such as strategic acquisitions in security-relevant sectors or concentrated foreign ownership.

Previous empirical studies of FDI motives were limited to aggregate cross-country analysis or to firm-level analysis for one country. This can mask differences in FDI motives across sectors: what drives FDI in textile manufacturing need not matter to the same extent for FDI in agriculture or the financial industry. Analyses using aggregated data thus risk masking opposing forces across sectors and may lead to incomplete or even misleading interpretations of global investment patterns. A novel dataset on FDI across 25 sectors in 184 countries over the period 2010 to 2020 allows us to uncover sector-specific FDI motives. We rely on a bilateral gravity equation, which relates bilateral FDI stocks to the economic sizes of countries, the distance between them, and other factors influencing cross-border investment decisions.

We pay particular attention to the role of geographic distance in explaining FDI, because it can reveal what drives firms to become multinational:

- If FDI of multinational firms is driven by market-seeking objectives, distance should be positively correlated with FDI: the further away customers are, the more difficult it becomes to serve them through exports and establishing a local subsidiary through FDI becomes a viable alternative.
- Conversely, if FDI of multinational firms is driven by efficiency-seeking objectives, distance should be negatively correlated with FDI: firms will aim to offshore production to low-wage countries but the further away they are, the more expensive it becomes to slice up their value chain due to intra-firm trade costs.

The relative importance of those aspects, and what additional aspects beyond trade costs physical distance captures, will differ across sectors. For example, there is more potential to slice up value chains in manufacturing than in personal services.

Our results suggest that aggregate analyses of FDI data fail to clearly identify market-seeking and efficiency-seeking FDI motives in the data as they mash up relevant differences in FDI motives across sectors that can go into opposite directions. In particular, the distance elasticity of FDI estimated in an aggregate gravity model masks sector-specific distance effects, because the economic concepts captured by geographic distance – such as trade costs, information frictions, or coordination needs – matter to very different degrees across sectors. We also show that no simple economic rationale can explain sector-specific differences in the distance effect on FDI.

From an academic perspective, this implies that aggregate analyses may misrepresent the underlying drivers of global investment flows and more refined studies of FDI motives at the sector level are essential for a correct interpretation. From a policy perspective, the findings highlight that successful strategies to attract FDI must be sector-specific: What works for attracting FDI in a country that is specialized in textile production, for example, may not work for a country that is specialized in providing financial services.

1. Introduction

Foreign direct investment (FDI) is a key component of international capital flows. As of 2022, FDI stocks represented 29.9% of global cross-border liabilities and 30.5% of cross-border assets (Lane and Milesi-Ferretti, 2018). FDI is chiefly characterised by the intent of an investor, usually a multinational enterprise (MNE), to acquire significant, active ownership in a foreign affiliate (Organisation for Economic Co-operation and Development, 2020). Uncovering the motives that drive FDI is hence relevant to understand financial globalisation and why firms choose FDI instead of other forms of market interaction, such as exporting or arms-length offshoring. A comprehensive empirical literature, to which our paper contributes, has therefore estimated those FDI determinants (e.g. Davies, 2008; Blonigen and Piger, 2014; Schneider and Wacker, 2022).

The FDI literature traditionally distinguishes between a market-seeking, horizontal FDI motive (HFDI) and an efficiency-seeking, vertical FDI motive (VFDI); see Davies and Markusen (2021), Yeaple and Antràs (2014), and section 2.1. HFDI involves firms that replicate production abroad to substitute exports and reduce trade costs. In contrast, VFDI fragments production internationally, which involves trade costs that are weighed against production cost reductions from exploiting countries' competitive advantages (Markusen, 2013). Kleinert and Toubal (2010; 'KT' hereafter) have shown that it is possible to build a micro-founded gravity model for both FDI motives, where different parameters give evidence of prevailing motives in the data. [1]

Distance plays a decisive role for FDI motives because trade costs increase with geographical distance. In the case of VFDI, greater distance hence raises intra-firm trade costs and makes production fragmentation ("offshoring") less attractive, suggesting a negative distance elasticity. In contrast, the 'proximity-concentration trade-off' (PCT; Brainard, 1997) predicts a positive effect of distance on HFDI: the higher the transport costs, the greater is the incentive to avoid them through HFDI. While empirical studies suggest that FDI is predominantly horizontal in nature (see Markusen, 2013; Davies and Markusen, 2021), gravity estimates of FDI usually find a negative distance effect on FDI. This gives rise to a paradox pointed out by Neary (2009): higher trade costs should encourage HFDI and if the bulk of FDI is horizontal, FDI should increase with distance, not decrease.

One leading explanation for this paradox is that actual FDI is driven by complex integration strategies which do not fit neatly into either the horizontal or vertical categories (Neary, 2009; Badinger and Egger 2010). For example, HFDI may require intermediate inputs from the home country (KT, 2010). Another key explanation is that distance raises fixed setup and information costs (KT, 2010), as also pointed out by the international finance literature (e.g. Portes et al., 2001; Daude and Fratzscher, 2008). Moreover, integration costs associated with FDI increase with geographical distance if it is correlated with cultural and institutional distance (Aleksynska and

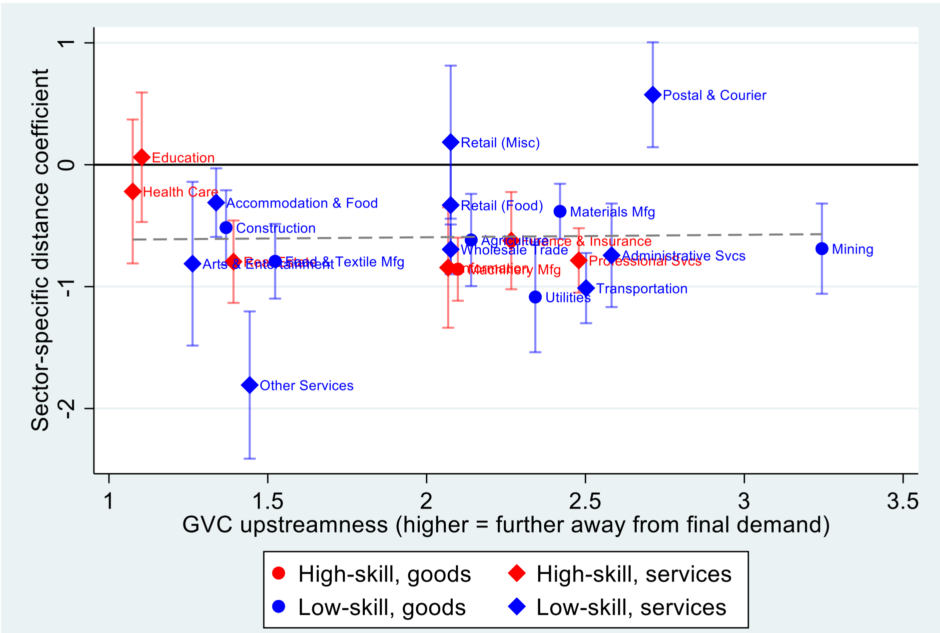
[1] Other theoretical foundations for a gravity model of FDI are possible. See, for example, Head and Ries (2008).

Havrylchyk, 2013; Beugelsdijk et al., 2017; van Hoorn and Maseland, 2016). This highlights that geographical distance captures several aspects beyond trade costs.

The relative importance of the factors captured by distance is likely to differ across sectors. For example, the importance of distance-related setup costs increase with sectors' fixed entry costs and integration costs are more relevant in contract-intensive sectors (Nunn and Trefler, 2013; Egger and Pfaffermayr 2004; Ottaviano and Turrini 2007). McCann (2011) points out that market proximity is particularly important in high-value knowledge-intensive activities and sectors that involve movement of people. For the same reason, FDI in the wholesale and retail sector are a case of export-supporting FDI (Krautheim, 2013).

Against this background, our paper investigates to what extent aggregate, as opposed to sector-level, data[2] mask FDI motives in bilateral analysis of FDI determinants, with a focus on the distance elasticity. In particular, we investigate to what extent distance is negatively associated with bilateral FDI positions and how robust this result is across econometric specifications and sectors. We find a clearly negative distance effect in aggregate and pooled data across various specifications but considerable heterogeneity across sectors, with some sectors exhibiting positive point estimates for the distance elasticity (see Figure 1). Since accounting for sector heterogeneity appears important, this leads us to the research question what plausible economic factors can explain the sector heterogeneity in FDI's distance elasticity?

Figure 1 / Sector-specific distance elasticities



95% confidence intervals based on clustered SEs (country pair). See section 5 for details.

[2] In the NAICS classification, 'sectors' refers to 1- or 2-digit categories ('industry' represents more granular classifications).

A back-of-the-envelope example illustrates why such sector-specific heterogeneity in the distance elasticity of FDI can be highly problematic. Consider Bangladesh's FDI in Malaysia: based on aggregate empirical results presented later, increasing the bilateral distance by switching the source country from Bangladesh to Israel would, on average, reduce total assets invested in Malaysian affiliates by 46.28%. However, for both Bangladesh and Israel, their FDI in Malaysia is highly concentrated in the warehouse and storage sector, accounting for 100% and 87% of total assets, respectively. A gravity model estimated for this specific sector shows that increasing the distance from Bangladesh-Malaysia to Israel-Malaysia would actually *increase* total assets by 87.18%, on average.[3] The effects from the aggregate estimate and the warehouse and storage-specific estimate therefore have opposite signs and the latter is twice as large as the former in absolute terms. Relying on aggregate gravity models can thus be highly misleading.

Our empirical analysis builds on the reduced-form 'gravity for FDI' model by KT (2010) and the novel Multinational Revenue, Employment, and Investment Database (MREID) by Ahmad et al. (2025). The MREID offers FDI data with extensive geographical and sectoral coverage, spanning both developing and developed countries as well as 25 sectors at the granular 2-digit North American Industry Classification System (NAICS) level. It thus provides a unique opportunity to assess the question of what motives drive FDI using sector-level data.

To address sector heterogeneity in FDI determinants, we estimate the FDI gravity model at three levels of aggregation. Initially, we aggregate the sectoral data up to the country level. This provides a benchmark, as it resembles what the relevant empirical literature has done. Second, we estimate a pooled sectoral model, where the data are included at sector level, but parameters are constrained to homogeneity across sectors. And third, we estimate the model separately for each two-digit sector, which allows regressors, parameters and fixed effects (FEs) to vary by sector (which is also the basis for Figure 1). This three-tier approach allows us to compare aggregate and pooled estimates to identify potential aggregation bias, and to examine the full sector heterogeneity in the distance elasticity.

The key finding of our paper is that aggregate or pooled gravity models for FDI fail to clearly distinguish between HFDI and VFDI motives in the data. The estimated parameters for investment motives, and particularly the distance elasticity, exhibit substantial sector heterogeneity. This suggests that uncovering FDI motives requires sector-specific gravity models, rather than aggregate or pooled models. The heterogeneity in FDI's distance elasticity across sectors that we document in our paper likely stems from sector-specific FDI motives and from the fact that geographical distance captures a variety of concepts beyond transport costs. Yet, we also document that an aggregate

3 The calculation is based on the distance coefficients from the pooled sectoral model in Table 2, col. (4), and from a gravity model estimated only for the warehouse and storage sector in Table B11, col. (12) in Appendix B. Moving from Bangladesh (distance: 2,613 km) to Israel as the source country (7,774 km) increases the bilateral distance by 197.5%. For such large changes in x, the formula: $((1+1.975)^{\beta}-1)\cdot 100$ following Kaplan (2023: 8.1.4) and Benoit (2011) is used to obtain the exact percentage change in total assets for an increase in distance by 197.5 %.

negative distance effect is robust to the inclusion of rich multiplicative fixed effect and to the inclusion of other distance components, including geopolitical distance.

A surprising result of our paper is that no clear economic rationale can explain a significant degree of variation in FDI's distance elasticity across sectors. This is also illustrated in Figure 1, which plots the sector-specific distance elasticities against a measure for upstreamness in global value chains (GVCs), showing no clear association (see section 5 for details). Likewise, we find no relevant correlation between sector-specific distance elasticities and sector classifications into high-skill vs. low-skill (red vs. blue in Figure 1), goods-producing vs. service-producing (different markers in Figure 1), and sector-specific trade costs—none of those variables explains more than 3% of the variation in the distance elasticity across sectors. At best, all of them together explain 5%. The Neary (2009) paradox hence does not seem to vanish once we take a sector-level perspective.

The findings of our paper contribute to the comprehensive literature on FDI determinants and links to a literature on aggregation biases in trade (Imbs and Mejean, 2015) and its workhorse gravity model (Borchert et al., 2022; Fontagné et al., 2022; Schreiber, 2022; Breinlich et al., 2024). In particular, Breinlich et al. (2024) have shown that for elasticities that vary at the sector level PPML estimation using aggregate data will approximately recover a weighted average of sector-level elasticities if the variables associated with this elasticity do not vary at the sector level, as is the case for distance. This averaging is highly problematic for FDI determinants because we may not be interested only in the quantitative magnitude of an elasticity, as is the case for trade costs (which are always negatively associated with trade), but also in its qualitative implication for what motive predominantly drives FDI. In cases like distance, the elasticities can be positive or negative, depending on whether a sector is predominantly horizontal or vertical. The above back-of-the-envelope example of Bangladesh's FDI in Malaysia shows that aggregation may lead to erroneous predictions of the direction how FDI is associated to distance.

To advance towards a genuinely sectoral gravity model, our key takeaway is to place more emphasis on theory: with the growing availability of high-quality, granular and globally comprehensive data, theoretical models are needed to accurately identify sector-specific effects and accommodate FDI motives beyond the horizontal-vertical dichotomy. These models should increasingly include regressors at sector level, where the effects of interest manifest. In particular, theoretical models should account for the various components captured by distance. What has previously been framed as an empirical puzzle – where theory-predicted effects do not appear in the data – should instead be reframed as a theoretical mismatch.

The remainder of this paper is organized as follows. We first review the related literature, with a focus on motives for HFDI vs VFDI, why they may vary at the sector level, and how they link to the gravity equation. Section 3 spells out our empirical gravity specification and how we bring it to the data. Section 4 presents our main empirical results and robustness checks; section 5 additionally explores possible reasons for sector heterogeneity in the distance elasticity. Section 6 concludes.

2. Theoretical background and related literature

2.1 Horizontal versus vertical motives for FDI

The theoretical FDI literature traditionally distinguishes two motives of FDI, each with distinct predictions on how distance affects FDI (see Table A1 in Appendix A). Market-seeking, or HFDI, formally introduced by Markusen (1984) and extended with firm heterogeneity in Helpman et al. (2004), involves firms that establish affiliates abroad, duplicating activities at the same production stage to directly supply the foreign market. HFDI substitutes exporting, thus circumventing trade costs, including tariffs. Firms face a 'proximity-concentration trade-off' (PCT; Brainard, 1997, p. 520), weighing trade costs and plant-level fixed costs against one another. The proximity motive prevails and encourages HFDI when trade costs, typically proxied by geographical distance, are high. Exporting is favoured when plant-level economies of scale are large, as it allows for the concentration of production in a single location. The PCT hence suggests a positive distance effect on HFDI that contrasts with a typical gravity effect, the empirical evidence for which is discussed in Section 0. This paper seeks to reconcile this mismatch between theory and empirical findings by examining sector-level data. The main studies supporting HFDI as the dominant FDI motive include Brainard (1997), who was the first to demonstrate that FDI is driven by country similarities rather than factor endowment differences, and Markusen and Maskus (2002), who provide empirical evidence for the horizontal, but not the vertical, model.

The second motive is efficiency-seeking, or VFDI, first formalised by Helpman (1984). VFDI views the production process as vertically fragmented into distinct stages with different factor intensities. For certain stages, the MNE establishes new production facilities in a foreign location to lower production costs by exploiting comparative advantages in factor endowments across countries. The standard view suggests that MNEs locate their skill-intensive headquarters in countries that are abundant in high-skilled labour and establish production in countries that are abundant in low-skilled labour (Fukao and Wei, 2008). Consequently, VFDI increases with high-skilled labour abundance of the home country relative to the host country. The home country remains the primary destination market, while foreign affiliates produce vertically linked goods and export them back to the parent firm. Intra-firm trade, therefore, complements VFDI and results in an unambiguously negative effect of distance on VFDI (Alfaro and Chen, 2018), conditional on endowment differences. The key trade-off in VFDI lies between reducing input costs and incurring higher trade costs, along with forgone economics of integration. Support for VFDI is provided by Carr et al. (2001) and the model refinements of Braconier et al. (2005a; 2005b) and Davies (2008), who find strong support for the vertical component in the knowledge-capital model. Alfaro and Charlton (2009) support the evidence of a vertical motive in FDI by accounting for intra-industry VFDI, where MNEs offshore activities close to the final production stage. This granular form requires 4-digit level data for identification, while inter-industry VFDI is visible in 2-digit level data, as it spans more distant production stages.

Theoretical models often draw a strict distinction between HFDI and VFDI. However, empirically, this separation is less clear-cut. FDI increasingly reflects complex corporate strategies that combine market access and factor cost motives (Badinger & Egger, 2010; Bricongne et al., 2023), or export-platform motives (Ekholm et al., 2007; Yeaple, 2003b), which reflect the multilateral nature of FDI.

2.2 Identifying FDI motives in the data

As the motive of investment is typically not disclosed by investors (Kox and Rojas-Romagosa, 2020), scholars have developed ex-post approaches to divide data into HFDI and VFDI. Most of these approaches, however, require firm-level data on the destination (origin) of affiliate sales (purchases).[4] Given that such data are not available on a global scale, a more flexible approach is to specify an empirical model that describes the determinants of bilateral FDI positions and allows discrimination between different motives.

A key empirical model to estimate bilateral determinants of FDI is the gravity model, which posits that flows between two countries are directly proportional to their economic sizes (the 'mass'; typically measured by their GDPs and weighted by a gravitational constant). Frictions countervail the attractive force of the 'mass', with distance as the most prominent proxy of such frictions (Yotov et al., 2016).[5] For a long time, gravity equations for FDI remained ad hoc (Dorakh, 2020), but KT (2010) made a seminal contribution by establishing a first theoretical foundation of 'gravity for FDI'. KT (2010) take two proximity-concentration models of purely horizontal MNEs, as well as a factor-proportions model of purely vertical MNEs, and derive gravity equations for foreign affiliate sales (FAS) from each model. The key outcome is the reduced-form gravity model presented in eq. (1) below, which accommodates both HFDI and VFDI, marking a significant advance from previous FDI gravity models, which were limited to HFDI. Critically, the horizontal and vertical FDI motives imply different predictions for the gravity parameters, owing to the different trade-offs they generate. Thus, eq. (1) provides testable predictions for the model parameters that allow for inference as to whether an HFDI or VFDI motive is dominant in the data. These predictions are summarised in Table 1. The gravity model for FDI in log-linear form can be written as:

$$Y_{ij,t} = \beta_1 \ln(GDP_{i,t}) + \beta_2 \ln(GDP_{j,t}) + \delta \ln(D_{ij}) + \vartheta_1 SkEDiff_{ij,t} + \vartheta_2 \ln(GDP_{i,t} + GDP_{j,t}) + \varphi_t + \varphi_i + \varphi_j + \varepsilon_{ij,t}, \quad (1)$$

where $Y_{ij,t}$ is FDI from source country *i* in host country *j* in year *t*. The first three regressors form the core of the gravity model: source and host country GDP in log form and bilateral distance in log form, D_{ij}. Furthermore, KT (2010) include a variable to capture skill endowment differences

4 For US outward FDI, such data are used by Alfaro and Charlton (2009), who identify VFDI by using the input-output linkages between the headquarter and affiliate industry, or by Ramondo et al. (2011), who identify VFDI based on intra-firm trade data.

5 In its pure, multiplicative form, the gravity model is written as follows: $Y_{ij} = G\frac{X_i^{\alpha}X_j^{\beta}}{D_{ij}^{\gamma}}$ (Yotov et al., 2016).

($SkEDiff_{ij,t}$) and the sum of GDPs. The estimated model parameters, when compared with the testable predictions for HFDI and VFDI summarised in Table 1, allow for determining which motive predominates in the data.

Table 1 / Predictions for the gravity variables based on Kleinert and Toubal (2010)

	Horizontal model	Vertical model
Source country market size β_1	1	<0
Host country market size β_2	1	>0
Bilateral distance δ	<0 ?	<0
Skill endowment difference ϑ_1	0	>0
Sum of GDPs ϑ_2	0	1

In the vertical model, the source country's GDP, which reflects its supply capacity, has a negative effect on FDI ($\beta_1 < 0$). This effect occurs because VFDI exploits factor abundances in foreign countries to minimise production costs. In contrast, the host country's GDP has a positive effect ($\beta_2 > 0$). The vertical motive strengthens with larger skill endowment differences between source and host ($\vartheta_1 > 0$) and is hindered by distance, which impedes intra-firm trade ($\delta < 0$). The sum of GDPs serves to pin down market demand and is expected to have a unitary elasticity ($\vartheta_2 = 1$) (KT, 2010; Venables, 1999).

In the horizontal model, host country GDP has a positive effect on FDI, as establishing production abroad is profitable only in a sufficiently large foreign market. Otherwise, it would be more efficient to concentrate production at home and serve the foreign market through exports. This coefficient is expected to equal one, similar to the source country's GDP, whose positive impact on HFDI arises from the need for domestic inputs in affiliate production ($\beta_{1,2} = 1$). No significant effect is anticipated from skill endowment differences or from the sum of GDPs ($\vartheta_{1,2} = 0$). KT (2010) expect a negative effect of distance ($\delta < 0$) because affiliates rely on some source-country inputs and because the fixed set-up costs of FDI are assumed to increase with distance. This is in contrast to the positive distance effect suggested by the theory of the 'proximity-concentration trade-off' (PCT), introduced in Section 2.1.

2.3 FDI motives and the distance elasticity of FDI

The key hypothesis of this paper is that the distance elasticity of FDI differs across sectors. This reflects that HFDI and VFDI motives, which are likely to vary across sectors, imply different distance effects and geographical distance proxies for various aspects that differ by sector and motive. For example, transport costs are more relevant for the utilities sector, while a shared culture and language plausibly play a larger role for services.

The dominant FDI motive across sectors has not yet been systematically analysed within the gravity framework,[6] but vertical motives should lead to a negative relation between distance and FDI as greater distance raises intra-firm trade costs. For sectors characterised by HFDI, however, the PCT framework predicts a positive effect of trade costs, and hence distance, on the absolute volume of HFDI and its share in total multinational activity (Alfaro and Chen, 2018; Egger and Pfaffermayr, 2004). Empirical evidence in favour of the PCT is, however, scarce and limited to the transport cost component of distance.[7] Most empirical studies find a gravity effect for HFDI, contradicting the PCT (Neary, 2009).[8] KT (2010) try to rationalise this with HFDI requiring headquarter (HQ) inputs from the source country, but we do not consider this explanation to be fully satisfactory. Instead, we propose that predictions on the sign of the distance elasticity of FDI should be made at the sector level. This also relaxes the assumption of the PCT framework that HFDI is a perfect substitute of exporting (Castellacci, 2014) in all sectors. In non-tradable industries, for instance, exporting is either prohibitively expensive or entirely unviable (Ramasamy and Yeung, 2010).

A central argument for considering sector heterogeneities is that distance captures a broader range of cost components beyond transport costs (Neary, 2009; McCann, 2011), which vary in relevance across sectors. These key components are outlined below. Each can be assumed to correlate with geographical distance (Hutzschenreuter et al., 2016), thereby contributing to the negative distance effect on HFDI while extending the explanation of KT (2010).

(i) **Trade costs.** The core assumption of the PCT is that HFDI entirely bypasses transport costs, tariffs and non-tariff barriers (NTBs), where distance is commonly used to proxy physical trade costs. KT (2010) and Irarrazabal et al. (2013), however, argue that foreign affiliate production still relies on intra-firm trade for inputs from their headquarters, which incurs trade costs and leads to a negative distance effect on HFDI.

(ii) **Fixed cost of HFDI.** According to KT (2010) and Irarrazabal et al. (2013), another argument for a negative distance effect is the fixed cost of HFDI. This fixed cost can be divided into the costs of evaluating a potential investment, plant set-up costs, and adaptation costs when starting to operate in the host country (Cezar and Escobar, 2015).

(iii) **Cultural distance**, closely tied to language, historical relationships and taste. Lankhuizen et al. (2011) and Beugelsdijk et al. (2018) emphasise that cultural distance complicates business

[6] Previous studies on sectoral FDI hence provide only incidental evidence, particularly for HFDI in most service sectors and more vertical motives in manufacturing and business services (e.g. Buch et al., 2005; Kolstad & Villanger, 2008; Riedl, 2010).

[7] Brainard (1993; 1997) demonstrates that higher freight costs positively affect the share of affiliate sales in total sales. Similarly, Carr et al. (2001) find a positive effect of the host country's trade cost index on the volume of real sales of US multinationals' affiliates. In contrast, Yeaple (2003a) identifies a negative effect of freight costs on total sales of US affiliates.

[8] This evidence is, however, limited to aggregate data: Carr et al. (2001) report a distance elasticity of affiliate sales ranging from -0.8 to -1.8, a magnitude similar to gravity of trade (Head and Mayer, 2014). Aggregate gravity equations with different dependent variables confirm this effect, including Head and Ries (2008) for global bilateral mergers and acquisitions (M&A) data, Ramondo et al. (2015) for the affiliate revenues of US MNEs, and Alfaro and Chen (2018), who use a bilateral dataset of MNE entry decisions.

operations abroad, owing to different organisational practices and a lack of trust. Cultural similarity, in turn, is bolstered by geographical proximity, but also by historical ties and a shared language, each of which facilitate communication and foster trust (Ly et al., 2018).

(iv) **Institutional distance.** Bénassy-Quéré et al. (2007), using economic freedom indices from the Fraser Institute, demonstrate that institutional distance in credit and labour market regulations negatively affects FDI. Cezar and Escobar (2015), employing an institutional distance index based on the Worldwide Governance Indicators (WGIs), further clarify that the negative effect occurs both at the extensive and the intensive margin of FDI: institutional distance reduces the likelihood of a firm investing abroad, as well as the volume of FDI among OECD countries.

(v) **Frictions for information flows**. This component reflects both pre- and post-investment information frictions. Before entry, firms face costs related to evaluating foreign markets (Hashimoto and Wacker, 2016; Carballo et al., 2021), assessing risks (Baier and Welfens, 2019), understanding regulations (Markusen and Strand, 2009), and navigating screening mechanisms (Harding and Javorcik, 2011; Gregori and Nardo, 2021). After entry, frictions persist owing to reliance on knowledge transfer from the parent firm – such as managerial oversight, marketing, or R&D – which becomes more difficult as geographical distance increases (Keller and Yeaple, 2013; Head and Ries, 2008).

We expect the relevance of these five distance components to vary across sectors. Therefore, we expect that the distance elasticity varies across sectors, which is particularly relevant for sectors that are typically subject to HFDI. Predictions for these sector-specific distance effects are formulated in the following section.

2.4 Sector-specific considerations for the distance elasticity

For sectors in which VFDI is likely to predominate, such as machinery or electronics manufacturing, geographical distance is expected to impede intra-firm trade, thereby exerting a clear negative effect on FDI, although with varying magnitude across sectors. As a notable exception, in R&D- or resource-based sectors such as mining, where access to specialised knowledge clusters or natural resources is key, resource availability may outweigh geographical frictions, leading to a less pronounced distance effect on VFDI (Castellani et al., 2013; Buch et al., 2005).

For sectors dominated by HFDI, the distance elasticity is more ambiguous and depends strongly on sector characteristics, as follows:

(1) A *small negative or even positive distance elasticity* is expected in tradable sectors where exporting is a viable alternative and the PCT develops. In these sectors, the HFDI motive strengthens with greater distance, especially where sectoral transport costs are high (e.g. in agriculture). Similarly, a small negative or positive effect is predicted for sectors where distance is integral to the business model, such as hospitality (Falk, 2016).

(2) A *stronger negative distance elasticity* is expected in: (i) sectors with high fixed set-up costs, such as utilities, or high adaptation costs, such as the energy sector, telecommunications infrastructure, or defence production; (ii) taste-dependent sectors that benefit from cultural proximity, such as the arts, entertainment and recreation industry; (iii) knowledge-intensive sectors that rely heavily on HQ inputs that are non-routine, difficult to codify and complex, such as legal services (Keller and Yeaple, 2013). In these sectors, firms face the PCT, but the gravity effect is likely to outweigh the positive distance effect.

(3) The *strongest negative distance effect* is anticipated in non-tradable service sectors where the PCT does not apply – e.g. in-person services (Dossani and Kenney, 2007) or highly customised service sectors such as business-specific software, and financial and telecommunication services (Castellacci, 2014).

2.5 Limitations of aggregate gravity models

These sector-level predictions highlight why it is problematic that FDI gravity models rely on data aggregated at the country level: the net effect of distance on FDI in an aggregate (or pooled) gravity model blends sector-specific effects, which result from the FDI motive and from the different distance-related components.

The trade literature has begun to assess this aggregation issue in gravity models. Breinlich et al. (2024) recently clarified that aggregate estimates provide an average of sectoral elasticities, weighted by the sector shares in total trade.[9] However, as illustrated by our earlier Bangladesh-Malaysia versus Israel-Malaysia example, such averages can lead to extremely misleading predictions on an aggregate level. Additionally, Imbs and Mejean (2015) highlighted an econometric bias in gravity models for trade.[10]

This paper is the first to extend the aggregation issue to the FDI literature. It fills a gap in this literature by estimating KT's (2010) gravity model at varying levels of aggregation, extending to the granular sub-sector level. This approach enables a clearer understanding of sectoral heterogeneities in FDI motives and the distance elasticity.

[9] The weighted average derivation applies to conventional gravity equations that do not include any sector-level regressor. When estimating a gravity model for each sector individually, the samples vary, based on the country pairs where FDI occurs within that sector. However, the regressors, including the FEs, can still be interpreted as a set of dummies that remain constant across sectors, although their coefficients differ by sector (Breinlich et al., 2024).

[10] This bias (which itself is not the focus of our paper) emerges when sector- or product-level heterogeneity in aggregate regressions is mechanically absorbed by the residual, and creates a correlation between the residual and the regressor. This econometric bias gave rise to the term 'elasticity pessimism' (Orcutt, 1950), which refers to an underestimation of the price elasticity of imports in aggregate data.

3. Data and methodology

3.1 A sectoral gravity model

Based on the reduced-form gravity model of KT (2010) in eq. (1) and guided by the results of Breinlich et al. (2024), we first estimate a gravity model using data aggregated across sectors at the country-pair-year level. This omits the sector dimension of the data and aligns with the aggregation level most commonly employed in the empirical literature (e.g. Schneider and Wacker, 2022).

$$Y_{ij,t} = exp\{\beta_1 \ln(GDP_{i,t}) + \beta_2 \ln(GDP_{j,t}) + \delta \ln(D_{ij}) + \vartheta_1 SkEDiff_{ij,t} + \vartheta_2 \ln(GDP_{i,t} + GDP_{j,t}) + \boldsymbol{Z}_{i,t}\theta + \boldsymbol{Z}_{j,t}\upsilon + \boldsymbol{Z}_{ij,t}\gamma + \boldsymbol{Z}_{ij}\zeta + \varphi_t + \varphi_i + \varphi_j\} \times \varepsilon_{ij,t} \quad (2)$$

$Y_{ij,t}$ is FDI from source country *i* in host country *j* in year *t*. The variables for whose parameters KT (2010) provide testable predictions regarding the FDI motive (Table 1) are included in explicit form: source and host GDP (log), bilateral distance (log), the skill endowment difference, and the sum of GDPs (log). Additionally, $\boldsymbol{Z}_{i,t}$ ($\boldsymbol{Z}_{j,t}$) is a vector of control variables that vary across time and source (host) country, such as the exchange rate or the host country tariff. $\boldsymbol{Z}_{ij,t}$ captures controls that vary across time and country pair, such as distance in financial development or institutions, and $\boldsymbol{Z}_{ij}$ captures time-invariant bilateral variables, such as colonial ties. All variables and their data sources are presented in Section 3.3 and summarised in Table 3 in Appendix A.

Breinlich et al. (2024) find that, when regressors are constant across sectors, estimates from aggregate or pooled sectoral models – where data are used at sector-level *s* but parameters are constrained to homogeneity – should be identical. In the second step, we hence estimate a pooled sectoral model to identify any potential econometric aggregation bias. Compared with eq. (2), this model includes additive sector FEs φ_s.

$$Y_{ij,t}^{s} = exp\{\beta_1 \ln(GDP_{i,t}) + \beta_2 \ln(GDP_{j,t}) + \delta \ln(D_{ij}) + \vartheta_1 SkEDiff_{ij,t} + \vartheta_2 \ln(GDP_{i,t} + GDP_{j,t}) + \boldsymbol{Z}_{i,t}\theta + \boldsymbol{Z}_{j,t}\upsilon + \boldsymbol{Z}_{ij,t}\gamma + \boldsymbol{Z}_{ij}\zeta + \varphi_t + \varphi_i + \varphi_j + \varphi_s\} \times \varepsilon_{ij,t} \quad (3)$$

The results of this pooled model serve as a benchmark for comparison with the sector-specific effects estimated in the final step. We estimate eq. (3) for each sector individually (omitting sector FEs, which are captured in the constant for each sector). This results in 23 separate regressions.[11] This approach allows all parameters of regressors and FEs to vary by sector, while the regressors themselves vary at the country or country-pair level but are constant across sectors.

[11] Table A2 displays all sectors with their description and their share in total assets. Sector 92, ‘public administration’, which provides too few observations for precise estimates, is excluded from the regressions, as is sector 99, ‘unclassified establishments’.

3.2 Estimation

We use a non-linear Poisson Pseudo Maximum Likelihood (PPML) estimator with various FEs. Year FEs φ_t eliminate the effect of common shocks that equally affect all country pairs or country-pair-sector combinations. Country (sector) FEs φ_i, φ_j (φ_s) control for unobserved, time-invariant characteristics that are specific to each country (sector). This identification strategy represents a rather unrestrictive FEs specification in the gravity literature.

More restrictive approaches are included as robustness checks. Country-pair FEs account for time-invariant, and potentially endogenous, components of bilateral trade costs (Egger and Nigai, 2015), but absorb the effect of the distance variable, which is essential to our analysis. A second robustness check augments the gravity equation with source-year and host-year FEs to control for the structural multilateral resistance (MR) terms.[12] As the theoretical interpretation of MR in the context of FDI is, however, ambiguous (Schneider and Wacker, 2022), and the country-time FEs absorb the effect of the market sizes, MR terms are not included in the baseline regressions.

To address potential serial correlation in the error term, standard errors are clustered at the directional country-pair level, following the standard approach in the gravity literature (Egger and Tarlea, 2015).[13] Previous research on gravity models has identified heteroskedasticity in the error term as a major issue. Furthermore, bilateral trade and FDI data typically include many zeros, which are omitted in a log-linear transformation, leading to biased estimates (Helpman et al., 2008).[14] To address these two issues, we estimate the gravity model using the PPML estimator, as implemented through the STATA command *ppmlhdfe* for high-dimensional FEs (Santos Silva and Tenreyro, 2006; Correia et al., 2020).

3.3 Data

We use panel data ranging from 2010 to 2020 and spanning 145 countries.[15] The final dataset combines the MREID with data from the Penn World Tables 10.0 (Feenstra et al., 2015), the CEPII gravity dataset (Conte et al., 2022), the World Development Indicators (World Bank, 2024a), the Fraser Institute (2024), Harms and Knaze (2021), and the International Monetary Fund (2024a). From the raw MREID, domestic investment was excluded, as well as observations where key gravity variables, particularly the skill difference variable, have missing values. The excluded countries are marked in Table A6. Eliminating observations with missing values reduces the total volume of assets by 13.37%. The final dataset includes data from 3,328 country pairs, comprising a

---

12 MR reflects the idea that trade or FDI between two countries depends not only on their bilateral relationship but also on their relative remoteness from the rest of the world (Anderson and van Wincoop, 2003; Yotov et al., 2016).

13 The need to cluster standard errors arises from the fact that a country-pair enters the sample several times, depending on the number of sectors and years in which FDI occurs within this pair (Castellani et al., 2013).

14 The final dataset used contains 77,502 observations with zero total assets, which account for 32.38% of the data.

15 Table A6 lists all countries included in the final dataset. The year 2021 is excluded from the analysis to prevent distortions caused by the COVID-19 pandemic.

total of 239,381 observations. Table A3 shows descriptions and data sources for all variables, and Table A4 presents summary statistics.

3.3.1 Dependent variable

Our dependent variable *Y* is total assets of foreign affiliates, which represents the sum of current assets and fixed assets on affiliates' balance sheets.[16] Total assets proxy the FDI stock, commonly used in the FDI literature because of its lower volatility compared with FDI flows and its reflection of multinational presence in an economy (Wacker et al., 2025). Although KT (2010) base their predictions on FAS rather than the FDI stock, Wacker (2013) and Wacker et al. (2025) document a high correlation between assets and sales, supporting the application of KT's (2010) elasticities to total assets. For robustness, we also use the total operating revenue as an alternative dependent variable, maintaining alignment with KT's (2010) FAS framework.

We retrieve both dependent variables from the novel MREID (Ahmad et al., 2025), which aggregates operational firm-level data from Orbis and thereby addresses key limitations of traditional balance-of-payments (BOP) statistics.[17] BOP data track the net financial capital that parents provide their affiliates, but do not necessarily reflect real economic activity (Beugelsdijk et al., 2010). Typical reasons for this deviation are profit-shifting or tax engineering motives of phantom FDI (Damgaard et al., 2024). In contrast, operational data, assets or sales offer more accurate proxies for MNE's real activities (Wacker, 2013). Total assets capture FDI from both parent funding and external host-country sources, mitigating biases linked to affiliate capital structures and avoiding the underestimation of real activity often seen in BOP data (Markusen, 2002).

The MREID's most notable feature is that it is the first dataset to achieve this combination of comprehensive country and sector coverage across 25 two-digit NAICS sectors. In this paper, we focus on exploiting this sector disaggregation, although, when estimating eq. (2), we aggregate the data at country-pair-year level across sectors.[18] Global heat maps in Figure A2 illustrate the broad country coverage.

16 In a robustness check, we use total assets deflated by the US GDP deflator (using the pl_gdpo series of the Penn World Tables) as the dependent variable for the pooled sectoral regressions from Panel B in Table 2. We thus follow Schneider and Wacker (2022) in using the price level of US GDP to deflate global nominal asset prices when they are combined in the model with real variables such as real GDP. The results prove to be highly robust to this adjustment (see Table B1 in Appendix B).

17 For definition as an affiliate, the MREID uses an ownership threshold of 50.01%, in line with the IMF and OECD definition of FDI to establish a lasting interest by a resident entity of one economy in another (Organisation for Economic Co-operation and Development, 2020; International Monetary Fund, 2024b).

18 This country-level dataset contains 49,115 observations.

3.3.2 Explanatory variables

Gravity variables. The key explanatory variables are KT's (2010) gravity variables: we proxy the market size (*GDP*) of source and host country by real GDP at constant 2015 USD from the World Development Indicators. Population-weighted distance (*D*) is retrieved from the CEPII gravity dataset. The sum of source and host country real GDP (in log form) is used to proxy market demand. We use the Human Capital Index (HCI) of the Penn World Tables to construct a bilateral skill endowment difference variable that proxies relative factor endowment (*SkEDiff*). The HCI is based on the average years of schooling from Barro and Lee (2013). After forecasting the value for 2020 by a country-specific linear trend, we constructed the skill level difference by host minus source and use its absolute value to ensure that the variable increases in dissimilarity. Our approach is consistent with KT's (2010) vertical model of FDI, in which high- and low-skilled labour constitute the production factors. VFDI is expected to increase with high-skilled labour abundance of the source relative to the host country. We prefer this approach to using a direct measure of factor costs, such as local wages, as wages are extremely heterogeneous by industry and skill level and are not globally available at the sector level. The skill difference variable primarily captures inter-industry VFDI driven by large gaps in skilled labour endowment, but fails to identify intra-industry VFDI at closely related production stages, which depends on more granular comparative advantages (Alfaro and Charlton, 2009). Additionally, the variable overlooks other factor endowment differences that can drive VFDI, such as capital abundance or environmental regulation differences, as highlighted by the pollution haven hypothesis (Rezza, 2013; Gerard and Chiappini, 2025).

3.3.3 Control variables

Surrounding market potential. Calculated as in eq. (4) (Schneider and Wacker, 2022), a positive effect of the SMP variable indicates the presence of export-platform FDI in the data, which occurs when a parent firm invests in a host country to serve third-party markets from there. The inclusion of SMP is critical, as its omission biases upwards the importance of the host country market size (Schneider and Wacker, 2022).

$$SMP_{ij,t} = \sum_{i_l \neq i}^{I} \frac{GDP_{i_l t}}{D_{i_l j}}, \qquad (4)$$

where $GDP_{i_l t}$ and $D_{i_l j}$ are defined as under "gravity variables" above.

Host country tariff. We proxy trade costs using the host country's weighted mean applied tariff (log) from the World Bank (2024a). Higher tariffs increase trade costs, with a negative predicted effect on VFDI and an ambiguous effect on HFDI. To capture NTBs, we include a 'contiguity' dummy, set to one if source and host country share a border, as neighbouring countries often facilitate FDI through streamlined processes.

Variables controlling for other distance components

Colonial relationships can reduce information frictions by shaping linguistic, legal, educational and social structures. From the CEPII gravity dataset, we include dummies for past colonial ties (post-1945), a common official language and a common coloniser after 1945, all of which can be expected to positively influence FDI.

Institutional distance. We use the difference in the Fraser Institute's (2024) Economic Freedom Summary Index between source and host country to proxy for institutional distance. Based on 26 components across five areas – legal system, size of government, sound money, freedom to trade and regulation – this index (scaled 0-10) is expected to show a negative effect on FDI.

International finance variables

Exchange rates (XRs) and volatility. XRs influence international asset prices and FDI by affecting relative investor wealth in imperfect capital markets (Froot and Stein, 1991). This wealth effect depends on the direction of the XR movement and the FDI type (Blonigen, 1997). XR volatility adds uncertainty about the return, but its impact on FDI remains ambiguous (Lee and Min, 2011). Harms and Knaze (2021) show that higher expected XR volatility (i.e. more flexible XR regimes) makes exporting more attractive than HFDI, while no clear prediction exists for VFDI. We include official XRs for host and source country from the World Development Indicators and the bilateral de jure XR regime index by Harms and Knaze (2021). However, no clear predictions regarding the expected signs emerge from the literature.

Distance in financial development. Although financial development generally boosts FDI (e.g. di Giovanni, 2005; Bilir et al., 2019), Donaubauer et al. (2021) find that bilateral FDI increases when both host and source countries have well-developed financial markets. We therefore include the difference in the IMF Financial Development Index (International Monetary Fund, 2024a), expecting a negative effect on FDI.

Appendix A presents and discusses descriptive statistics of our sample, including moments of the variables, top ten source and host countries, and Sankey plots.

4. Empirical results

Section 4.1 compares the results of the aggregate and the pooled sectoral gravity model, evaluating the predominant FDI motive in the data. Section 4.2 shows alternative specifications that demonstrate a robust physical distance effect on FDI. Section 4.2 analyses the results of the sector-specific gravity regressions in relation to the estimates from the pooled model, revealing how the latter masks sector heterogeneities in both the FDI motive and distance elasticity.

4.1 Aggregate versus pooled sectoral model

The key finding in this section is that PPML estimates remain largely insensitive to the use of aggregate data versus pooled sectoral data with parameters constrained to homogeneity. Table 2 presents this result by contrasting the aggregate model (Panel (a)), where data are aggregated up to the country-level, with the pooled sectoral model (Panel (b)). Confirming Breinlich et al. (2024), the estimated coefficients and standard errors are invariant to the level of aggregation, with only minor changes in the coefficients when going from column (1) to (3) or column (2) to (4). Hence, there is no evidence for an econometric aggregation bias in this particular case, which does not allow for heterogeneity across sectors.[19]

Based on the full specification of the pooled sectoral model in col. (4), a HFDI motive appears to be dominant in the data, which is consistent with previous literature. However, there is no unequivocal support for the theoretical predictions from Table 1. This discrepancy could stem from homogenous parameters that pool sector-specific effects and mask the true underlying FDI motives at the sector level.

The coefficients of source and host GDP are both positive and significant, although the host GDP estimate notably deviates from the unitary elasticity predicted for HFDI.[20] The positive coefficient of source country GDP contradicts the VFDI prediction. As expected, the distance elasticity is negative. The sign contradicts the PCT argument, but does not permit a distinction between VFDI and HFDI. This elasticity remains robust when controlling for additional distance components (col. (3) versus col. (4)). None of these proxies is significant, which may reflect a blending of sector-specific effects that potentially offset one another. The highly significant negative coefficient for skill endowment differences challenges the VFDI motive, as greater endowment differences are expected to incentivise efficiency-seeking VFDI. A vertical motive is further rejected by the negative sign of the sum of GDPs coefficient. Yet, both coefficients also fail to align with HFDI. In line with Schneider and Wacker (2022), we find evidence for export-platform FDI, as indicated by the positive and significant effect of the SMP variable.

---

[19] Notably, when including country-pair FEs (see Table B3 in Appendix B), aggregate and pooled estimates are almost identical, as the FEs account for a varying number of sectors across country pairs.

[20] Notably, when using revenue as the dependent variable, in a robustness check displayed in Table B4, these two coefficients move closer to unity, suggesting that the horizontal motive is stronger when using the FAS proxy as a measure of FDI.

Table 2 / Aggregate and pooled sectoral gravity model. Dependent variable: Total assets

	Panel (a): Aggregate model		Panel (b): Pooled sectoral model	
	(1)	(2)	(3)	(4)
DV: Total assets	Aggregated at country level	Aggregated at country level With controls	NAICS 2-digit level	NAICS 2-digit level With controls
Source GDP (log)	**1.286****	**1.451****	**1.292****	**1.417****
	(0.538)	(0.639)	(0.538)	(0.642)
Host GDP (log)	**1.997*****	**1.817*****	**1.996*****	**1.811*****
	(0.327)	(0.355)	(0.323)	(0.353)
Distance (log)	**-0.641*****	**-0.649*****	**-0.548*****	**-0.570*****
	(0.100)	(0.120)	(0.0969)	(0.117)
(abs.) Skill endowment diff.	**-0.662*****	**-0.695*****	**-0.771*****	**-0.785*****
	(0.199)	(0.225)	(0.204)	(0.235)
Sum of GDPs (log)	**-0.285**	**-0.384***	**-0.327**	**-0.416***
	(0.203)	(0.231)	(0.215)	(0.242)
Surr. market potential (log)	0.970*	2.162***	0.864*	2.005***
	(0.521)	(0.571)	(0.448)	(0.530)
Common coloniser		0.335		0.253
		(0.399)		(0.391)
Colonial relationship		0.516		0.150
		(0.317)		(0.324)
Common off. language		0.178		0.203
		(0.230)		(0.220)
Contiguity		0.439		0.432
		(0.270)		(0.267)
XR of host country		0.000100*		9.48e-05
		(6.07e-05)		(5.88e-05)
XR of source country		-6.88e-06		5.57e-06
		(3.46e-05)		(3.76e-05)
XR regime (de jure)		0.0370		0.0471
		(0.0314)		(0.0306)
Dist. in financial development		-0.873		-0.853
		(0.741)		(0.735)
Dist. Economic Freedom Index		-0.121		-0.118
		(0.105)		(0.104)
Host country tariff (log)		-0.0413		-0.0410
		(0.0489)		(0.0491)
Constant	-34.84***	-45.50***	-32.54***	-45.18***
	(11.30)	(13.45)	(10.86)	(13.20)
N	36,095	30,903	239,276	214,155
Country pairs	3,337	3,211	3,337	3,211
Pseudo R2	0.851	0.874	0.601	0.765

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Panel (a) regressions are estimated using eq. (2), Panel (b) regressions using eq. (3) with parameters constrained to homogeneity across sectors. Panel (a) regressions include additive year, host and source country FEs, while Panel (b) regressions additionally include additive sector FEs. All regressions include additive FEs for outlier country pairs (see Appendix A). N decreases, owing to missing values in some control variables. Additional observations are automatically dropped from the PPML estimation that singletons.

The results in this section rule out an econometric aggregation bias. However, the overall findings from the pooled model do not provide a definitive conclusion regarding a dominant FDI motive and unambiguous distance elasticity in the data. This ambiguity could arise from sector-specific effects that are explored in section 4.3.

4.2 Robustness of the distance effect

Several robustness checks validate our main findings. These checks include using revenue as the dependent variable (Table B4) and the introduction of the time-varying main regressors lagged by one period to address potential reverse causality (Table B6), which accounts for reverse causality as, for instance, FDI might boost current GDP via productivity spillovers (Riedl, 2010).[21] F-tests on joint insignificance of the main gravity variables reject the null hypothesis that their combined effect is equal to zero (Table B8, Table B10). Furthermore, Ramsey regression specification error tests (RESETs) yield no evidence of general misspecification (Table B7, Table B9), which lends credibility to the robustness of our results.

4.2.1 Multiplicative fixed effect structures

Additionally, we controlled for source-year and host-year FEs to control for conventional MR terms (Table B5), and for an even more demanding source-sector-year and host-sector-year fixed effects. In the latter case, the analogues to the distance coefficients in Table 2, columns (3) and (4), are estimated to be -0.637*** and -0.672*** respectively and, hence, appear to be even somewhat stronger in magnitude. Yet, those point estimates fall within a 95% confidence interval of the corresponding baseline parameters in Table 2. Results are available upon request.

4.2.2 Distance deciles instead of continuous variable

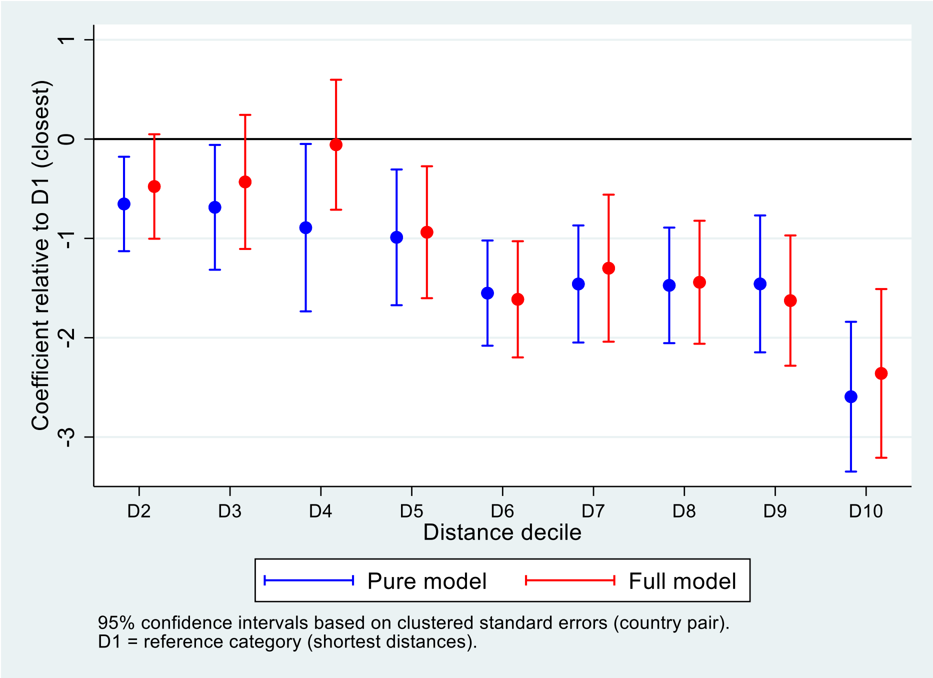
It may be the case that the inclusion of a continuous (log) distance variable into the bilateral FDI model is an over-simplification. We hence generated deciles of distance (in terms of country pairs) and included them as dummy variables into our regression equations for the “pure” (basic) and “full” model (with additional controls). The results are graphically displayed in Figure 2, with decile 1 (D1) as the reference category, and higher deciles indicating larger bilateral distance.

The results suggest that a continuous distance variable may be a “good enough” approximation but does not necessarily capture the full story about FDI’s distance elasticity. For the pure model (blue coefficient dots), FDI continuously declines with distance deciles up to D6, where the negative effect seems to plateau, before becoming again increasingly negative for decile D10. For the full model (red coefficient dots), deciles D2-D4 seem to experience a bit less of FDI compared to the reference (D1), but imprecisely estimated and with no clear trend within the first four deciles. From D4/D5 onward, there is a continuous decline in bilateral FDI for each decile except for the minor outlier of D6.

---

[21] Even with lagged regressors, past values of the regressor may still be correlated with the error term in the current period, especially if feedback effects occur over time. Fully controlling for reverse causality thus requires an instrumental variable approach (Wooldridge, 2020).

Figure 2 / Deciles of distance



Notes: The figure shows dummy variable coefficients for deciles of distance for models that resemble Table 2, col (3) and (4), but replace the continuous (log) distance variable with distance deciles.

4.2.3 Accounting for geopolitical distance

Among the several potential aspects that distance captures is geopolitical alignment – physically close countries are often politically well-aligned.[22] Recent studies such as Aiyar et al. (2024) and Gopinath et al. (2025) have indeed highlighted cross-border investment fragmentation along geopolitical lines. We follow those studies in calculating "ideal point distances" (IPDs) from United Nations voting patterns based on the July 2025 data vintage from Bailey et al. (2017). This data reflects UN General Assembly voting alignment with the US-led liberal order. The IPD between two countries is then simply the absolute difference between their two ideal points, with smaller values indicating closer geopolitical alignment. For example, the IPD between the United States and, respectively, Israel, France, Germany, Russia, and China, in our data is 0.2, 1.1, 1.4, 2.5, 3.1 (averages over sample years).

We include IPD in raw form and as natural logs into the models corresponding to Table 2, columns (3) and (4). In all four cases, physical distance remains statistically different from 0 at the 1% significance level, with parameter estimates ranging from -0.531*** to -0.649***. Conversely, none of the IPD variables comes out as statistically significant. These results, available in detail upon request, question the robustness of geopolitical alignment for bilateral capital allocations[23] and confirm the role of geographical distance in bilateral FDI positions.

[22] In our sample, ln(IPD) and ln(distance) show a correlation coefficient of 0.54.
[23] A possible reason for this discrepancy (among many others) may be the use of fDi Markets data by Aiyar et al. (2024) and Gopinath et al. (2025), which is mostly limited to greenfield FDI (or considerable expansions).

4.3 Separate regressions at sector level

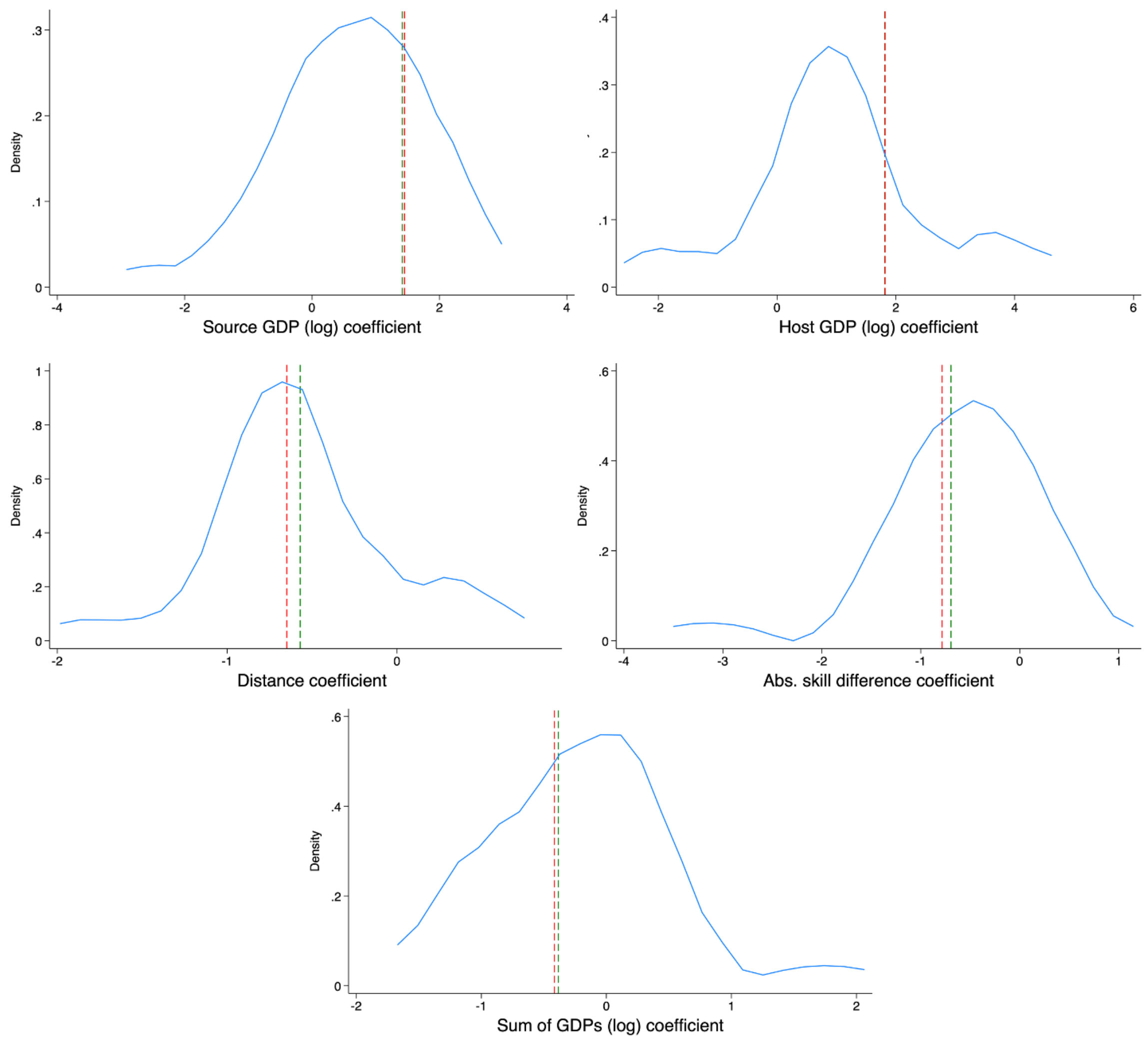
Estimating eq. (3) individually for each of the 23 sectors reveals substantial sector heterogeneity in the FDI determinants, especially in the distance elasticity.[24] Figure 3 illustrates this variation of parameters across sectors. It shows kernel density plots for the distribution of the total estimated main gravity coefficients for each of the 2-digit level sectors, together with the aggregate and pooled estimates as a red and green line, respectively. The plots visually display the key result of the section 4.1, as they demonstrate that the aggregate and pooled estimates align closely with one another. Additionally, confirming Breinlich et al. (2024), both estimates lie reasonably close to the mode of most sector-level coefficient distributions.

The sector-specific gravity regressions do not pinpoint any sub-sector where FDI clearly adheres to a purely horizontal or vertical motive. This result emerges when comparing the coefficients of the sector-specific regressions (Table B11) with the predictions in Table 1. This ambiguity surrounding the FDI motive, which remains when estimating at sector level, is likely due to methodological limitations. The parameters of the reduced-form gravity model do not provide perfect indicators for a clear-cut classification into HFDI and VFDI. Additionally, the 2-digit disaggregation of the dataset is still quite coarse, for instance masking intra-industry VFDI (Alfaro and Charlton, 2009). Finally, the mixed results may reflect the hybrid strategies of complex FDI, which are not captured by the binary horizontal-vertical framework.

Below, we discuss results for the distance elasticity, which, as expected, stands out as one of the elasticities with the greatest sectoral heterogeneity.

[24] The results of the 23 sector-specific regressions are presented in Table B11 in Appendix B.

Figure 3 / Distribution of key gravity parameters across sectors (kernel density plots)



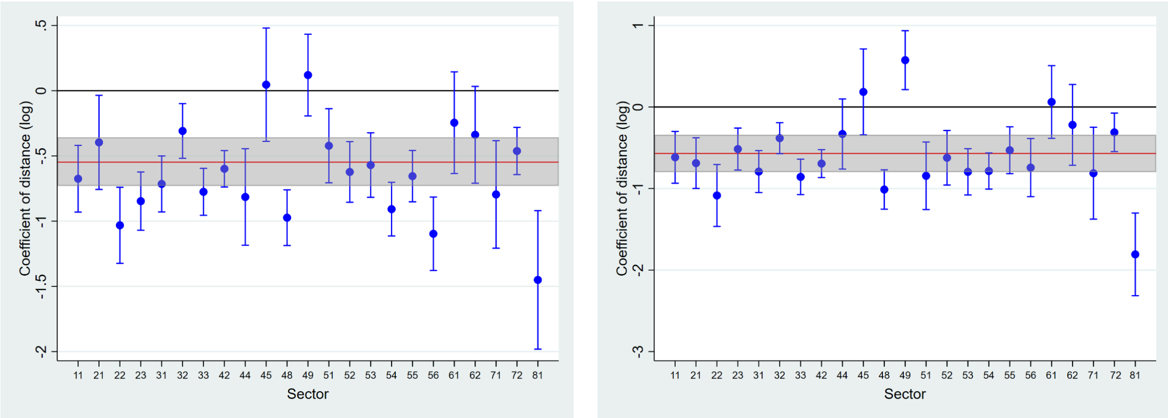
Notes: The figure shows the distribution of parameter estimates obtained from regression of eq. (3) performed for each of the 23 NAICS 2-digit sectors individually. The detailed regression results can be found in Table B11. The weighting function used in the kernel density estimation is Epanechnikov kernel. The red dashed line is the aggregate estimate from Table 2, col. (2), and the green line is the pooled estimate from Table 2, col (4).

4.4 The distance elasticity of FDI across sectors

Figure 4 compares the 23 sector-level coefficients to the pooled distance elasticity from Table 2. The sectoral elasticities are widely dispersed; this is obscured by the aggregate or pooled gravity model. Sectors with substantial deviations from the pooled estimate – those where the confidence intervals (CIs) do not overlap – are highlighted in green. Panel (a) shows the distance coefficients from the sparse regression, only with the main gravity variables. In this model, distance is likely to capture the combined effects of multiple related components, the relevance of which varies across sectors. The regressions in Panel (b) control for the variables that proxy for such components (e.g. for institutional distance), to examine whether the effect of geographical distance changes. As discussed in Section 4.1, at the aggregate and pooled levels, the distance effect remains highly robust to adding these variables. At the sector level, we observe some changes between Panel (a) and Panel (b), but only for a few sectors.

Figure 4 / Sector-specific coefficient of bilateral distance (log) with 95% CIs

(a) Pure model | (b) Full model, including other distance components



Notes: The coefficients are retrieved from regression of eq. (3) separately for each of the 23 NAICS 2-digit sectors. The sector classification can be found in Table A2. The regression in Panel (a) only includes the main KT (2010) gravity variables, the regression in Panel (b) includes all other distance-related components and the full set of controls. The pooled estimate is indicated by the red line, with the 95% CIs as the shaded area.

In Panel (b), sector 49, 'warehouse and storage', stands out with a positive distance coefficient. A vertical motive explains this positive effect: MNEs invest in local warehouses to manage decentralised inventories, reduce supply-chain risks and ensure faster deliveries. The further the target market is from the production site, the stronger the incentive for such vertical FDI. Consistent with prediction (3) in Section 0, in Panel (a) we see a more pronounced negative distance effect in the non-tradable sectors 22, 'utilities', 48, 'transportation', and 81, 'other services'. In these sectors, exporting is not viable and firms do not face the PCT, leading to an unambiguously negative distance effect. For instance, services such as electricity and water (Jenkins et al., 2011), and in-person activities such as repairs (sector 81) or transport (sector 48) must be provided locally, necessitating HFDI.

In the utilities sector, several factors add to the strong negative distance effect. High transport costs for resources such as natural gas deter VFDI, reinforcing the non-tradable nature of utilities (Knutsson and Flores, 2022). For horizontal FDI in utilities, fixed entry and adaptation costs are substantial, and tend to increase with distance. Setting up infrastructure from scratch involves high instalment and sunk costs (Bergara et al., 1998). Additionally, utilities FDI is frequently subject to regulatory barriers and investment screening, owing to its systemic importance (Rajavuori and Huhta, 2020). In turn, investors often demand high government accountability and political stability to mitigate risks of opportunistic behaviour by local governments (Mahbub and Jongwanich, 2019). This adds to the informational and organisational costs, which are partly reflected in geographical distance, explaining its negative effect on FDI. The negative effect weakens in Panel (b) when controlling for institutional distance and colonial ties, indicating that distance captures these factors.

Kolstad and Villanger (2008) identify a VFDI component in business services. The strong negative distance effect in the operational business services sector (sector 56) aligns with the theoretical prediction for VFDI. Also, the effect for knowledge-intensive business services (sector 54) shows a downward deviation from the pooled estimate. Controlling for colonial ties in Panel (b), which are particularly relevant for language- and knowledge-intensive sectors, substantially diminishes the negative effect for both business services sectors.

These results support our notion that the predicted positive distance effect from the PCT can be empirically rejected. However, KT's (2010) assumption that distance solely reflects entry costs and HQ-affiliate input linkages also appears too simplistic.

5. Are there systematic patterns in sector heterogeneity?

Our evidence for considerable sector heterogeneity in the distance elasticity raise the question to what extent economic factors can explain this heterogeneity. To investigate possible determinants of cross-sector heterogeneity in the estimated distance coefficients, we employ a two-stage estimation strategy: after having obtained the 23 NAICS 2-digit sector-specific distance coefficients from regression eq. (3), which are illustrated in Figure 4, panel (b), we regress those coefficients on four sector characteristics *H* that are explained below:

$$\hat{\delta}_s = a + \sum_{k=1}^{4} b_k H_{k,s} + e_s \qquad (5)$$

Such two-stage approaches are occasionally applied in the literature (e.g., MaCurdy, 1981; Couttenier et al. 2019) and recent work by Muris and Wacker (2026) highlights that it provides superior bias protection in the context of linear models than including interactions with variables *H* directly into regression eq. (3). We are not aware of similar work for interaction variables in the context of PPML models. In any case, the second-stage regression is estimated by OLS with heteroskedasticity-robust standard errors. Given the small number of sectors (*N* = 23, excluding Public Administration) and additional limitations in sector-specific *H* variables, the second-stage specifications are kept parsimonious, and we explore the four regressors individually and included simultaneously. These sector-specific *H* variables are:

- **Goods versus services production.** A binary indicator distinguishes goods-producing from services-producing sectors.[25] Some services are conventionally regarded as less tradable than goods and thus more prone to proximity-based FDI motives. On the other hand, the digitalisation of many service activities has contributed to significant growth in services trade over the past decade (Baldwin, 2022) and trade costs for such tradeable services are usually low. The goods dummy provides a first-pass test of whether tradability is systematically associated with stronger distance deterrence.

[25] Goods-producing sectors are defined as NAICS2 codes 11 (Agriculture, Forestry, Fishing and Hunting), 21 (Mining, Quarrying, and Oil and Gas Extraction), 22 (Utilities), 23 (Construction), 31 (Food and Textile Manufacturing), 32 (Materials Manufacturing), and 33 (Machinery and Equipment Manufacturing). All remaining sectors — NAICS2 codes 42 and above — are classified as services-producing.

- **Skill intensity.** A binary indicator distinguishes high-skill from low-skill sectors.[26] Skill intensity is particularly relevant in a sectoral FDI context because firms in low-skill-intensive industries tend to offshore labour-intensive production stages to countries with lower wage costs — a characteristically vertical FDI motive — while high-skill-intensive sectors are more likely to engage in horizontal FDI aimed at replicating knowledge-intensive activities in foreign markets. This variable also serves as a partial proxy for skill endowment complementarity at the sector level, addressing the lack of wage data disaggregated by sector and country pair in our dataset.
- **GVC upstreamness** based on a measure compiled by Mancini et al. (2024) that gauges the average number of production stages a sector is removed from final demand (see Antràs et al., 2012; Antràs and Chor 2013, 2019). We build sector-level averages of the data across all 82 available countries over the period 2010–2020 available by Mancini et al. (2024).[27] A higher value indicates that a sector's output is more intensively used as an intermediate input rather than consumed directly. Sectors with high upstreamness — such as mining and basic materials manufacturing — are hence more likely to be associated with vertical FDI, where distance may matter less because the investment responds to cost differentials rather than market proximity. Conversely, downstream sectors serving final consumers are expected to exhibit stronger distance sensitivity.
- **Sector-specific trade costs.** We employ the WTO Trade Cost Index (TCI) developed by Egger, Larch, Nigai, and Yotov (2021) as a measure of sector-specific trade costs.[28] The TCI captures the overall bilateral trade cost faced by a sector, including tariff and non-tariff barriers, and is constructed from a partial equilibrium gravity framework using combined WIOD and ADB inter-country input-output data. We use the 2018 cross-section — the most recent year available — collapsed to a sector-level average across all available economies. The TCI provides a theoretically grounded measure of trade cost barriers that complements the bilateral distance variable in the first stage: sectors facing higher trade costs are those for which the proximity-concentration motive for FDI is strongest, and where one may therefore expect distance deterrence to be more pronounced.

Table 3 presents the second-stage regression results. The most striking feature across all columns, representing different specifications, is that no simple economic explanation can make sense of sector-specific heterogeneity in the distance elasticity. The goods vs. services classification achieves the highest (but still low) R squared and is estimated to be negative in all specifications (columns) but

[26] High-skill sectors are defined as NAICS2 codes 33 (Machinery and Equipment Manufacturing, reflecting its engineering and precision-intensive character), 51 (Information), 52 (Finance and Insurance), 53 (Real Estate), 54 (Professional, Scientific, and Technical Services), 55 (Management of Companies and Enterprises), 61 (Educational Services), and 62 (Health Care and Social Assistance). This classification broadly follows Eurostat's Knowledge-Intensive Services (KIS) framework, adapted to NAICS2. All remaining sectors are classified as low-skill. The assignment of NAICS 33 is a borderline case: the sector spans both high-skill machinery and aerospace production and lower-skill metal fabrication.

[27] The concordance from ISIC Rev. 4 to NAICS2 is applied manually. Where multiple ISIC sectors map to a single NAICS2 code, simple averages are taken. TiVA sector 31 (Wholesale and retail trade; repair of motor vehicles) maps simultaneously to NAICS 42, 44, and 45, all of which therefore receive the same upstreamness value. NAICS 55 (Management of Companies) has no TiVA counterpart and is missing from this variable.

[28] The TCI covers 43 economies and 31 sectors classified according to ISIC Rev. 3.1. Since the TCI is not available at NAICS2 resolution, we apply a manual concordance from ISIC Rev. 3.1 to NAICS2. Where multiple ISIC sectors map to a single NAICS2 code — as is the case for NAICS 31 (ISIC 3 and 4&5), NAICS 32 (ISIC 6–11), NAICS 33 (ISIC 12–16), and NAICS 48 (ISIC 23–26) — we take simple averages of the ISIC-level TCI values. NAICS 22 (Utilities) has no ISIC counterpart in the TCI and remains missing. The TCI data are available at http://tradecosts.wto.org.

statistically indistinguishable from 0. The coefficient for skill intensity is also statistically insignificant and switches its prefix from positive in column (2) to negative in the multivariate specifications in columns (4) and (6). Upstreamness in GVCs and trade costs also do not seem to explain much of FDI's sector-specific heterogeneity in the distance elasticity. Interpreted at face value, however, the positive point estimate for upstreamness indicates that sectors further from final demand — and therefore more likely to engage in vertical rather than horizontal FDI — exhibit less negative distance coefficients. Such a result would be consistent with vertical FDI being driven by cost differentials that are not strongly attenuated by geographic distance.

Table 3 / Second stage regression for sector heterogeneity

	Dependent variable: Sector-specific distance elasticity (estimated in first stage)					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
goods (dummy)	-0.163			-0.177		-0.144
	(0.163)			(0.217)		(0.288)
high skill (dummy)		0.0282		-0.00930		-0.0304
		(0.184)		(0.231)		(0.292)
GVC upstreamness			0.0202	0.0495		0.139
			(0.196)	(0.208)		(0.335)
TCI (trade costs)					0.0478	0.0666
					(0.0683)	(0.132)
Constant	-0.543***	-0.602***	-0.635	-0.634	-0.728**	-1.019
	(0.139)	(0.143)	(0.407)	(0.463)	(0.318)	(1.269)
Observations	23	23	22	22	18	18
R-squared	0.026	0.001	0.001	0.029	0.013	0.048
R2	0.0259	0.000833	0.000592	0.0288	0.0132	0.0483

We also experimented with more complex second-stage specifications, none of which provided a clear result that would allow us to make clear economic sense of the sector heterogeneity of FDI's distance elasticity. The low explanatory power of all second stage regressions also underscores the multidimensional nature of sectoral heterogeneity in FDI gravity and, perhaps, the limitations inherent in a 23-observation cross-section.

To overcome this limitation in sample size, we also included sector groupings into our original bilateral annual sector data. In particular, we tested whether sector heterogeneities in the distance elasticity corresponds to the sector group considerations from Section 0. We therefore employed an interaction model with dummies for the following six sector groups. Group 1 consists of highly tradable manufacturing sectors (31-33). Group 2 contains sectors in which the business model centres on distance (warehouse and storage (49) and accommodation and food services (72)). These two groups are both related to prediction (1). Group 3 contains sectors with particularly high fixed set-up or adaptation costs (mining, oil and gas extraction (21) and utilities (22)), related to prediction (2 i). Group 4 contains the taste-driven arts, entertainment and recreation sector (71), related to prediction (2 ii). Group 5 contains business services sectors (54, 56), related to prediction (2 iii). Group 6 contains non-tradable in-person services sectors (transportation (48) and 'other services' (81)) as well as financial services (sector 52), related to prediction (3). We augmented the pooled sectoral model from eq. (3) to include a summation term which captures interaction terms between the distance variable D_{ij} and six dummy variables, $S_{k=1,\dots,6}$, one for each of the six sector groups.

This interaction model can be written as follows:

$$Y_{ij,t}^{s} = exp\left\{\boldsymbol{X}_{i,t}\beta + \boldsymbol{X}_{j,t}\pi + \boldsymbol{X}_{ij,t}\vartheta + \delta \ln\left(D_{ij}\right) + \sum_{k=1}^{6}(D_{ij} \times S_k)\lambda_k + \boldsymbol{Z}_{i,t}\theta + \boldsymbol{Z}_{j,t}\upsilon + \boldsymbol{Z}_{ij,t}\gamma + \boldsymbol{Z}_{ij}\zeta + \varphi_t + \varphi_i + \varphi_j + \varphi_s\right\} \times \varepsilon_{ij,t}, \tag{6}$$

where $\boldsymbol{X}_{i,t}$, $\boldsymbol{X}_{j,t}$, $\boldsymbol{X}_{ij,t}$ and D_{ij} contain the gravity variables, $\boldsymbol{Z}$ and φ contain the control variables and FEs, and λ_k captures the parameters of the interaction terms, and thus the differential effects of the distance variable for each of the six sector groups relative to the remaining sectors.

The regression results are reported in Table B12. An F-test on joint insignificance (Table B13) finds no systematic improvement in explanatory power compared with the reduced pooled gravity model without interactions. This suggests that the distance elasticity does not vary meaningfully across the six sector groups. The observed heterogeneity in Figure 4 hence appears to reflect rather random variation around the pooled estimate, with the exception of sector group 4 (arts and entertainment), where a negative interaction coefficient is significant at the 10% level. However, when the distance variable is interacted with 22 individual sector dummies (sector 81, 'other services', as base category), the F-test shows a systematic improvement in explanatory power.[29] This formal econometric test thus confirms systematic variation in the distance elasticity across individual sectors. This variation is exemplified by the Malaysian warehouse and storage sector in Section 1. Despite this evidence for sector-specific heterogeneity, further refinement of theoretical sector categories is needed to determine whether this variation aligns with theory.

Overall, the results presented in this section reveal that sector heterogeneity in the distance elasticity is present in bilateral FDI data, which is obscured in aggregate or pooled models. From a statistical perspective, the inclusion of sector-specific heterogeneity significantly improves model fit. Yet, we have also highlighted that no single economic reason can plausibly explain the heterogeneity of FDI's distance elasticity across sectors.

6. Conclusion

This paper examines whether aggregate or pooled gravity models mask sector heterogeneities in FDI motives, i.e. horizontal versus vertical FDI, and the distance elasticity of FDI. We anticipate sectoral variations in the distance elasticity, because the distance effect varies by FDI motive, and because the importance of different aspects covered broadly by distance is likely to differ across sectors. The study leverages the novel MREID by Ahmad et al. (2025), which provides operational FDI data at the 2-digit NAICS level for 145 countries over the period 2010-2020. Using this dataset, supplemented with gravity variables, the reduced-form gravity model of Kleinert and Toubal (2010) is estimated at three levels of aggregation: (i) aggregated at the country-pair level, (ii) at the sector level but with parameters constrained to homogeneity, and (iii) for all 23 sectors separately, allowing for full parameter

[29] The regression results are presented in Table B14 and the F-test in Table B15.

heterogeneity. The gravity model's testable predictions are used to identify the dominant FDI motive and to examine variations in the distance elasticity.

The central finding of this paper is that neither a clear horizontal nor vertical FDI motive emerges from the conventional aggregate or pooled sectoral models. The ambiguity in the baseline models stems primarily from parameter homogeneity, while any econometric bias from aggregating data at the country-pair level is negligible. Notably, enforcing homogeneity in the distance elasticity across sectors conceals significant heterogeneity. This likely heterogeneity arises from both underlying FDI motives and from other distance-related factors, such as cultural distance, which vary in importance across sectors. However, we could not prove that the heterogeneities follow a clear theory-driven pattern.

This paper makes two key contributions to existing research. First, it exploits a global, sector-level dataset, marking a decisive advance in meeting the data requirements necessary for identifying FDI motives. Second, it introduces a novel methodological approach that combines KT's (2010) gravity model with a systematic aggregation framework for gravity models based on recent findings from the trade literature. By addressing a critical gap in the FDI literature, which has traditionally been limited to aggregate-level analyses or narrow industry case studies (Hecock and Jepsen, 2014), this study provides a more nuanced and robust understanding of FDI determinants.

Sector-specific FDI motives should be seen as indicative tendencies rather than clear-cut classifications, owing to methodological limitations in identifying FDI motives, pointing to important future research paths. One limitation of KT's (2010) methodology is the strict distinction between HFDI and VFDI, which does not align well with the reality of hybrid strategies such as 'complex FDI'. Additionally, even within the horizontal-vertical dichotomy, the current model parameters lack precision in providing definitive classifications. The skill difference variable only captures traditional inter-industry VFDI, driven by large gaps in skilled labour, and does not account for comparative advantages in other factors, such as environmental regulation (e.g. Gerard and Chiappini, 2025). The model, therefore, misses non-traditional VFDI strategies relevant to certain sectors. Addressing these limitations requires more granular measures of comparative advantage, accounting for the factor intensities of sector-specific production technologies.

These challenges highlight the need for a theoretical refinement of gravity models for FDI. Sector-specific motives should be integrated into models by adding sector-level regressors, along with country-level variables, such as market size. Parameter predictions for these regressors must be revised to enable empirical parameter tests. Model refinements should further go beyond the horizontal-vertical dichotomy, for example by incorporating SMP.

Identifying FDI motives is further complicated by the limitations of MREID. The dataset specifies the sector of the affiliate, but not the parent firm, complicating a granular data-driven differentiation between HFDI and VFDI that relies on firm- or plant-level data. Accurate identification of intra-industry VFDI requires 4-digit level disaggregation and data on foreign subsidiaries' sales and input purchases (Alfaro and Charlton, 2009). Additionally, the MREID data's overreliance on finance and management sectors raises concern. Our sector-level approach helps to separate such sectors, while sector imbalances are obscured in aggregate data that is commonly used in previous studies.

Our paper finds a negative distance elasticity for both HFDI and VFDI, masking sector-specific variation. However, we could not reconcile these heterogeneities with sector groups based on a priori economic reasoning. Further refinement of theoretical predictions for sector group-specific effects is needed, alongside rigorous econometric model selection. Additionally, various distance-related components were controlled for in the regressions. Future research should refine gravity models to explicitly account for these components and estimate their sector-specific effects.

A key question for future research is whether the focus should shift from the sector level to the firm level when examining FDI motives, given the firm-specific as well as industry-specific nature of production technologies. Buch et al. (2005) identify firm heterogeneity in determinants of German outward FDI, for example in the host market size. A firm-level perspective could further enhance our understanding of the distance elasticity. Bricongne et al. (2023) show that, across French firms, the distance effect varies by product category. However, firm-level studies remain limited to a few source countries (Forte and Silva, 2017). Expanding global, publicly accessible firm-level FDI datasets could enable deeper insights into micro-level FDI dynamics.

The sector heterogeneities identified in this thesis express a clear need for sector-specific FDI policies. A sectoral approach throughout the investment lifecycle is crucial to align a country's comparative advantages with the determinants of FDI in each sector. For example, VFDI in business services requires a reduction of barriers to intra-firm trade, while HFDI in utilities benefits from policies that lower entry costs. Additionally, FDI's impact on the host economy varies by motive and sector (Beugelsdijk et al., 2008). Countries aiming for economic upgrading through technology spillovers and high-skilled job creation, for instance, should tailor policies to attract VFDI in knowledge-intensive sectors (Blalock and Gertler, 2008). Although institutions such as the World Bank traditionally emphasise a sectoral approach (e.g. World Bank, 2006), this remains rather disconnected from academic research. Our introductory example of Bangladesh-Malaysia versus Israel-Malaysia reveals how aggregate models can mislead predictions. Thus, there is an urgent need to develop a genuinely sectoral gravity model to guide evidence-based policy making, particularly in developing countries where FDI is a crucial development strategy (World Bank, 2024c).

References

Ahmad, S., Bergstrand, J., Paniagua, J. & Wickramarachi, H. (2025). The Multinational Revenue, Employment, and Investment Database (MREID). *Review of International Economics*, 33(4), 817-836.

Aleksynska, M. & Havrylchyk, O. (2013). FDI from the South: The role of institutional distance and natural resources. *European Journal of Political Economy*, 29, 38–53. https://doi.org/10.1016/j.ejpoleco.2012.09.001

Alfaro, L. & Charlton, A. (2009). Intra-industry foreign direct investment. *The American Economic Review*, 99(5), 2096-2119. https://doi.org/10.1257/aer.99.5.2096

Alfaro, L. & Chen, M.X. (2018). Transportation cost and the geography of foreign investment, in: Blonigen, B.A. & Wilson, W.W. (eds), *Handbook of international trade and transportation*. Edward Elgar Publishing, Ch. 12, 369-406.

Anderson, J.E. & van Wincoop, E. (2003). Gravity with gravitas. A solution to the border puzzle. *The American Economic Review,* 93(1), 170-192. https://doi.org/10.1257/000282803321455214

Antràs, P., Chor, D., Fally, T., & Hillberry, R. (2012), Measuring the Upstreamness of Production and Trade Flows. *American Economic Review Papers & Proceedings* 102(3): 412-416. https://doi.org/10.1257/aer.102.3.412

Antràs, P., & Chor, D. (2013). Organizing the Global Value Chain. *Econometrica* 81(6): 2127-2204. https://doi.org/10.3982/ECTA10813

Antràs, P. and Chor, D. (2019). On the Measurement of Upstreamness and Downstreamness in Global Value Chains. In L. Y. Ing and M. Yu (Eds.), World Trade Evolution: Growth, Productivity and Employment, Chapter 5, pp. 126–194. Routledge.

Aiyar, S., Malacrino, D., & Presbitero, A.F. (2024). Investing in friends: The role of geopolitical alignment in FDI flows. *European Journal of Political Economy*, 83: 102508. https://doi.org/10.1016/j.ejpoleco.2024.102508

Aykut, D., Sanghi, A. & Kosmidou, G. (2017). What to do when foreign direct investment is not direct or foreign: FDI round tripping. World Bank Policy Research Working Paper No. 8046. https://ssrn.com/abstract=2960424

Badinger, H. & Egger, P. (2010). Horizontal vs. vertical interdependence in multinational activity. *Oxford Bulletin of Economics and Statistics*, 72(6), 744-768. https://doi.org/10.1111/j.1468-0084.2010.00600.x

Baier, F.J. & Welfens, P.J.J. (2019). The UK's banking FDI flows and total British FDI: A dynamic Brexit analysis. *International Economics and Economic Policy*, 16(1), 193-213. https://doi.org/10.1007/s10368-018-00426-x

Bailey, M. A., Strezhnev, A., & Voeten, E. (2017). Estimating dynamic state preferences from United Nations voting data. *Journal of Conflict Resolution*, 61(2):430–456. https://doi.org/10.1177/0022002715595700

Baldwin, R. (2022). Globotics and Macroeconomics: Globalisation and Automation of the Service Sector. NBER Working Paper No. 30317. National Bureau of Economic Research. https://doi.org/10.3386/w30317

Barro, R.J. & Lee, J.W. (2013). A new data set of educational attainment in the world, 1950-2010. *Journal of Development Economics*, 104, 184-198. https://doi.org/10.1016/j.jdeveco.2012.10.001

Bénassy-Quéré, A., Coupet, M. & Mayer, T. (2007). Institutional determinants of foreign direct investment. *The World Economy*, 30(5), 764-782. https://doi.org/10.1111/j.1467-9701.2007.01022.x

Benoit, K. (2011). Linear regression models with logarithmic transformations. Methodology Institute, London School of Economics. Retrieved 10 January 2025, from https://kenbenoit.net/assets/courses/ME104/logmodels2.pdf

Bergara, M.E., Henisz, W.J. & Spiller, P.T. (1998). Political institutions and electric utility investment: A cross-nation analysis. *California Management Review*, 40(2), 18-35. https://doi.org/ 10.2307/41165931

Beugelsdijk, S., Smeets, R. & Zwinkels, R. (2008). The impact of horizontal and vertical FDI on host's country economic growth. *International Business Review*, 17(4), 452-472. https://doi.org/ 10.1016/j.ibusrev.2008.02.004

Beugelsdijk, S., Hennart, J-F., Slangen, A. & Smeets, R. (2010). Why and how FDI stocks are a biased measure of MNE affiliate activity. *Journal of International Business Studies*, 41(9), 1444-1459. https://doi.org/10.1057/jibs.2010.29

Beugelsdijk, S., Kostova, T., & Roth, K. (2017). An overview of Hofstede-inspired country-level culture research in international business since 2006. *Journal of International Business Studies*, 48(1), 30–47. https://doi.org/10.1057/s41267-016-0038-8

Beugelsdijk, S., Kostova, T., Kunst, V.E., Spadafora, E. & van Essen, M. (2018). Cultural distance and firm internationalization: A meta-analytical review and theoretical implications. *Journal of Management*, 44(1), 89-130. https://doi.org/10.1177/0149206317729027

Bilir, L.K., Chor, D. & Manova, K. (2019). Host-country financial development and multinational activity. *European Economic Review*, 115, 192-220. https://doi.org/10.1016/j.euroecorev.2019.02.008

Blalock, G. & Gertler, P.J. (2008). Welfare gains from foreign direct investment through technology transfer to local suppliers. *Journal of International Economics*, 74(2), 402-421. https://doi.org/10.1016/j.jinteco.2007.05.011

Blonigen, B.A. (1997). Firm-specific assets and the link between exchange rates and foreign direct investment. *The American Economic Review*, 87(3), 447-465. https://www.jstor.org/stable/2951354

Blonigen, B.A., Davies, R.B. & Head, K. (2003). Estimating the knowledge-capital model of the multinational enterprise: Comment. *The American Economic Review*, 93(3), 980-994. https://doi.org/10.1257/000282803322157214

Blonigen, B.A. & Piger, J. (2014). Determinants of foreign direct investment. *Canadian Journal of Economics*, 47(3), 775-812. https://doi.org/10.1111/caje.12091

Borchert, I., Larch, M., Shikher, S. & Yotov, Y.V. (2022). Disaggregated gravity: Benchmark estimates and stylized facts from a new database. *Review of International Economics*, 30(1), 113-136. https://doi.org/ 10.1111/roie.12555

Braconier, H., Norbäck, P-J. & Urban, D. (2005a). Reconciling the evidence on the knowledge-capital model. *Review of International Economics*, 13(4), 770-786. https://doi.org/10.1111/j.1467-9396.2005.00536.x

Braconier, H., Norbäck, P-J. & Urban, D. (2005b). Multinational enterprises and wage costs: Vertical FDI revisited. *Journal of International Economics*, 67(2), 446-470. https://doi.org/10.1016/j.jinteco.2004.08.011

Brainard, S.L. (1993). A simple theory of multinational corporations and trade with a trade-off between proximity and concentration. National Bureau of Economic Research (NBER) Working Paper 4269.

Brainard, S.L. (1997). An empirical assessment of the proximity-concentration trade-off between multinational sales and trade. *The American Economic Review*, 87(4), 520-544.

Breinlich, H., Leromain, E., Novy, D. & Sampson, T. (2020). Voting with their money: Brexit and outward investment by UK firms. *European Economic Review*, 124: 103400. https://doi.org/10.1016/j.euroecorev.2020.103400

Breinlich H., Novy, D. & Santos Silva, J.M.C. (2024). Trade, gravity, and aggregation. *The Review of Economics and Statistics*, 106(5), 1418-1426. https://doi.org/10.1162/rest_a_01234

Bricongne, J-C., Bedoya, S.F. & Forero, M.L. (2023). The proximity-concentration trade-off with multi-product firms: Are exports and FDI complements or substitutes? *The World Economy*, 46(5), 1264-1289. https://doi.org/10.1111/twec.13339

Buch, C.M., Kleinert, J., Lipponer, A., Toubal, F. & Baldwin, R. (2005). Determinants and effects of foreign direct investment: Evidence from German firm-level data. *Economic Policy*, 20(41), 52-110. https://doi.org/10.1111/j.1468-0327.2005.00133.x

Carballo, J., Artiñano, I.M. & Martincus, C.V. (2021). Information frictions, investment promotion, and multinational production: Firm-level evidence. CESifo Working Paper No. 9043. https://ssrn.com/abstract=3837771

Carr, D.L., Markusen, J.R. & Maskus, K.E. (2001). Estimating the knowledge-capital model of the multinational enterprise. *The American Economic Review*, 91(3), 693-708. https://doi.org/10.1257/aer.91.3.693

Castellacci, F. (2014). Service innovation and the proximity-concentration trade-off model of trade and FDI. *Economics of Innovation and New Technology*, 23(1), 92-108. http://dx.doi.org/10.1080/10438599.2013.828890

Castellani, D., Jimenez, A. & Zanfei, A. (2013). How remote are R&D labs? Distance factors and international innovative activities. *Journal of International Business Studies*, 44(7), 649-675. https://doi.rg/10.1057/jibs.2013.30

Cezar, R. & Escobar, O.R. (2015). Institutional distance and foreign direct investment. *Review of World Economics*, 151(4), 713-733. https://doi.org/10.1007/s10290-015-0227-8

Clarke, D. & Tapia Schythe, K. (2020). Implementing the panel event study. IZA Discussion Paper No. 13524, IZA Institute of Labor Economics. https://docs.iza.org/dp13524.pdf

Conte, M., Cotterlaz, P. & Mayer, T. (2022). The CEPII Gravity database. CEPII Working Paper No. 2022-05. https://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=8

Correia, S., Guimarães, P. & Zylkin, T. (2020). Ppmlhdfe. Fast Poisson estimation with high-dimensional fixed effects. *The Stata Journal*, 20(1), 95-115. https://doi.org/10.1177/1536867X20909691

Couttenier, M., Petrencu, V, Rohner, D. & Thoenig, M. (2019). The Violent Legacy of Conflict: Evidence on Asylum Seekers, Crime, and Public Policy in Switzerland. *American Economic Review* 109(12): 4378–4425. https://doi.org/10.1257/aer.20170263

Crowley, M., Exton, O. & Han, L. (2020). Renegotiation of trade agreements and firm exporting decisions: Evidence from the impact of Brexit on UK exports. Cambridge Working Papers in Economics: 1839. https://www.inet.econ.cam.ac.uk/working-paper-pdfs/wp1810.pdf

Damgaard J., Elkjaer, T. & Johannesen, N. (2024). What is real and what is not in the global FDI network? *Journal of International Money and Finance*, 140: 102971. https://doi.org/10.1016/j.jimonfin.2023.102971

Daude, C., & Fratzscher, M. (2008). The pecking order of cross-border investment. *Journal of International Economics*, *74*(1), 94–119.

Davies, R.B. (2008). Hunting high and low for vertical FDI. *Review of International Economics*, 16(2), 250-267. https://doi.org/10.1111/j.1467-9396.2008.00735.x

Davies, R.B. & Markusen, J.R. (2021). What do multinationals do? The structure of multinational firms' international activities. *The World Economy*, 44(12), 3444-3481. https://doi.org/10.1111/twec.13199

di Giovanni, J. (2005). What drives capital flows? The case of cross-border M&A activity and financial deepening. *Journal of International Economics*, 65(1), 127-149. https://doi.org/ 10.1016/j.jinteco.2003.11.007

Donaubauer, J., Neumayer, E. & Nunnenkamp, P. (2021). Financial market development in host and source countries and their effects on bilateral foreign direct investment. *The World Economy*, 43(3), 534-556. https://doi.org/10.1111/twec.12884

Dorakh, A. (2020). A gravity model analysis of FDI across EU member states. *Journal of Economic Integration*, 35(3), 426-456. https://www.jstor.org/stable/26927873

Dossani, R. & Kenney, M. (2007). The next wave of globalization: Relocating service provision to India. *World Development*, 35(5), 772-791. https://doi.org/10.1016/j.worlddev.2006.09.014

Egger, P.. & Pfaffermayr, M. (2004). Distance, trade and FDI: A Hausman-Taylor SUR Approach. *Journal of Applied Econometrics*, 19(2): 227-246.

Egger, P.H. & Nigai, S. (2015). Structural gravity with dummies only: Constrained ANOVA-type estimation of gravity models. *Journal of International Economics*, 97(1), 86-99. https://doi.org/10.1016/j.jinteco.2015.05.004

Egger, P.H. & Tarlea, F. (2015). Multi-way clustering estimation of standard errors in gravity models. *Economics Letters*, 134, 144-147. https://doi.org/10.1016/j.econlet.2015.06.023

Egger, P. H., Larch, M., Nigai, S., & Yotov, Y. V. (2021). Trade Costs in the Global Economy: Measurement, Aggregation and Decomposition. WTO Staff Working Paper ERSD-2021-2. World Trade Organization. https://doi.org/10.30875/e6c4c0b1-en

Ekholm, K., Forslid, R. & Markusen, J.R. (2007). Export-platform foreign direct investment. *Journal of the European Economic Association*, 5(4), 776-795.

Executive Office of the President, Office of Management and Budget (2017). *North American Industry Classification System* [Report]. Retrieved 5 September 2024, from https://www.census.gov/naics/reference_files_tools/2017_NAICS_Manual.pdf

Falk, M. (2016). A gravity model of foreign direct investment in the hospitality industry. *Tourism Management*, 55, 225-237. https://doi.org/10.1016/j.tourman.2016.02.012

Feenstra, R.C., Inklaar, R. & Timmer, M.P. (2015). The next generation of the Penn World Table. *The American Economic Review,* 105(10), 3150-3182.

Fontagné, L., Guimbard, H. & Orefice, G. (2022). Tariff-based product-level trade elasticities. *Journal of International Economics*, 137: 103953. https://doi.org/10.1016/j.jinteco.2022.103593

Forte, R. & Silva, V. (2017). Outward FDI and home country exports: Theoretical approaches and empirical evidence. *The International Trade Journal*, 31(3), 245-271. https://doi.org/10.1080/08853908.2016.1271731

Fraser Institute (2024). Economic Freedom Index. Retrieved 15 November 2024, from https://efotw.org/economic-freedom/dataset

Frenkel, M. & Stefan, H. (2024). The impact of the brexit referendum on greenfield foreign direct investments from countries within and outside the European single market. *Applied Economics Letters*, 31(12), 1125-1128. https://doi.org/10.1080/13504851.2023.2176437

Froot, K.A. & Stein, J.C. (1991). Exchange rates and foreign direct investment: An imperfect capital markets approach. *The Quarterly Journal of Economics*, 106(4), 1191-1217. https://doi.org/10.2307/2937961

Fukao, K. & Wei, Y. (2008). How do the location determinants of vertical FDI and horizontal FDI differ? Hi-Stat Discussion Paper No. 233, Hitotsubashi University Research Unit for Statistical Analysis in Social Sciences. https://hi-stat.ier.hit-u.ac.jp/research/discussion/2007/pdf/D07-233.pdf

Gerard, E. & Chiappini, R. (2025). Environmental regulation and foreign direct investments: Evidence from a new measure of environmental stringency. *Macroeconomic Dynamics*, 29, e136. https://doi.org/10.1017/S1365100525100436

Gopinath, G., Gourinchas, P.-O., Presbitero, A. F., & Topalova, P. (2025). Changing global linkages: A new cold war? *Journal of International Economics*, 153:104042. https://doi.org/10.1016/j.jinteco.2024.104042

Gregori, W.D. & Nardo, M. (2021). The effect of restrictive measures on cross-border investment in the European Union. *The World Economy*, 44(7), 1914-1943. https://doi.org/10.1111/twec.13104

Harding, T. & Javorcik, B.S. (2011). Roll out the red carpet and they will come: Investment promotion and FDI inflows. *The Economic Journal*, 121(557), 1445-1476. https://doi.org/10.1111/j.1468-0297.2011.02454.x

Harms, P. & Knaze, J. (2021). Bilateral de-jure exchange rate regimes and foreign direct investment: A gravity analysis. *Journal of International Money and Finance*, 117: 102438. https://doi.org/ 10.1016/j.jimonfin.2021.102438

Hashimoto, Y. & Wacker, K.M. (2016). The role of information for international capital flows: New evidence from the SDDS. *Review of World Economics*, 152(3), 529-557. https://doi.org/10.1007/s10290-016-0250-4

Head, K. & Ries, J. (2008). FDI as an outcome of the market for corporate control: Theory and evidence. *Journal of International Economics*, 74(1), 2-20. https://doi.org/10.1016/j.jinteco.2007.04.004

Head, K. & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook, in: Gopinath, G., Helpman, E. & Rogoff, K. (eds), *Handbook of International Economics*, *Vol. 4*. Elsevier, Ch. 3, 131-195. https://doi.org/10.1016/B978-0-444-54314-1.00003-3

Hecock, R.D. & Jepsen, E.M. (2014). The political economy of FDI in Latin America 1986-2006: A sector-specific approach. *Studies in Comparative International Development*, 49(4), 426-447. https://doi.org/10.1007/s12116-013-9143-x

Helpman, E. (1984). A simple theory of international trade with multinational corporations. *Journal of Political Economy*, 92(3), 451-471. https://doi.org/10.1086/261236

Helpman, E., Melitz, M.J. & Yeaple, S.R. (2004). Export versus FDI with heterogeneous firms. *The American Economic Review*, 94(1), 300-316. https://doi.org/ 10.1257/000282804322970814

Helpman, E., Melitz, M. & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics*, 123(2), 441-487. https://doi.org/10.1162/qjec.2008.123.2.441

Hutzschenreuter, T., Kleindienst, I. & Lange, S. (2016). The concept of distance in international business research: A review and research agenda. *International Journal of Management Reviews*, 18(2), 160-179. https://doi.org/10.1111/ijmr.12065

Iliopoulos, P., Ioannou, S. & Wójcik, D. (2024). The City of London after Brexit: Sticky power in the global financial network. *Progress in Economic Geography*, 2(1): 100011. https://doi.org/10.1016/j.peg.2024.100011

Imbs, J. & Mejean, I. (2015). Elasticity optimism. *American Economic Journal: Macroeconomics*, 7(3), 43-83. http://dx.doi.org/10.1257/mac.20130231

International Monetary Fund (2024a). Financial Development Index Database. Retrieved 10 September 2024, from https://data.imf.org/?sk=f8032e80-b36c-43b1-ac26-493c5b1cd33b

International Monetary Fund (2024b). Glossary of foreign direct investment terms. Retrieved 12 September 2024, from https://www.imf.org/external/np/sta/di/glossary.pdf

Irarrazabal, A., Moxnes, A. & Opromolla, L.D. (2013). The margins of multinational production and the role of intrafirm trade. *Journal of Political Economy*, 121(1), 74-126. https://doi.org/10.1086/669877

Jenkins, G.P., Kuo, C-Y. & Harberger, A.C. (2011). Economic prices for non-tradable goods and services. JDI Executive Programs, Development Discussion Paper 2011-11. https://cri-world.com/publications/qed_dp_204.pdf

Kalaitzake, M. (2022). Resilience in the City of London: The fate of UK financial services after Brexit. *New Political Economy*, 27(4), 610-628. https://doi.org/10.1080/13563467.2021.1994540

Kaplan, D.M. (2023). Introductory Economics. Description, prediction, and causality. Retrieved 10 January 2025, from https://kaplandm.github.io/Kaplan_Intro_Econometrics_text.pdf

Keller, W. & Yeaple, S.R. (2013). The gravity of knowledge. *The American Economic Review*, 103(4), 1414-1444. http://dx.doi.org/10.1257/aer.103.4.1414

Kleinert, J. & Toubal, F. (2010). Gravity for FDI. *Review of International Economics*, 18(1), 1-13. https://doi.org/10.1111/j.1467-9396.2009.00869.x

Knutsson, P. & Flores, P.I. (2022). Trends, investor types and drivers of renewable energy FDI. OECD Working Papers on International Investments, No. 2022/02, OECD Publishing. https://doi.org/10.1787/4390289d-en

Kolstad, I. & Villanger, E. (2008). Determinants of foreign direct investment in services. *European Journal of Political Economy*, 24(2), 518-533. https://doi.org/10.1016/j.ejpoleco.2007.09.001

Kox, H.L.M. & Rojas-Romagosa, H. (2020). How trade and investment agreements affect bilateral foreign direct investment: Results from a structural gravity model. *The World Economy,* 43(12), 3203-3242. https://doi.org/10.1111/twec.13002

Krautheim, S. (2013). Export-supporting FDI. *Canadian Journal of Economics/Revue Canadienne d'Économique*, 46(4), 1571–1605.

Lane, P.R. & Milesi-Ferretti, G.M. (2018). The external wealth of nations revisited: International financial integration in the aftermath of the global financial crisis. *IMF Economic Review*, 66(1), 189-222. https://doi.org/10.1057/s41308-017-0048-y

Lankhuizen, M., de Groot, H.L.F. & Linders, G-J.M. (2011). The trade-off between foreign direct investments and exports: The role of multiple dimensions of distance. The World Economy, 34(8), 1395-1416. https://doi.org/10.1111/j.1467-9701.2011.01335.x

Lee, B-S. & Min, B. S. (2011). Exchange rates and FDI strategies of multinational enterprises. *Pacific-Basin Finance Journal*, 19(5), 586-603. https://doi.org/10.1016/j.pacfin.2011.05.003

Ly, A., Esperança, J. & Davcik, N.S. (2018). What drives foreign direct investment: The role of language, geographical distance, information flows and technological similarity. *Journal of Business Research*, 88, 111-122. https://doi.org/10.1016/j.jbusres.2018.03.007

MaCurdy, T. E. (1981). An Empirical Model of Labor Supply in a Life-Cycle Setting. *Journal of Political Economy* 89(6): 1059–1085. https://doi.org/10.1086/261023

Mahbub, T. & Jongwanich, J. (2019). Determinants of foreign direct investment in the power sector: A case study of Bangladesh. *Energy Strategy Reviews*, 24, 178-192. https://doi.org/ 10.1016/j.esr.2019.03.001

Mancini M., Montalbano P., Nenci S. & Vurchio D., (2024). Positioning in Global Value Chains: World Map and Indicators, a new dataset available for GVC analyses. *World Bank Economic Review*, 38(4): 669–690. https://doi.org/10.1093/wber/lhae005

Markusen, J.R. (1984). Multinationals, multi-plant economies, and the gains from trade. *Journal of International Economics,* 16(3-4), 205-226. https://doi.org/10.1016/S0022-1996(84)80001-X

Markusen, J.R. (2002). *Multinational firms and the theory of international trade*. MIT Press, Cambridge, Massachusetts.

Markusen, J.R. & Maskus, K.E. (2002). Discriminating among alternative theories of the multinational enterprise. *Review of International Economics*, 10(4), 694-707. https://doi.org/10.1111/1467-9396.00359

Markusen, J.R. & Strand, B. (2009). Adapting the knowledge-capital model of the multinational enterprise to trade and investment in business services. *The World Economy*, 32(1), 6-29. https://doi.org/10.1111/j.1467-9701.2009.01155.x

Markusen, J.R. (2013). Multinational firms, in: Bernhofen, D., Falvey, R., Greenaway, D. & Kreickemeier, U. (eds), *Palgrave Handbook of International Trade*. Palgrave Macmillan, London and New York, Ch. 8, 236-262.

McCann, P. (2011). International business and economic geography: knowledge, time and transactions costs. *Journal of Economic Geography*, 11(2): 309-317.

Miller, D.L. (2023). An introductory guide to event study models. *The Journal of Economic Perspectives*, 37(2), 203-230. https://doi.org/10.1257/jep.37.2.203

Muris, C. & Wacker, K.M. (2026). Estimating Interaction Effects With Panel Data. *Journal of Applied Econometrics*, forthcoming. https://doi.org/10.1002/jae.70040

Navaretti, G.B. & Venables, A.J. (eds) (2004). *Multinational firms in the world economy*. Princeton University Press, Princeton, New Jersey.

Neary, J.P. (2009). Trade costs and foreign direct investment. *International Review of Economics and Finance*, 18(2), 207-218. https://doi.org/10.1016/j.iref.2008.06.004

Nunn, N. (2007). Relationship-Specificity, Incomplete contracts and the Pattern of Trade. *The Quarterly Journal of Economics*, 122(2), 569-600. https://doi.org/10.1162/qjec.122.2.569

Orcutt, G. (1950). Measurement of price elasticities in international trade. *The Review of Economics and Statistics*, 32(2), 117-132.

Organisation for Economic Co-operation and Development (2020). Recommendation of the Council on the OECD benchmark definition of foreign direct investment. OECD/LEGAL/0363.

Ottaviano, G.I.P. & Turrini, A. (2007). Distance and Foreign Direct Investment When Contracts Are Incomplete. *Journal of the European Economic Association*, 5(4): 796–822, https://doi.org/10.1162/JEEA.2007.5.4.796

Portes, R., Rey, H., & Oh, Y. (2001). Information and capital flows: The determinants of transactions in financial assets. *European Economic Review*, 45(4–6), 783–796.

Rajavuori, M. & Huhta, K. (2020). Investment screening: Implications for the energy sector and energy security. *Energy Policy*, 144: 111646. https://doi.org/10.1016/j.enpol.2020.111646

Ramasamy, B. & Yeung, M. (2010). The determinants of foreign direct investment in services. *The World Economy*, 33(4), 573-596. https://doi.org/10.1111/j.1467-9701.2009.01256.x

Ramondo, N., Rappoport, V. & Ruhl, K.J. (2011). Horizontal versus vertical FDI: New empirical evidence about U.S. multinationals. 2011 Meeting Papers 654, Society for Economic Dynamics.

Ramondo, N., Rodríguez-Clare, A. & Tintelnot, F. (2015). Multinational production: Data and stylized facts. *The American Economic Review*, 105(5), 530-536. http://dx.doi.org/10.1257/aer.p20151046

Rezza, A.A. (2013). FDI and pollution havens: Evidence from the Norwegian manufacturing sector. *Ecological Economics*, 90, 140-149. https://doi.org/10.1016/j.ecolecon.2013.03.014

Riedl, A. (2010). Location factors of FDI and the growing services economy. *Economics of Transition*, 18(4), 741-761. https://doi.org/10.1111/j.1468-0351.2010.00391.x

Santos Silva, J.M.C. & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4), 641-658. https://doi.org/10.1162/rest.88.4.641

Schneider, S.T. & Wacker, K.M. (2022). Explaining the global landscape of foreign direct investment: Knowledge capital, gravity, and the role of culture and institutions. *The World Economy*, 45(10), 3080-3108. https://doi.org/10.1111/twec.13267

Schreiber, S. (2022). A practical approach to estimating sector-level substitution elasticites with PPML. Economics Working Paper 2022-03-A, U.S. International Trade Commission Office of Economics. https://www.usitc.gov/publications/332/working_papers/2022-03-a_estimating_substitution_elasticities_with_ppml.pdf

Serwicka, I. & Tamberi, N. (2018). Not backing Britain: FDI inflows since the Brexit referendum. Briefing Paper No. 23, UK Trade Policy Observatory.

Sirr, G., Garvey, J. & Gallagher, L.A. (2017). Bilateral investment treaties and foreign direct investment: Evidence of asymmetric effects on vertical and horizontal investments. *Development Policy Review*, 35(1), 93-113. https://doi.org/10.1111/dpr.12202

UNCTAD (2011). World Investment Report: Non-equity modes of international production and development. Retrieved 5 October 2024, from https://unctad.org/system/files/official-document/wir2011_en.pdf

van Hoorn, A., & Maseland, R. (2016). How institutions matter for international business: Institutional distance effects vs institutional profile effects. *Journal of International Business Studies*, 47(3), 374–381. https://doi.org/10.1057/jibs.2016.2

Venables, A.J. (1999). Fragmentation and multinational production. *European Economic Review*, 43(4-6), 935-945.

Wacker, K.M. (2013). On the measurement of foreign direct investment and its relationship to activities of multinational corporations. Working Paper No. 1614, European Central Bank.

Wacker, K.M., Casella, B. & Borga, M. (2025). Measuring multinational production with foreign direct investment statistics. Recent trends, challenges, and developments. *Journal of Economic Surveys*, 39(4), 1462-1487. https://doi.org/10.1111/joes.12662

Wooldridge, J.M. (2020). *Introductory Economics. A modern approach* (7th edn). Cengage Learning, Boston, Massachusetts.

World Bank (2006). Snapshot Africa. Benchmarking FDI competitiveness in Sub-Saharan African countries. World Bank Group/MIGA, Washington, DC. https://documents1.worldbank.org/curated/en/465861468212386977/pdf/389050AFR0Snapshot1africa01PUBLIC1.pdf

World Bank (2024a). Data bank: World Development Indicators. Retrieved 10 September 2024, from https://databank.worldbank.org/reports.aspx?source=2&series=SH.DYN.MORT&country=WLD

World Bank (2024b). Data Catalog: World Bank Official Boundaries. Retrieved 22 November 2024, from https://datacatalog.worldbank.org/search/dataset/0038272

World Bank (2024c). Investment policy and promotion: Attracting foreign direct investment and maximizing impact for the local economy. https://www.worldbank.org/content/dam/Worldbank/document/Trade/InvestClimate_InvestPolicyPromotion.pdf

Yeaple, S.R. (2003a). The role of skill endowments in the structure of US outward foreign direct investment. *The Review of Economics and Statistics*, 85(3), 726-734.

Yeaple, S.R. (2003b). The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment. *Journal of International Economics*, 60(2), 293-314. https://doi.org/10.1016/S0022-1996(02)00051-X

Yeaple, S.R. & Antràs, P. (2014). Multinational Firms and the Structure of International Trade, in: Gopinath, G., Helpman, E. & Rogoff, K. (eds), *Handbook of International Economics*, *Vol. 4.* Elsevier, Ch. 2, 55-130.

Yotov, Y.V., Piermartini, R., Monteiro, J. & Larch, M. (2016). *An advanced guide to trade policy analysis. The structural gravity model.* World Trade Organization, Geneva. https://doi.org/10.30875/abc0167e-en

Appendix

APPENDIX A: DATA AND DESCRIPTIVES

Table A1 / Comparison of the two traditional motives for FDI

		Horizontal FDI	Vertical FDI
	Strategy	Duplicate activities at the same production stage abroad (with the same factor intensity of production)	Geographically fragment activities by factor intensities
	Example	Duplicate assembly in a foreign country	Produce a specific intermediate in a plant abroad
Trade-off	**Benefits**	› Firm-level economies of scale › Better market access › Economise on trade costs › Strategic advantage over local competitors	› Lower production costs by exploiting factor endowment differences
Trade-off	**Costs**	› Plant-level economies of scale forgone	› Economies of integration forgone › Intra-firm trade and transaction costs

Notes: Compilation based on Navaretti and Venables (2004).

Table A2 / Sector coverage of the MREID

NAICS 2-digit code	Sector name	Share of total assets (%)	Share of observations (%)
11	Agriculture, forestry, fishing and hunting	0.11	2.29
21	Mining, oil and gas extraction	2.47	2.69
22	Utilities	1.69	3.09
23	Construction	0.80	4.99
31	Food and textile manufacturing	1.41	4.67
32	Materials manufacturing	2.57	6.24
33	Finished product manufacturing	4.29	6.63
42	Wholesale trade	3.16	9.20
44	Food and beverage stores	0.44	4.02
45	Miscellaneous store retailers	0.20	2.73
48	Transportation	0.70	5.07
49	Warehouse and storage	0.07	2.00
51	Information	1.69	4.32
52	Finance and insurance	5.31	6.46
53	Real estate	2.36	5.53
54	Professional, scientific and technical services	2.20	7.91
55	Management of companies	27.43	4.71
56	Administrative and support, and waste management and remediation services	2.12	5.54
61	Educational services	0.02	1.45
62	Health care and social assistance	0.18	1.97
71	Arts, entertainment and recreation	0.22	1.71
72	Accommodation and food services	0.26	0.31
81	Other services (except public administration)	0.31	3.30
92	Public administration	0.20	0.39

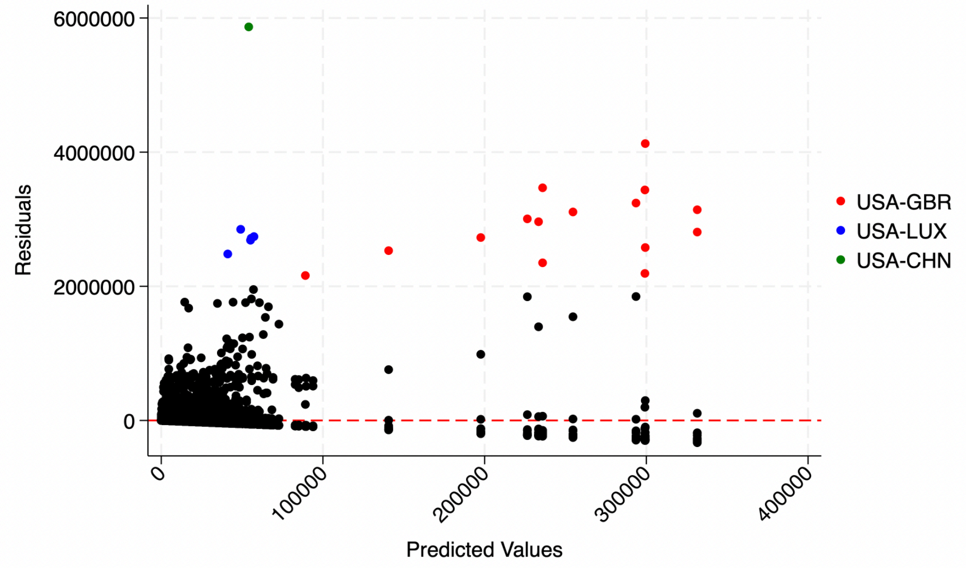
Notes: Taken from Executive Office of the President, Office of Management and Budget (2024). Primary sector in orange; secondary sector in blue; tertiary sector in green.

Table A3 / Description of variables and their sources

Variable	Definition	Source
Total assets	Total assets, i.e. the sum of current and fixed assets, including intangibles (in mn USD)	MREID dataset Release 1 (Ahmad et al., 2025).
Revenue	Operating revenue turnover, excluding taxes (in mn USD)	MREID
Gravity variables		
GDP	Real GDP at constant 2015 mn USD	World Development Indicators database (World Bank, 2024a).
Distance	Population-weighted distance between most populated cities (harmonic mean)	CEPII gravity dataset 202211 version (Conte et al., 2022).
Skill endowment difference	Skill endowment levels are measured by the Human Capital Index. Source index is subtracted from host index. The absolute value of this difference is used.	Penn World Table 10.01 (Feenstra et al., 2015). Values for 2020 are forecast based on a country-specific linear trend.
Surrounding market potential	The surrounding market potential is calculated as the sum of the GDPs of all neighbouring countries, weighted by the inverse of their distances, excluding the source and host countries, for each year.	GDP from the World Development Indicators and distance from the CEPII gravity dataset
Control variables for other dimensions of distance		
Tariff	Weighted mean applied tariff of the host country	World Development Indicators database
Common coloniser (post-1945)	Dummy variable equal to one if a pair had a common coloniser after 1945	CEPII gravity dataset
Common official language	Dummy variable equal to one if a pair shares the same official language	CEPII gravity dataset
Colonial relationship (post-1945)	Dummy variable equal to one if a pair had a colonial relationship after 1945	CEPII gravity dataset
Contiguity	Dummy variable equal to one if a pair has a common border	CEPII gravity dataset
Institutional distance	Measured as the difference in the Economic Freedom Summary Index, which is based on 26 components in the five areas of size of government, legal system, sound money, freedom to trade and regulation. Source index is subtracted from host index.	Fraser Institute (2024)
International finance variables		
Exchange rate	Official exchange rate in local currency per USD calculated as an annual average based on monthly averages. The XR is defined as local currency units per USD. Hence, the variable rises with local currency depreciation.	World Development Indicators database
De jure exchange rate regime	De jure exchange rate regime based on the IMF's Annual Report on Exchange Arrangements and Exchange Restrictions. The index ranges from 1 to 10 (1 = hard peg; 10 = free floating)	Harms and Knaze (2021).
Distance in financial development	Based on the Financial Development Index by the IMF (2024a), which ranks countries according to the depth, access, and efficiency of their financial institutions and markets. The source index is subtracted from the host index.	Financial Development Index Database (IMF, 2024a).

Notes: Own compilation.

Figure A1 / Residuals plotted against predicted values from PPML regression with pooled sectoral data with inclusion of full controls



Notes: Observations marked as outliers if the corresponding residual exceeded the value of 2,000,000.

Table A4 / Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Total assets (mn USD)	239,381	2,223.686	40,335.498	0	5,921,893
Operating revenue (mn USD)	239,381	588.347	5,074.737	0	1,113,036.3
Source GDP (USD m)	239,381	2,357,310.5	4,395,426.9	1,240.786	20,136,688
Host GDP (USD m)	239,381	1,606,296.3	3,484,076.6	1,986.736	20,136,688
Weighted distance	239,381	5,070.22	4,518.032	55	19,571
Absolute skill difference	239,381	.499	.413	0	2.731
Sum of GDPs (log)	239,381	14.444	1.273	8.834	17.355
Surr. market potential (log)	239,381	9.72	.66	.273	10.821
Common coloniser	239,381	.028	.165	0	1
Common language	239,381	.148	.355	0	1
In colonial relationship (post-1945)	239,381	.034	.18	0	1
Contiguity	239,381	.099	.299	0	1
XR host country	239,271	360.203	2,445.437	.276	42,000
XR source country	239,379	136.088	1,400.618	.276	42,000
De jure XR regime	226,431	8.186	3.261	1	10
Dist. in financial development	239,175	-.094	.284	-.878	.847
Dist. in economic freedom	239,194	.208	1.037	-4.99	4.81
Host country tariff (%)	227,924	2.653	2.086	0	23.97

Notes: The data sources of the different variables are listed in Table A3.

Table A5 / Pairwise correlation matrix of regression variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
(1) Total assets	1.000																
(2) Source GDP (log)	0.052	1.000															
(3) Host GDP (log)	0.038	-0.028	1.000														
(4) Distance (log)	0.002	0.277	0.228	1.000													
(5) Abs. skill difference	-0.019	0.141	0.070	0.301	1.000												
(6) Sum of GDPs (log)	0.057	0.702	0.548	0.391	0.196	1.000											
(7) SMP (log)	0.010	-0.118	0.004	-0.371	-0.267	-0.157	1.000										
(8) Common coloniser	-0.006	-0.173	-0.118	-0.089	0.046	-0.208	-0.126	1.000									
(9) Colonial relationship	0.003	-0.055	0.017	0.043	0.139	0.029	-0.053	-0.032	1.000								
(10) Common language	0.038	-0.033	0.044	-0.005	0.071	0.030	-0.190	0.167	0.228	1.000							
(11) Contiguity	0.009	-0.079	-0.011	-0.488	-0.168	-0.092	0.064	0.017	0.015	0.200	1.000						
(12) XR host country	-0.006	0.037	-0.048	0.059	0.075	-0.019	-0.090	-0.018	0.005	-0.045	-0.025	1.000					
(13) XR source country	-0.004	-0.024	0.040	0.023	0.062	0.002	-0.011	-0.007	0.004	-0.025	-0.005	0.004	1.000				
(14) De jure XR regime	0.006	0.157	0.138	0.411	0.130	0.184	-0.202	-0.053	0.048	0.009	-0.177	0.060	0.041	1.000			
(15) Dist. in fin. dev.	0.005	-0.397	0.532	-0.024	-0.054	0.015	0.189	0.064	0.067	0.089	0.081	-0.136	0.099	-0.042	1.000		
(16) Dist. in economic freedom	-0.006	0.057	-0.098	0.043	0.079	-0.007	-0.271	-0.042	-0.021	-0.052	-0.034	0.203	-0.133	0.062	-0.600	1.000	
(17) Host country tariff (log)	-0.009	0.023	0.095	0.067	0.153	0.093	-0.250	-0.079	0.038	-0.077	0.014	0.073	-0.034	0.128	-0.196	0.464	1.000

Notes: The data sources of the different variables are listed in

Table A3.

Table A6 / List of countries in dataset, ISO-3-code in parentheses

East Asia and Pacific	Europe and Central Asia	Latin America and the Caribbean	Middle East and North Africa	North America	South Asia	Sub-Saharan Africa
Australia (AUS)	Albania (ALB)	Aruba (ABW)*	United Arab Emirates (ARE)	Canada (CAN)	Bangladesh (BGD)	Angola (AGO)
Brunei (BRN) – *only source*	Andorra (AND)*	Anguilla (AIA)*	Bahrain (BHR)	United States (USA)	India (IND)	Burundi (BDI) – *only source*
China (CHN)	Armenia (ARM)	Argentina (ARG)	Djibouti (DJI)*		Sri Lanka (LKA)	Benin (BEN) – *only source*
Fiji (FJI)	Austria (AUT)	Antigua and Barbuda (ATG)*	Algeria (DZA)		Nepal (NPL)*	Burkina Faso (BFA) – *only source*
Hong Kong (HKG)	Azerbaijan (AZE)	Bahamas (BHS)*	Egypt (EGY)			Botswana (BWA)
Indonesia (IDN)	Belgium (BEL)	Belize (BLZ)	Iran (IRN)			Central African Republic (CAF) – *only source*
Japan (JPN)	Bulgaria (BGR)	Bermuda (BMU)*	Iraq (IRQ)			Côte d'Ivoire (CIV)
Cambodia (KHM)	Bosnia and Herzegovina (BIH)*	Bolivia (BOL)	Israel (ISR)			Cameroon (CMR) – *only host*
South Korea (KOR)	Belarus (BLR)*	Brazil (BRA)	Jordan (JOR)			Democratic Republic of Congo (COD)
Laos (LAO)	Switzerland (CHE)	Barbados (BRB)	Kuwait (KWT)			Congo (COG) – *only host*
Macao (MAC)	Cyprus (CYP)	Chile (CHL)	Lebanon (LBN)*			Cape Verde (CPV)*
Maldives (MDV) – *only host*	Czechia (CZE)	Colombia (COL)	Libya (LBY)*			Ethiopia (ETH) – *only host*
Marshall Islands* (MHL)	Germany (DEU)	Costa Rica (CRI)	Morocco (MAR)			Gabon (GAB)
Myanmar (MMR) – *only source*	Denmark (DNK)	Cuba (CUB)*	Oman (OMN)*			Ghana (GHA) – *only host*
Mongolia (MNG)	Spain (ESP)	Curaçao (CUW)*	Pakistan (PAK)			Guinea (GIN)*
Malaysia (MYS)	Estonia (EST)	Cayman Islands (CYM)*	Palestine (PSE)*			Gambia (GMB)
New Zealand (NZL)	Finland (FIN)	Dominica (DMA)*	Qatar (QAT)			Guinea-Bissau (GNB)
Philippines (PHL)	France (FRA)	Dominican Republic (DOM)	Saudi Arabia (SAU)			Kenya (KEN)
Papua New Guinea (PNG)	United Kingdom (GBR)	Ecuador (ECU)	Sudan (SDN) – *only host*			Liberia (LBR)
North Korea (PRK) – *only source*	Georgia (GEO)	Grenada (GRD)	South Sudan (SSD)*			Lesotho (LSO) – *only host*
Singapore (SGP)	Gibraltar (GIB)*	Guatemala (GTM)	Syria (SYR)			Madagascar (MDG)
Thailand (THA)	Greece (GRC)	Guyana (GUY) – *only source*	Tunisia (TUN)			Mali (MLI) – *only host*
Taiwan (TWN)*	Croatia (HRV)	Honduras (HND)				Mozambique (MOZ) – *only source*

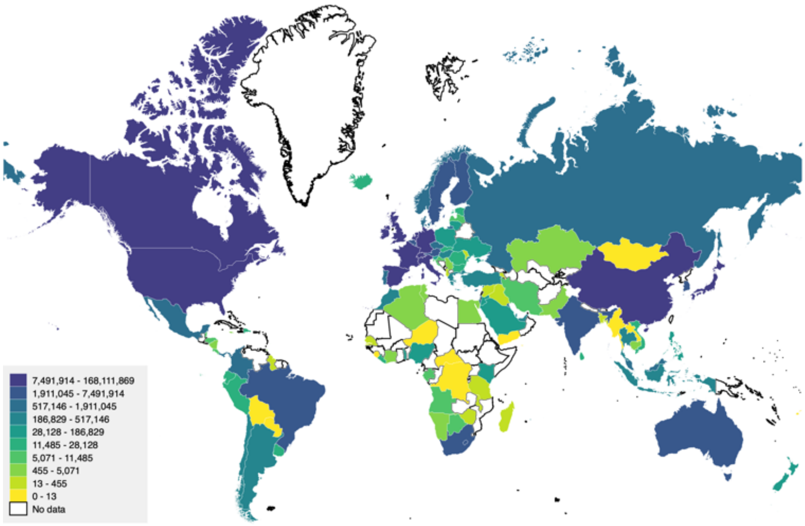
Table A6 / Continued

East Asia and Pacific	Europe and Central Asia	Latin America and the Caribbean	Middle East and North Africa	North America	South Asia	Sub-Saharan Africa
Vietnam (VNM)	Hungary (HUN)	Jamaica (JAM)				Mauritania (MRT) – *only source*
Samoa (WSM)*	Ireland (IRL)	Saint Kitts and Nevis (KNA)*				Mauritius (MUS)
	Iceland (ISL)	Saint Lucia (LCA)*				Malawi (MWI)
	Italy (ITA)	Mexico (MEX)				Namibia (NAM)
	Kazakhstan (KAZ)	Nicaragua (NGA)				Niger (NER)
	Kyrgyz Republic (KGZ) – *only source*	Panama (PAN)				Nigeria (NGA)
	Liechtenstein (LIE)*	Peru (PER)				Rwanda (RWA) – *only host*
	Lithuania (LTU)	Paraguay (PRY) – *only source*				Senegal (SEN)
	Luxembourg (LUX)	El Salvador (SLV) – *only host*				Sierra Leone (SLE)
	Latvia (LVA)	Trinidad and Tobago (TTO)				São Tomé and Príncipe (STP)*
	Monaco (MCO)*	Uruguay (URY)				Swaziland (SWZ) – *only host*
	Moldova (MDA)	Saint Vincent and the Grenadines (VCT)*				Seychelles (SYC)
	North Macedonia (MKD)*	Venezuela (VEN)*				Chad (TCD)*
	Malta (MLT)	British Virgin Islands (VGB)*				Togo (TGO)
	Montenegro (MNE)*					Tanzania (TZA)
	Netherlands (NLD)					Uganda (UGA) – *only host*
	Norway (NOR)					South Africa (ZAF)
	Poland (POL)					Zambia (ZMB) – *only host*
	Portugal (PRT)					Zimbabwe (ZWE)
	Romania (ROU)					
	Russia (RUS)					
	San Marino (SMR)*					
	Slovakia (SVK)					
	Slovenia (SVN)					
	Sweden (SWE)					
	Turkmenistan (TKM)*					
	Turkey (TUR)					
	Ukraine (UKR)					
	Uzbekistan (UZB)*					

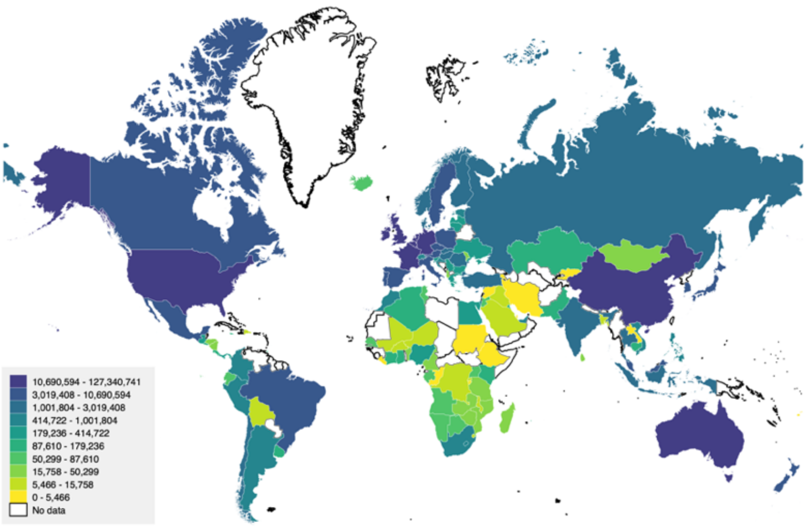
Notes: * indicates countries that are included in the MREID but not in the final dataset, as at least one of the gravity variables has a missing value for this country.

Figure A2 / Heat map of global pattern of FDI (total assets in mn USD) covered by the MREID

Panel (a): By FDI source country



Panel (b): By FDI host country



Notes: Panel (a) presents the global distribution of total assets by FDI source country; Panel (b) by host country. Total assets are summed over the period 2010-2020.
Data sources: MREID (Ahmad et al., 2025); World Bank Official Boundaries (World Bank, 2024b).

Descriptive analysis

Figure A3 presents the top 10 source and host countries for FDI in the sample, with the value of total assets averaged over the sample period. The Sankey plot in The top source and the top host countries are predominantly major industrialised nations, with the US the largest source and the UK the largest host country. However, the graphs also highlight the outsized role of several small offshore financial centres in the global FDI network. For instance, Luxembourg, with a population of less than a million, records an FDI inward position of USD 3.2trn, surpassing those of Germany and China (USD 2.1trn and USD 1.6trn, respectively). Investments of this magnitude are unlikely to reflect productive assets employed within Luxembourg's small domestic economy. Instead, they are likely to represent examples of 'phantom FDI': investments originate from an ultimate investor and pass through shell entities in low-tax jurisdictions, such as Luxembourg, Ireland and the Netherlands, without contributing to real activity, before ending up in the host country (Damgaard et al., 2024). Another form of phantom FDI is 'round-tripping', as seen when Chinese investments are routed through Hong Kong and back into China to profit from preferential treatment granted to foreign investors (Aykut et al., 2017).

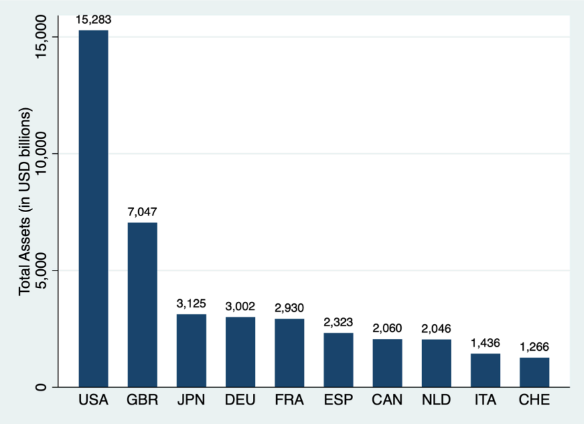
The MREID tries to minimise the share of phantom FDI in the data through two approaches. First, it seeks to exclude dormant affiliates by including only those with turnover or total assets over USD 1mn in at least one year of the sample period. Second, when aggregating firm-level data with explicit ownership details from Orbis, the MREID focuses on FDI from the ultimate investor, rather than the immediate investor. This approach minimises the share of multinational activity driven by the addition of layers of empty corporate shells in ownership chains, ensuring a focus on real activity (Ahmad et al., 2025).

To add greater scrutiny to this approach, and to avoid distortions in examining global FDI determinants, we identify individual outliers by estimating the standard pooled gravity model in eq. (3), then plotting the residuals against the predicted values in Figure A1, and visually pinpointing country pairs with the most significant deviations from the point cloud – unsurprisingly, US-UK, US-Luxembourg and US-China. Following Schneider and Wacker (2022), we account for these outliers by including FEs for these country pairs in our regressions.

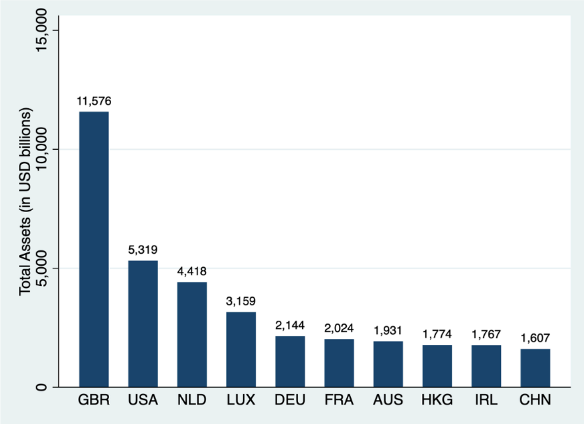
Figure A4 shows the country pairs with the largest bilateral FDI positions, where the width of the connecting lines reflects the volume of total assets invested between each source-host pair over time. The largest average FDI position is invested from the US in UK affiliates, followed by US-Netherlands and US-Luxembourg. Total assets in the dataset are concentrated in the sectors 'finance and insurance' (45.31%) and 'management of companies' (27.43%).

Figure A3 / Top 10 source and host FDI countries by total assets (bn USD)

(a) Top 10 source countries by total assets



(b) Top 10 host countries by total assets



Notes: Total assets (in bn USD) are summed across source and host countries, respectively, and averaged over the period 2010-2020. Panel (a) presents the top 10 source countries, and Panel (b) the top 10 host countries.
Data source: MREID (Ahmad et al., 2025).

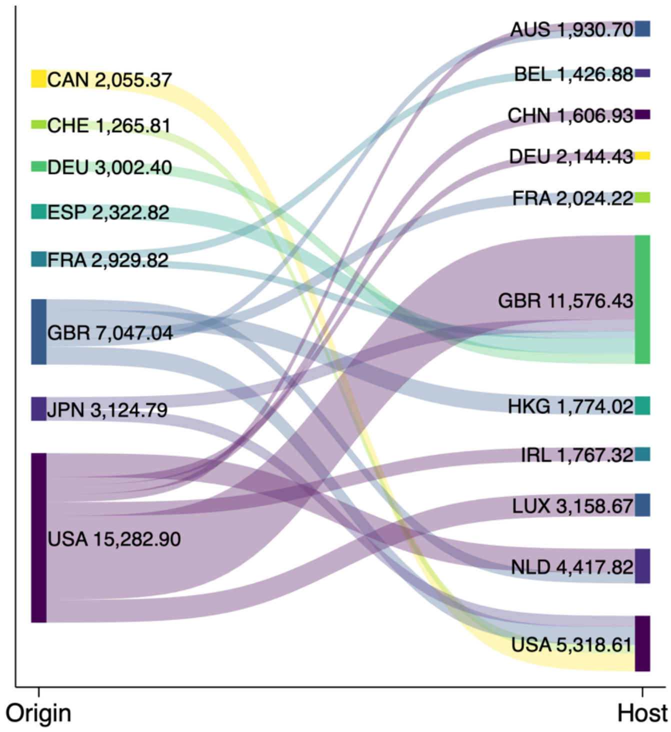
The top source and the top host countries are predominantly major industrialised nations, with the US the largest source and the UK the largest host country. However, the graphs also highlight the outsized role of several small offshore financial centres in the global FDI network. For instance, Luxembourg, with a population of less than a million, records an FDI inward position of USD 3.2trn, surpassing those of Germany and China (USD 2.1trn and USD 1.6trn, respectively). Investments of this magnitude are unlikely to reflect productive assets employed within Luxembourg's small domestic economy. Instead, they are likely to represent examples of 'phantom FDI': investments originate from an ultimate investor and pass through shell entities in low-tax jurisdictions, such as Luxembourg, Ireland and the Netherlands, without contributing to real activity, before ending up in the host country (Damgaard et al., 2024). Another form of phantom FDI is 'round-tripping', as seen when Chinese investments are routed through Hong Kong and back into China to profit from preferential treatment granted to foreign investors (Aykut et al., 2017).

The MREID tries to minimise the share of phantom FDI in the data through two approaches. First, it seeks to exclude dormant affiliates by including only those with turnover or total assets over USD 1mn in at least one year of the sample period. Second, when aggregating firm-level data with explicit ownership details from Orbis, the MREID focuses on FDI from the ultimate investor, rather than the immediate investor. This approach minimises the share of multinational activity driven by the addition of layers of empty corporate shells in ownership chains, ensuring a focus on real activity (Ahmad et al., 2025).[30]

[30] Although the dataset lags behind approaches such as that of Damgaard et al. (2024), which are able to distinguish between investment in special purpose entities (SPEs) and non-SPEs, the selection strategy can be considered successful. Offshore financial centres such as the Cayman Islands and the British Virgin Islands do not appear among the largest FDI countries in the raw dataset.

To add greater scrutiny to this approach, and to avoid distortions in examining global FDI determinants, we identify individual outliers by estimating the standard pooled gravity model in eq. (3), then plotting the residuals against the predicted values in Figure A1, and visually pinpointing country pairs with the most significant deviations from the point cloud – unsurprisingly, US-UK, US-Luxembourg and US-China. Following Schneider and Wacker (2022), we account for these outliers by including FEs for these country pairs in our regressions.

Figure A4 / Sankey plot with the top 20 country pairs with the largest bilateral FDI positions



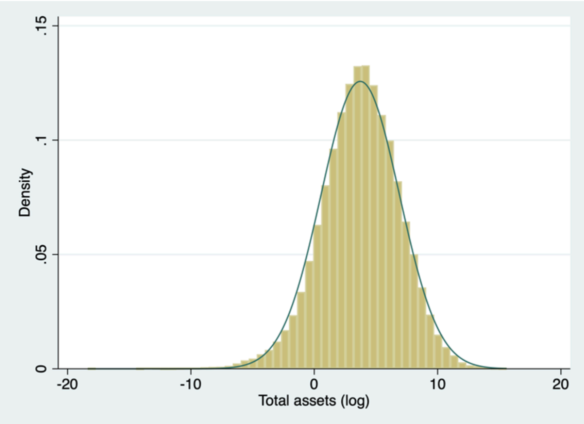
Notes: Largest bilateral FDI positions are calculated by total assets (in bn USD), averaged over 2010-2020. Next to the Iso3-Code of the origin (host) country, the outward (inward) FDI stock of this country is presented, expressed as the average of total assets over 2010-2020 (in bn USD).
Data source: MREID (Ahmad et al., 2025).

APPENDIX B: ADDITIONAL ESTIMATION RESULTS

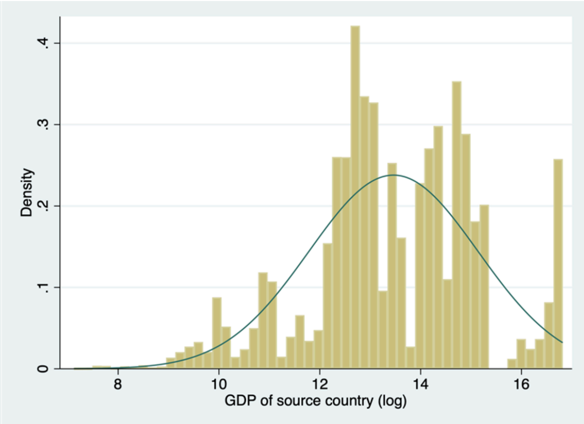
Histograms

Figure B1 / Histograms of the dependent variable and the main explanatory variables

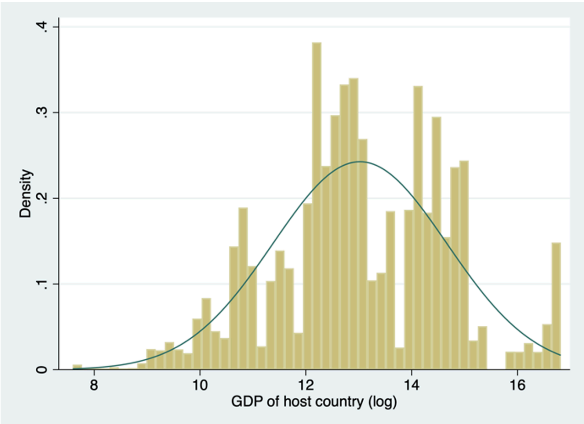
Panel (a): Total assets (log)



Panel (b): GDP of source country (log)



Panel (c): GDP of host country (log)



Panel (d): Bilateral distance (log)

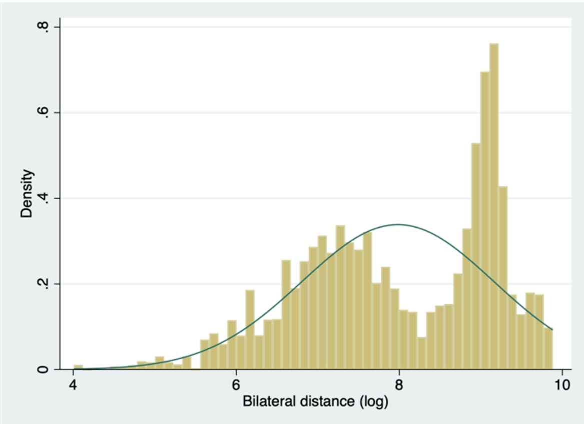
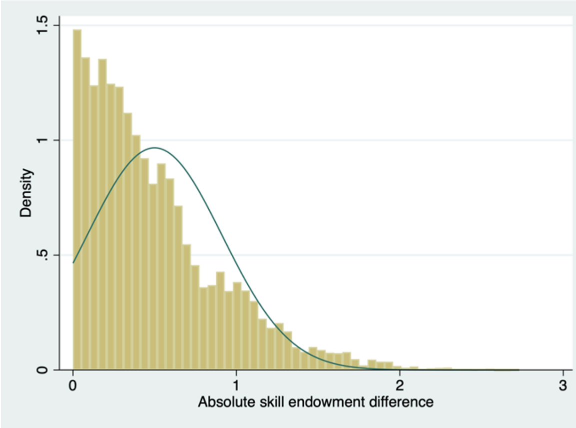
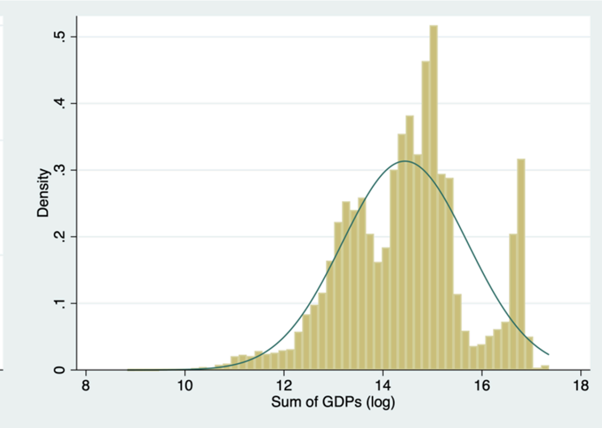


Figure B1 / Continued

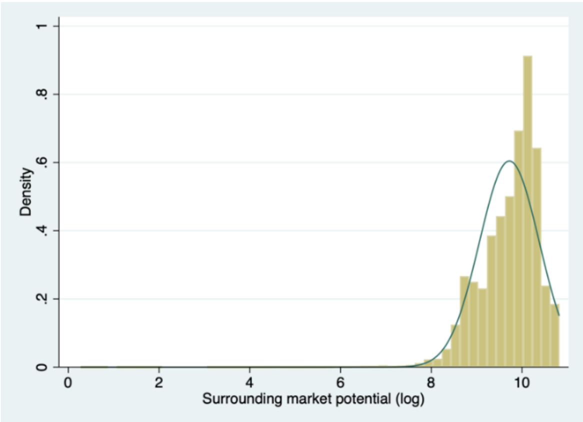
Panel (e): Absolute skill endowment difference



Panel (f): Sum of GDPs (log)



Panel (g): Surrounding market potential (log)



Notes: Histograms of the dependent variable (Panel a) and the main explanatory variables (Panels b-g) included in the gravity regressions. Normal distribution in green.
Data source: MREID (Ahmad et al., 2025).

Robustness checks

Table B1 / Pooled sectoral model with total assets deflated by the US GDP deflator

DV: Total assets (deflated)	(1) NAICS 2-digit level	(2) NAICS 2-digit level
Source GDP (log)	1.296**	1.416**
	(0.547)	*(0.653)*
Host GDP (log)	2.003***	1.817***
	(0.330)	*(0.358)*
Distance (log)	-0.545***	-0.567***
	(0.0974)	*(0.118)*
(abs.) Skill endowment diff.	-0.776***	-0.788***
	(0.206)	*(0.237)*
Sum of GDPs (log)	-0.333	-0.420*
	(0.215)	*(0.243)*
Surr. market potential (log)	0.848*	1.966***
	(0.443)	*(0.528)*
Common coloniser		0.261
		(0.395)
Colonial relationship		0.146
		(0.325)
Common off. language		0.205
		(0.222)
Contiguity		0.428
		(0.269)
XR of host country		9.10e-05
		(5.92e-05)
XR of source country		3.28e-06
		(3.72e-05)
XR regime (de jure)		0.0464
		(0.0308)
Dist. in financial development		-0.837
		(0.750)
Dist. Economic Freedom Index		-0.123
		(0.104)
Host country tariff (log)		-0.0441
		(0.0504)
Constant	-35.46***	-44.82***
	(10.98)	*(13.34)*
N	239,276	214,155
Pseudo R2	0.746	0.763

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Regressions are the same as in Table 2, Panel (b), but with total assets deflated by the US GDP deflator (using the pl_gdpo series of the PWT). Regressions include additive year, host country, source country, sector and outlier pair FEs.

Table B2 / Gradual addition of control variables

DV: Total assets	(1) NAICS 2-digit level	(2) NAICS 2-digit level	(3) NAICS 2-digit level	(4) NAICS 2-digit level
Source GDP (log)	1.292**	1.258**	1.316**	1.417**
	(0.538)	*(0.562)*	*(0.562)*	*(0.642)*
Host GDP (log)	1.996***	1.914***	1.942***	1.811***
	(0.323)	*(0.355)*	*(0.357)*	*(0.353)*
Distance (log)	-0.548***	-0.536***	-0.482***	-0.570***
	(0.0969)	*(0.114)*	*(0.116)*	*(0.117)*
(abs.) Skill endowment diff.	-0.771***	-0.851***	-0.848***	-0.785***
	(0.204)	*(0.223)*	*(0.234)*	*(0.235)*
Sum of GDPs (log)	-0.327	-0.361	-0.458*	-0.416*
	(0.215)	*(0.221)*	*(0.248)*	*(0.242)*
Surr. market potential (log)	0.864*	0.741	1.015*	2.005***
	(0.448)	*(0.456)*	*(0.528)*	*(0.530)*
Common coloniser		3.04e-05	4.54e-05	9.48e-05
		(5.72e-05)	*(5.94e-05)*	*(5.88e-05)*
Colonial relationship		5.84e-05	5.90e-05	5.57e-06
		(6.30e-05)	*(6.37e-05)*	*(3.76e-05)*
Common off. language		0.0360	0.0631**	0.0471
		(0.0296)	*(0.0315)*	*(0.0306)*
Contiguity			0.0611	0.253
			(0.591)	*(0.391)*
XR of host country			1.231***	0.150
			(0.373)	*(0.324)*
XR of source country			0.271	0.203
			(0.225)	*(0.220)*
XR regime (de jure)			0.444	0.432
			(0.271)	*(0.267)*
Dist. in financial development				-0.853
				(0.735)
Dist. Economic Freedom Index				-0.118
				(0.104)
Host country tariff (log)				-0.0410
				(0.0491)
Constant	-35.54***	-32.56***	-35.86***	-45.18***
	(10.86)	*(11.49)*	*(11.88)*	*(13.20)*
N	239,276	226,325	226,325	214,155
Pseudo R2	0.747	0.750	0.757	0.765

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Regressions are estimated using eq. (3), while stepwise adding the control variables. All regressions include additive year, host country, source country, sector and outlier pair FEs. N decreases owing to missing values in some explanatory and control variables. Additional observations are automatically dropped from the PPML estimation if singletons.

Table B3 / Aggregate and pooled sectoral gravity regressions from Table 2 with country-pair FEs

DV: Total assets	Panel (a): Aggregate model (1) Aggregated at country-pair level	(2) Aggregated at country pair level With controls	Panel (b): Pooled sectoral model (3) NAICS 2-digit level	(4) NAICS 2-digit level With controls
Source GDP (log)	1.407**	1.589**	1.406**	1.585**
	(0.556)	*(0.617)*	*(0.554)*	*(0.615)*
Host GDP (log)	2.069***	1.887***	2.067***	1.886***
	(0.357)	*(0.389)*	*(0.358)*	*(0.390)*
(abs.) Skill endowment diff.	-0.506	-1.313**	-0.500	-1.304**
	(0.323)	*(0.589)*	*(0.324)*	*(0.589)*
Sum of GDPs (log)	-0.788	-0.941	-0.781	-0.932
	(0.650)	*(0.817)*	*(0.649)*	*(0.816)*
Surr. market potential	-0.341	-0.860	-0.333	-0.836
	(0.578)	*(1.042)*	*(0.580)*	*(1.037)*
XR of host country		-4.44e-05		-4.18e-05
		(8.14e-05)		*(8.16e-05)*
XR of source country		1.60e-05		1.59e-05
		(4.14e-05)		*(4.11e-05)*
XR regime (de jure)		0.00837		0.00845
		(0.0536)		*(0.0536)*
Dist. in financial development		-0.522		-0.518
		(0.684)		*(0.684)*
Dist. Economic Freedom Index		-0.142		-0.142
		(0.100)		*(0.1000)*
Host country tariff (log)		-0.0562		-0.0558
		(0.0556)		*(0.0555)*
Constant		-14.19	-23.47**	-16.19
		(15.77)	*(11.49)*	*(15.70)*
N	31,822	27,266	227,703	203,710
Pseudo R2	0.975	0.975	0.834	0.838

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Panel (a) regressions are estimated using eq. (2), and Panel (b) regressions are estimated using eq. (3). Columns (1) and (3) present the pure model only with the KT (2010) variables and SMP. Columns (2) and (4) present regressions that include all control variables. Panel (a) regressions include additive year and country-pair FEs which absorb the host and source country FEs and the outlier FEs. Panel (b) regressions additionally include additive sector FEs. N decreases owing to missing values in some explanatory and control variables. Additional observations are automatically dropped from the PPML estimation if singletons.

Table B4 / Aggregate and pooled sectoral gravity regressions from Table 2 with revenue as dependent variable

DV: Revenue	Panel (a): Aggregate model (1) Aggregated at country-pair level	(2) Aggregated at country-pair level With controls	Panel (b): Pooled sectoral model (3) NAICS 2-digit	(4) NAICS 2-digit With controls
Source GDP (log)	1.055**	0.771***	1.106**	0.776***
	(0.476)	*(0.256)*	*(0.476)*	*(0.252)*
Host GDP (log)	1.131***	0.790**	1.177***	0.827**
	(0.252)	*(0.372)*	*(0.248)*	*(0.365)*
Distance (log)	-0.668***	-0.597***	-0.578***	-0.513***
	(0.0582)	*(0.0799)*	*(0.0551)*	*(0.0761)*
(abs.) Skill endowment diff.	-0.368***	-0.325**	-0.423***	-0.357***
	(0.130)	*(0.130)*	*(0.128)*	*(0.127)*
Sum of GDPs (log)	0.0857	-0.215	-0.0177	-0.300**
	(0.120)	*(0.145)*	*(0.120)*	*(0.141)*
Surr. market potential (log)	0.889*	1.850***	0.783*	1.663***
	(0.530)	*(0.549)*	*(0.465)*	*(0.493)*
Common coloniser		-0.117		-0.183
		(0.302)		*(0.284)*
Colonial relationship		0.291		0.0570
		(0.330)		*(0.322)*
Common off. language		0.756***		0.713***
		(0.166)		*(0.153)*
Contiguity		0.0569		0.0481
		(0.194)		*(0.191)*
XR of host country		0.000116		0.000106
		(8.31e-05)		*(8.50e-05)*
XR of source country		-5.30e-05		-4.49e-05
		(8.02e-05)		*(7.67e-05)*
XR regime (de jure)		0.0180		0.0190
		(0.0219)		*(0.0212)*
Dist. in financial development		-0.175		-0.178
		(0.311)		*(0.309)*
Dist. Economic Freedom Index		-0.115		-0.116
		(0.0825)		*(0.0816)*
Host country tariff (log)		0.0256		0.0223
		(0.0369)		*(0.0366)*
Constant	-25.63**	-22.49**	-27.06***	-22.98***
	(10.11)	*(9.218)*	*(9.906)*	*(8.780)*
N	36,121	30,912	239,320	214,186
Pseudo R2	0.842	0.863	0.597	0.614

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Dependent variable is operating revenue turnover. Panel (a) regressions are estimated using eq. (2), and Panel (b) regressions are estimated using eq. (3). Columns (1) and (3) present the pure model only with the KT (2010) variables and SMP. Columns (2) and (4) present regressions that include all control variables. Panel (a) regressions include additive year, host country and source country FEs. Panel (b) regressions additionally include additive sector FEs. All regressions include additive FEs for the country pairs that have been identified as outliers (see Appendix A).

N decreases owing to missing values in some control variables. Additional observations are automatically dropped from the PPML estimation if singletons.

Table B5 / Aggregate and pooled sectoral gravity regressions from Table 2 controlling for multilateral resistance

DV: Total assets	Panel (a): Aggregate model		Panel (b): Pooled sectoral model	
	(1) Aggregated at country-pair level	(2) Aggregated at country-pair level With controls	(3) NAICS 2-digit level	(4) NAICS 2-digit level With controls
Distance (log)	-0.658***	-0.676***	-0.564***	-0.595***
	(0.102)	*(0.122)*	*(0.0980)*	*(0.118)*
(abs.) Skill endowment diff.	-0.676***	-0.679***	-0.790***	-0.767***
	(0.206)	*(0.221)*	*(0.209)*	*(0.231)*
Sum of GDPs (log)	-0.283	-0.381*	-0.327	-0.415*
	(0.196)	*(0.226)*	*(0.207)*	*(0.237)*
Surr. market potential (log)	1.100*	2.456***	0.982*	2.274***
	(0.602)	*(0.701)*	*(0.514)*	*(0.621)*
Common coloniser		0.271		0.207
		(0.393)		*(0.384)*
Colonial relationship		0.496		0.130
		(0.314)		*(0.322)*
Common off. language		0.197		0.219
		(0.220)		*(0.211)*
Contiguity		0.450*		0.444*
		(0.268)		*(0.266)*
XR regime (de jure)		0.0357		0.0455
		(0.0309)		*(0.0303)*
Constant	11.61*	-0.401	11.14*	-0.411
	(6.546)	*(7.494)*	*(6.112)*	*(6.993)*
N	35,361	30,258	237,964	212,970
Pseudo R2	0.863	0.887	0.757	0.775
Source-year FEs	YES	YES	YES	YES
Host-year FEs	YES	YES	YES	YES
Sector FEs	NO	NO	YES	YES
Outlier FEs	YES	YES	YES	YES

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Dependent variable total assets. Panel (a) regressions are estimated using eq. (2), and Panel (b) regressions are estimated using eq. (3). Columns (1) and (3) present the pure model only with the KT (2010) variables and SMP. Columns (2) and (4) present regressions that include all control variables. Panel (a) regressions include additive source country-year, host-country year and outlier FEs. Panel (b) regressions additionally include additive sector FEs. N decreases owing to missing values in some explanatory and control variables. Additional observations are automatically dropped from the PPML estimation if singletons. The inclusion of the country-time FEs omits the estimates of all variables that are country-specific and time-varying, such as GDP (log), as well as bilateral variables that are likely to display little variation over time, such as the distance in economic freedom.

Table B6 / Pooled sectoral gravity model from Table 2, column (4), controlling for reverse causality

DV: Total assets	NAICS 2-digit level
Source GDP (log)$_{t-1}$	1.513**
	(0.684)
Host GDP (log)$_{t-1}$	1.982***
	(0.335)
Distance (log)	-0.581***
	(0.115)
(abs.) Skill endowment diff.$_{t-1}$	-0.804***
	(0.233)
Sum of GDPs (log)$_{t-1}$	-0.406*
	(0.242)
SMP (log)$_{t-1}$	2.132***
	(0.542)
Common coloniser	0.173
	(0.375)
Colonial relationship	0.191
	(0.324)
Common off. language	0.189
	(0.216)
Contiguity	0.452*
	(0.259)
XR of host country	9.88e-05*
	(5.61e-05)
XR of source country	1.07e-05
	(3.90e-05)
XR regime (de jure)	0.0490*
	(0.0298)
Dist. in financial development	-0.796
	(0.609)
Dist. Economic Freedom Index	0.131
	(0.0939)
Host country tariff (log)	-0.00861
	(0.0446)
Constant	-50.36***
	(12.90)
N	192,944
Pseudo R2	0.766

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Regressions are estimated using eq. (3), but the time-varying main gravity variables are included with a one-period lag. Regression includes additive year, country, sector and outlier FEs. N decreases, as, owing to the lag structure, the year 2020 is not included in the regression.

Post-estimation tests

Table B7 / Ramsey regression specification error test (RESET) for the aggregate gravity regression in Table 2, Panel (a), Column 2

(1) fit2 = 0
chi2(1) = 0.11
Prob > chi2 = 0.7452

Table B8 / F-test on joint insignificance of the main gravity variables from the aggregate gravity regression in Table 2, Panel (a), Column 2

(1) lngdp_o = 0
(2) lngdp_d = 0
(3) lndistw = 0
(4) lnsumgdp = 0
(5) abs_sk_diff = 0
(6) lnsmp_dest = 0
chi2(6) = 92.82
Prob > chi2 = 0.0000

Table B9 / RESET for the pooled sectoral gravity regression in Table 2, Panel (b), Column 4

(1) fit2 = 0
chi2(1) = 1.51
Prob > chi2 = 0.2193

Table B10 / F-test on joint insignificance of the main gravity variables from the pooled sectoral gravity regression in Table 2, Panel (b), Column 4

(1) lngdp_o = 0
(2) lngdp_d = 0
(3) lndistw = 0
(4) lnsumgdp = 0
(5) lnsmp_dest = 0
chi2(5) = 50.24
Prob > chi2 = 0.0000

Table B11 / Sector-specific gravity regressions

DV: Total assets	(1) 11: Agriculture, forestry, fishing and hunting	(2) 21: Mining, oil and gas extraction	(3) 22: Utilities	(4) 23: Construction of buildings	(5) 31: Food and textile manufacturing
Source GDP (log)	-0.201	-0.928	0.713	0.855	1.479
	(0.657)	*(1.024)*	*(0.960)*	*(0.894)*	*(1.773)*
Host GDP (log)	-0.172	-1.772*	1.003	-0.578	1.449
	(0.798)	*(1.047)*	*(0.769)*	*(0.483)*	*(1.220)*
Distance (log)	-0.617***	-0.689***	-1.085***	-0.516***	-0.792***
	(0.193)	*(0.189)*	*(0.231)*	*(0.156)*	*(0.156)*
(abs.) Skill endowment diff.	-0.403	-0.974**	-1.489***	0.796**	-1.103***
	(0.387)	*(0.473)*	*(0.363)*	*(0.404)*	*(0.409)*
Sum of GDPs (log)	0.570	-0.272	0.423	-0.789**	0.113
	(0.353)	*(0.806)*	*(0.365)*	*(0.339)*	*(0.275)*
Surr. market potential (log)	3.441*	-2.786***	-0.534	-1.793	1.536**
	(1.776)	*(0.763)*	*(1.263)*	*(1.588)*	*(0.646)*
Constant	-26.46	84.75***	-8.635	36.11**	-44.74
	(17.11)	*(22.75)*	*(19.98)*	*(17.12)*	*(35.04)*
N	4,710	5,355	6,605	10,354	9,641
Year FE	YES	YES	YES	YES	YES
Source FEs	YES	YES	YES	YES	YES
Host FEs	YES	YES	YES	YES	YES
Pseudo R2	0.626	0.840	0.739	0.714	0.666
DV: Total assets	**(6) 32: Materials manufacturing**	**(7) 33: Finished product manufacturing**	**(8) 42: Wholesale trade**	**(9) 44: Food and beverage stores**	**(10) 45: Miscellaneous store retailers**
Source GDP (log)	1.874	0.209	0.590	-2.370**	1.817
	(1.214)	*(0.638)*	*(0.497)*	*(1.182)*	*(2.593)*
Host GDP (log)	1.478***	0.831	0.976**	0.435	0.256
	(0.536)	*(0.717)*	*(0.466)*	*(0.781)*	*(1.051)*
Distance (log)	-0.382***	-0.857***	-0.694***	-0.331	0.185
	(0.115)	*(0.132)*	*(0.104)*	*(0.261)*	*(0.320)*
(abs.) Skill endowment diff.	0.0377	0.157	-0.161	-0.971*	-0.969
	(0.198)	*(0.222)*	*(0.165)*	*(0.513)*	*(0.714)*
Sum of GDPs (log)	0.151	0.0611	-0.000154	0.105	-1.170**
	(0.188)	*(0.156)*	*(0.132)*	*(0.295)*	*(0.579)*
Surr. market potential (log)	3.067**	4.724***	6.847***	2.597**	-4.224***
	(1.344)	*(1.347)*	*(1.354)*	*(1.322)*	*(1.237)*
Constant	-70.56***	-46.45***	-77.15***	10.92	32.08
	(23.91)	*(16.53)*	*(15.25)*	*(23.50)*	*(36.40)*
N	13,132	13,673	19,140	8,293	5,488
Year FEs	YES	YES	YES	YES	YES
Source FEs	YES	YES	YES	YES	YES
Host FEs	YES	YES	YES	YES	YES
Pseudo R2	0.763	0.799	0.839	0.592	0.698

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Regressions are estimated using eq. (3) separately for each sector, without sector FEs. The regressions include year, source and host country FEs. For space reasons, the table only reports coefficients of the main gravity variables. Outlier FEs are omitted as for each sector sample, there are different outlier country pairs, which would render comparability of the results difficult.

Table B11 / Continued

DV: Total assets	(11) 48: Air transportation	(12) 49: Warehouse and storage	(13) 51: Information services	(14) 52: Finance and insurance	(15) 53: Real estate, renting and leasing
Source GDP (log)	0.755	1.342	2.355	0.360	1.187
	(0.628)	*(1.497)*	*(1.461)*	*(1.032)*	*(0.899)*
Host GDP (log)	1.344*	0.769	0.895	1.974***	1.786**
	(0.762)	*(0.707)*	*(0.553)*	*(0.558)*	*(0.839)*
Distance (log)	-1.012***	0.575***	-0.844***	-0.623***	-0.795***
	(0.147)	*(0.220)*	*(0.252)*	*(0.203)*	*(0.173)*
(abs.) Skill endowment diff.	0.0144	-0.785*	-0.567	-0.499	-0.452
	(0.336)	*(0.472)*	*(0.444)*	*(0.432)*	*(0.307)*
Sum of GDPs (log)	-0.330	-0.0842	-0.916**	-1.319***	-0.251
	(0.344)	*(0.321)*	*(0.465)*	*(0.407)*	*(0.214)*
Surr. market potential (log)	6.113***	0.790	2.913	1.916***	4.919***
	(1.639)	*(2.335)*	*(1.976)*	*(0.713)*	*(1.773)*
Constant	-70.79***	-36.87	-48.12	-14.46	-72.87***
	(21.81)	*(32.96)*	*(33.28)*	*(19.43)*	*(27.46)*
N	10,226	4,022	8,933	13,094	11,672
Year FEs	YES	YES	YES	YES	YES
Source FEs	YES	YES	YES	YES	YES
Host FEs	YES	YES	YES	YES	YES
Pseudo R2	0.659	0.715	0.725	0.749	0.682
DV: Total assets	**(16) 54: Professional, scientific and technical services**	**(17) 55: Management of companies and enterprises**	**(18) 56: Administrative, support, waste management services**	**(19) 61: Educational services**	**(20) 62: Health care and social assistance**
Source GDP (log)	1.333*	2.432***	-0.588	-0.807	0.684
	(0.707)	*(0.832)*	*(1.004)*	*(1.486)*	*(0.835)*
Host GDP (log)	2.833***	4.148***	0.322	0.110	4.145***
	(1.093)	*(0.707)*	*(0.331)*	*(1.366)*	*(1.038)*
Distance (log)	-0.785***	-0.530***	-0.708***	0.0614	-0.219
	(0.120)	*(0.175)*	*(0.199)*	*(0.271)*	*(0.301)*
(abs.) Skill endowment diff.	-0.407*	-1.306***	0.352	-0.237	-0.680
	(0.211)	*(0.368)*	*(0.271)*	*(0.543)*	*(0.527)*
Sum of GDPs (log)	0.437**	-0.653**	0.399	-0.305	-0.0345
	(0.213)	*(0.323)*	*(0.273)*	*(0.456)*	*(0.655)*
Surr. market potential (log)	2.713**	2.879***	5.327***	1.450	7.053*
	(1.054)	*(0.594)*	*(1.537)*	*(2.885)*	*(3.899)*
Constant	-77.80***	-97.10***	-40.06*	5.066	-128.5***
	(23.71)	*(17.14)*	*(21.50)*	*(44.94)*	*(46.64)*
N	17,302	9,954	11,772	2,987	4,156
Year FEs	YES	YES	YES	YES	YES
Source FEs	YES	YES	YES	YES	YES
Host FEs	YES	YES	YES	YES	YES
Pseudo R2	0.862	0.898	0.906	0.798	0.704

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Regressions are estimated using eq. (3) separately for each sector, without sector FEs. The regressions include year, source and host country FEs. For space reasons, the table only reports coefficients of the main gravity variables. Outlier FEs are omitted as for each sector sample, there are different outlier country pairs, which would render comparability of the results difficult.

Table B11 / Continued

DV: Total assets	(21) 71: Arts, entertainment and recreation	(22) 72: Accommodation and food services	(23) 81: Other services
Source GDP (log)	1.849	-0.0694	0.289
	(2.152)	*(0.465)*	*(1.639)*
Host GDP (log)	3.184***	0.982*	1.008
	(0.853)	*(0.570)*	*(1.962)*
Distance (log)	-1.269***	-0.311**	-1.808***
	(0.362)	*(0.144)*	*(0.308)*
(abs.) Skill endowment diff.	0.0634	-1.195***	0.0535
	(0.825)	*(0.338)*	*(0.471)*
Sum of GDPs (log)	-1.359**	-0.838***	-0.716*
	(0.558)	*(0.169)*	*(0.366)*
Surr. market potential (log)	2.743	-2.857**	15.68***
	(1.837)	*(1.252)*	*(4.183)*
Constant	-62.77*	36.87***	-138.8***
	(36.15)	*(13.67)*	*(49.53)*
N	3,666	6,517	6,452
Year FEs	YES	YES	YES
Source FEs	YES	YES	YES
Host FEs	YES	YES	YES
Pseudo R2	0.781	0.827	0.827

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Regressions are estimated using eq. (3) separately for each sector, without sector FEs. The regressions include year, source and host country FEs. For space reasons, the table only reports coefficients of the main gravity variables. Outlier FEs are omitted as for each sector sample, there are different outlier country pairs, which would render comparability of the results difficult. Sector 92, ‘public administration’, which provides too few observations for precise estimates, is excluded from the regressions.

Table B12 / Distance interacted with six sector groups

VARIABLES	(1) Total assets
Source GDP (log)	1.418** *(0.642)*
Host GDP (log)	1.811*** *(0.354)*
Distance (log)	-0.533*** *(0.126)*
Distance (log) x Sector group 1	-0.114 *(0.105)*
Distance (log) x Sector group 2	0.00644 *(0.123)*
Distance (log) x Sector group 3	-0.116 *(0.163)*
Distance (log) x Sector group 4	-0.288* *(0.164)*
Distance (log) x Sector group 5	0.0407 *(0.144)*
Distance (log) x Sector group 6	-0.0501 *(0.113)*
(abs.) Skill endowment diff.	-0.786*** *(0.235)*
Sum of GDPs (log)	-0.418* *(0.241)*
Surr. market potential (log)	2.008*** *(0.530)*
Common coloniser	0.257 *(0.391)*
Colonial relationship	0.149 *(0.324)*
Common off. language	0.205 *(0.219)*
Contiguity	0.429 *(0.267)*
XR of host country	9.55e-05 *(5.87e-05)*
XR of source country	5.53e-06 *(3.76e-05)*
XR regime (de jure)	0.0465 *(0.0305)*
Dist. in financial development	-0.851 *(0.735)*
Dist. Economic Freedom Index	0.118 *(0.104)*
Host country tariff (log)	-0.0409 *(0.0491)*
Constant	-45.19*** *(13.18)*
N	214,155
Pseudo R2	0.765

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Dependent variable are total assets. The regression is estimated using eq. (6). Only the distance variable is interacted with six dummy variables for six sector groups, as follows. Group 1 for manufacturing sectors 31-33. Group 2 for the warehouse and storage sector (49) and accommodation and food services (72). Group 3 for mining, gas and oil extraction (21) and utilities (22). Group 4 for the arts and entertainment sector (71). Group 5 for the business service sectors 54 and 56. Group 6 for in-person services sectors 48 and 81, IT services (51), and financial services (sector 52). The coefficients of the main gravity variables reflect the average effect for all other sectors. The sector

classification can be found in Table A2. The regression includes additive year, host, source country, sector and outlier pair FEs.

Table B13 / F-test on joint insignificance of the interaction terms between distance (log) and each sector group dummy from the regression in Table B12

(1) lndistw_1 = 0
(2) lndistw_2 = 0
(3) lndistw_3 = 0
(4) lndistw_4 = 0
(5) lndistw_5 = 0
(6) lndistw_6 = 0
chi2(6) = 5.75
Prob > chi2 = 0.4519

Table B14 / Distance interacted with 22 sector dummies

VARIABLES	(1) Total assets
Source GDP (log)	1.418**
	(0.642)
Host GDP (log)	1.811***
	(0.354)
Distance (log)	-0.393***
	(0.151)
Distance (log) x Sector 11	-0.249
	(0.177)
Distance (log) x Sector 21	-0.196
	(0.294)
Distance (log) x Sector 22	-0.357**
	(0.173)
Distance (log) x Sector 23	-0.415**
	(0.172)
Distance (log) x Sector 31	-0.331**
	(0.153)
Distance (log) x Sector 32	-0.232*
	(0.135)
Distance (log) x Sector 33	-0.248*
	(0.127)
Distance (log) x Sector 42	-0.106
	(0.117)
Distance (log) x Sector 44	-0.376**
	(0.169)
Distance (log) x Sector 45	-0.116
	(0.240)
Distance (log) x Sector 48	-0.346***
	(0.119)
Distance (log) x Sector 49	-0.186
	(0.144)
Distance (log) x Sector 51	-0.278**
	(0.114)
Distance (log) x Sector 52	-0.189
	(0.125)

Table B14 / Continued

VARIABLES	(1) Total assets
Distance (log) x Sector 53	-0.430***
	(0.134)
Distance (log) x Sector 54	0.0700
	(0.204)
Distance (log) x Sector 55	-0.118
	(0.136)
Distance (log) x Sector 56	-0.249
	(0.156)
Distance (log) x Sector 61	-0.0323
	(0.104)
Distance (log) x Sector 62	0.101
	(0.165)
Distance (log) x Sector 71	-0.431***
	(0.162)
Distance (log) x Sector 72	-0.118
	(0.141)
(abs.) Skill endowment diff.	-0.788***
	(0.235)
Sum of GDPs (log)	-0.417*
	(0.242)
Surr. market potential (log)	2.006***
	(0.530)
Common coloniser	0.254
	(0.391)
Colonial relationship	0.150
	(0.325)
Common off. language	0.203
	(0.220)
Contiguity	0.425
	(0.267)
XR of host country	9.54e-05
	(5.87e-05)
XR of source country	5.38e-06
	(3.76e-05)
XR regime (de jure)	0.0463
	(0.0304)
Dist. in financial development	-0.850
	(0.735)
Dist. Economic Freedom Index	0.119
	(0.104)
Host country tariff (log)	-0.0410
	(0.0491)
Constant	-45.20***
	(13.18)
N	214,155
Pseudo R2	0.766

Standard errors clustered at country-pair level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Dependent variable are total assets. The regression is estimated using a version of eq. (6). The distance variable is interacted with 22 dummy variables for each individual sector (using sector 81, ‘other services’, as the base category). The coefficients of the main gravity variables reflect the effect for the base category, sector 81. The sector

classification can be found in Table A2. The regression includes additive year, host, source country, sector and outlier pair FEs.

Table B15 / F-test on joint insignificance of the interaction terms between distance (log) and each sector dummy from the regression in Table B14

(1) lndistw_11 = 0	(11) lndistw_48 = 0	(21) lndistw_71 = 0
(2) lndistw_21 = 0	(12) lndistw_49 = 0	(22) lndistw_72 = 0
(3) lndistw_22 = 0	(13) lndistw_51 = 0	
(4) lndistw_23 = 0	(14) lndistw_52 = 0	
(5) lndistw_31 = 0	(15) lndistw_53 = 0	
(6) lndistw_32 = 0	(16) lndistw_54 = 0	
(7) lndistw_33 = 0	(17) lndistw_55 = 0	
(8) lndistw_42 = 0	(18) lndistw_56 = 0	
(9) lndistw_44 = 0	(19) lndistw_61 = 0	
(10) lndistw_45 = 0	(20) lndistw_62 = 0	
chi2(22) = 57.29		
Prob > chi2 = 0.0001		

Acknowledgements

This paper builds on Nicola Rebmann's thesis at the University of Groningen and its content does not necessarily represent the views of the European Central Bank and the Eurosystem.

Nicola Rebmann
European Central Bank, Frankfurt am Main, Germany; email: nicola.rebmann@web.de

Konstantin M. Wacker
University of Groningen, Groningen, The Netherlands; Vienna Institute for International Economic Studies (wiiw), Vienna, Austria; email: k.m.wacker@rug.nl

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-7897-2 ISSN 1725-2806 doi:10.2866/7988404 QB-01-26-146-EN-N