DISCUSSION PAPER SERIES

DP13491

GRAVITY, COUNTERPARTIES, AND FOREIGN INVESTMENT

Tarun Ramadorai, and Chihiro Shimizu

FINANCIAL ECONOMICS, INTERNATIONAL MACROECONOMICS AND FINANCE AND INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

GRAVITY, COUNTERPARTIES, AND FOREIGN INVESTMENT

Tarun Ramadorai, and Chihiro Shimizu

Discussion Paper DP13491
Published 26 January 2019
Submitted 24 January 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

GRAVITY, COUNTERPARTIES, AND FOREIGN INVESTMENT

Abstract

Gravity models excel at explaining international trade and investment flows; their success poses a continuing puzzle. In a comprehensive dataset of global commercial real-estate investments, we find that the role of distance in the gravity model is well-explained by preferential matching between counterparties of the same nationality. This tendency for same-country matching is widespread, robust, and increases in poorly-governed locations. We structurally estimate an equilibrium matching model with a friction affecting different-nationality transactions. The model explains the persistent success of gravity using a combination of this friction and the spatial distribution of same-nationality counterparties, which is well-predicted by current and historical linguistic, cultural, and trade links between countries.

JEL Classification: D83, F14, F30, G11

Keywords: Gravity, Foreign investment, Commercial real estate, Trust, Matching, Cross-border flows

Tarun Ramadorai - t.ramadorai@imperial.ac.uk
Imperial College London and CEPR

 - cristian.badarinza@nus.edu.sg
National University of Singapore

Chihiro Shimizu - shimizu.chihiro@nihon-u.ac.jp
Nihon University and MIT

Gravity, Counterparties, and Foreign Investment*

Cristian Badarinza, Tarun Ramadorai, and Chihiro Shimizu†

January 24, 2019

Abstract

Gravity models excel at explaining international trade and investment flows; their success poses a continuing puzzle. In a comprehensive dataset of global commercial real-estate investments, we find that the role of distance in the gravity model is well-explained by preferential matching between counterparties of the same nationality. This tendency for same-country matching is widespread, robust, and increases in poorly-governed locations. We structurally estimate an equilibrium matching model with a friction affecting different-nationality transactions. The model explains the persistent success of gravity using a combination of this friction and the spatial distribution of same-nationality counterparties, which is well-predicted by current and historical linguistic, cultural, and trade links between countries.

*We thank seminar participants at the National University of Singapore, Warwick Business School, London Business School, Imperial College Business School, Tinbergen Institute, University of Southern California, UCLA, the Atlanta ASSA Meeting and Patrick Bolton, John Campbell, Darrell Duffie, Andra Ghent, Chris Hansman, Ralph Koijen, Elias Papaioannou, Steven Poelhekke, Helene Rey, Eva Steiner, Paolo Surico, Raman Uppal, and Ansgar Walther for comments and useful conversations. We gratefully acknowledge Hoang Minh Duy for excellent research assistance.

†Badarinza: Department of Real Estate, Institute of Real Estate Studies, 4 Architecture Drive, Singapore 117566, National University of Singapore, and CEPR. Email cristian.badarinza@nus.edu.sg. Ramadorai (corresponding author): Imperial College London, Tanaka Building, South Kensington Campus, London SW7 2AZ, and CEPR. Tel.: +44 207 594 99 10. Email: t.ramadorai@imperial.ac.uk. Shimizu: University of Tokyo, and Center for Real Estate, MIT, Cambridge, Massachusetts. Email: shimizu.chihiro@nihon-u.ac.jp.

1 Introduction

Gravity models have been very successful at explaining international trade and investment flows,[1] though the underlying reasons for their success pose a continuing puzzle. Empirical gravity equations reveal, somewhat mysteriously, that trade and foreign investment flows decline substantially as the physical distance between origin and destination countries increases. A promising line of research highlights the important role of informational and contracting frictions between counterparties, with consequences for the formation of cross-border networks and the spatial distribution of cross-border flows.[2]

In this paper, we bring new evidence to bear on these questions, and set up and estimate a structural model to rationalize the puzzling role of distance in estimated gravity equations. The evidence comes from comprehensive data covering all high-value transactions in over 70 countries in the global commercial real estate market. This is an important venue for cross-border investment, with a global transaction volume of US$ 660BN in 2016. A unique feature of these data is that they identify both counterparties in all transactions, as well as the nation in which these counterparties are incorporated. This information allows us to uncover how counterparty matching frictions give rise to observed gravity relationships in the data.

We find that buyers of commercial real estate have an unusually strong tendency to transact with sellers who hail from their country of origin.[3] We term this pronounced preference to transact with counterparties from the same country *nationality bias*. As we describe in the paper, we identify this preference using the substantial over-representation, relative to their prevalence in any given location, of sellers from a given country in transactions with buyers from the *same* country. This tendency shows up for virtually all nationalities, is present

[1]Anderson (2011) and Head and Mayer (2014) survey the literature on gravity models, and see, for example, Portes and Rey (2005), who show that gravity models can help to explain the behaviour of cross-border capital flows.

[2]See, for example, Rauch (1999), Rauch and Trindade (2002), and Chaney (2014).

[3]We use the domicile status of firms interchangeably with the term "nationality" in what follows.

when transactions occur at home or overseas, and is economically large and statistically robust.[4]

Nationality bias is particularly strong when buyers venture overseas—buyers' propensity to match with same-nationality sellers is 44% greater, on average, than the occurrence of same-nationality sellers in foreign countries. The prices of property transactions occurring between same-country counterparties are also higher on average (by 7.36%) controlling for a range of hedonic characteristics, time, and region effects. This preferential matching tendency is restricted to same-nationality matches, with no greater tendency for matching between counterparties hailing from countries with cultural or linguistic links, or those that are physically proximate. We do find, however, that same-country matching rates rise substantially in locations in which the rule of law is weak, suggesting that contracting frictions and trust are the principal driver of this tendency (see e.g., Nunn (2007)).

How does nationality bias affect estimated gravity equations? We begin by estimating a standard naïve gravity equation to explain the log volume of cross-border commercial property investment flowing from (buyer) origin countries towards investment destination countries. As with most estimated gravity equations in the literature, we find that the effect of log physical distance between origin and destination locations on bilateral flows is strong and negative.[5] We then add to this equation the transactions volume in the location country generated by sellers of the *same* nationality as the buyer, as a simple proxy for the location-specific density of same nationality counterparties. The inclusion of this new variable renders the estimated coefficient on log distance in the gravity equation indistinguishable from zero. This attenuation in the role of distance is robust to the inclusion of a range of controls and

[4]As an illustrative example, consider a location in which transactions occur, and assume that a particular country's (say Indian) sellers account for one-tenth of transactions in this location, regardless of buyer nationality. This fraction can be compared to the representation of Indian sellers in transactions that also have Indian buyers; we find that the latter fraction is far greater than a tenth. This could, of course, arise because of preferential matching of particular nationalities to particular property characteristics or locations. However, as we describe, a range of additional parametric and non-parametric tests provide little evidence to support this conjecture.

[5]A useful point to note here is that the nature of the asset being traded—commercial real estate—immediately rules out transportation-cost based explanations for gravity in this context. Nevertheless, the data reveal a strong role for physical distance. This result is similar to Blum and Goldfarb (2006), who find gravity relationships for digital goods consumed over the internet that have no trading costs.

fixed effects, to using a number of alternative estimation techniques currently used in the gravity literature,[6] and is not mechanical, as we confirm using placebo simulations.

To explain this striking result, we more closely study the spatial distribution of same-nationality counterparties. Intriguingly, we find that the density of same-nationality sellers, on average, declines log-linearly in distance from buyer countries, i.e., in a "gravity-like" manner. Digging deeper, we find that the spatial distribution of same-nationality sellers is well-explained by historical shipping and commercial trading routes, colonial and linguistic ties, as well as patterns of trade flows and the existence of free trade agreements between countries. The combination of this spatial distribution of counterparties and nationality bias is responsible for the empirical gravity relationships that we observe in the data.

To better understand the underlying economics driving these new facts, we set up an equilibrium model with heterogeneous buyers and sellers, random matching, and endogenous determination of volumes and prices in a rational expectations equilibrium. The main assumption in the model is that transactions with different-nationality counterparties are subject to a friction which affects their expected value.[7] Buyers thus prefer same-nationality matches to cross-nationality matches, which determines the likelihood that meetings between buyers and sellers will result in a successful transaction. Sellers also experience valuation uncertainty, which may lead them to post a lower price in an attempt to avoid losses arising from failed matches. In equilibrium, buyers and sellers act optimally given the frictions in the model, and form rational expectations about their counterparties' decisions when accepting, rejecting, or posting offers.

We solve the model in closed form, and structurally estimate that the friction is equivalent to an expected value reduction of 9.4%, meaning that buyers in the model are willing to pay this amount to avoid transacting with different-nationality counterparties. In the counterfactual frictionless economy, volumes increase by 6.5% as a result of new transactions between different nationalities. There is also a predicted increase in the average price

[6] Head and Mayer (2014) provide a useful survey of current challenges and the state of the art in the estimation of gravity models in international trade and investment flows.

[7] We interpret this friction as a generic representation of difficulties in contracting, or a lack of trust that affects transactions with different nationality counterparties.

per transaction of 7.4%, which is substantial, given that the average transaction size in the data is US$ 10MM.

How does the model explain the puzzling role of distance in estimated gravity equations? We assume that buyers rationally anticipate the frictions that they will encounter in equilibrium, and pre-filter countries based on their expected likelihood of encountering same-nationality counterparties. This pre-filtering process leads them to scale their desired capital flows towards and away from particular markets depending on the size of the friction and the location-specific densities of same-nationality sellers. These location-specific densities are exogenous to our model, but as we show in the data, the establishment of physical "beachheads" with high densities of same-nationality counterparties is well-predicted by historical and current links and patterns of trade between pairs of countries. The combination of this spatial distribution of beachheads and nationality bias delivers the observed role of distance between origin and destination countries both in the model and in estimated gravity equations. Put differently, given an initial/historical stock of bilateral investment that declines with physical distance between origin and destination countries, nationality bias is a strong force which perpetuates historically observed gravity effects.

In addition to the large literature on gravity models,[8] our work is related to the growing literature on the role of networks, affinity, and trust in international trade and finance.[9] It is also related to the literature on home bias at home and abroad.[10] Our use of commercial real estate market data connects the paper to the growing literature on information asymmetries and social networks[11] in real estate markets. Our theoretical model builds on frameworks developed by Han and Strange (2015), Landvoigt et al. (2015), and Piazzesi et al. (2017) on segmented housing search, and extends this literature in two ways, introducing a new

[8] Other important papers in this literature include Anderson and Van Wincoop (2003), and Antràs (2003).

[9] See, for example, Combes et al. (2005), Guiso et al. (2009), Garmendia et al. (2012), Burchardi and Hassan (2013), and Burchardi et al. (2017).

[10] See, for example, French and Poterba (1991), Tesar et al. (1995), Coval and Moskowitz (1999), Huberman (2001), Ahearne et al. (2004), Nieuwerburgh and Veldkamp (2009), and Coeurdacier and Rey (2013). Branikas et al. (2017) show that the phenomenon of home bias in the allocation of households' investment portfolios is significantly reduced when accounting for the households' endogenous residential location decision.

[11] See, for example, Garmaise and Moskowitz (2004), Levitt and Syverson (2008), Chinco and Mayer (2016), Kurlat and Stroebel (2015), and.Bailey et al. (2018).

matching friction to capture nationality bias in the model, and explicitly modelling the distribution of buyer valuations rather than assuming random arrival rates of inventory on the market. Finally, our work contributes to a new and growing literature on capital flows in global real estate markets. For example, Badarinza and Ramadorai (2018) document the impact of foreign buyers on the London real estate market using a new cross-sectional identification approach based on different nationalities' preferred locations with the city, and Van Nieuwerburgh and Favilukis (2017) propose a welfare-cost approach to understanding the market impact of foreign investors in the market for residential real estate.[12]

The paper is organized as follows. Section 2 describes the dataset that we employ in our empirical work, and Section 3 outlines the empirical methodology to identify nationality bias, reports estimates of this bias, and investigates its drivers. Section 4 estimates gravity equations, and connects nationality bias with gravity. Section 5 introduces the equilibrium matching model, and Section 6 describes how we structurally estimate the model and use it to evaluate counterfactuals. Section 7 concludes.

[12]See also Sa (2015), Cvijanovic and Spaenjers (2015), Miyakawa et al. (2016), and Agarwal et al. (2017).

2 Data

2.1 Commercial Real Estate Transactions

Our main dataset contains transaction-level information which covers 87,679 individual deals in a total of 123,648 commercial properties. These properties are located in 434 metropolitan areas in 70 countries, and the transactions occur over the period from January 2007 to October 2017. Real Capital Analytics (RCA) provide these data, with the aim of capturing the universe of global commercial real estate deals with a value above USD$ 10 million. For each property, we know the exact location, total floor space area, the year of construction, the type of functional use (office, retail, business apartments, industrial facilities and hotels), and the transaction price.

In addition to information on properties, the dataset contains details about the buying and selling entities in these transactions, which comprise a total of 42,923 firms. For these buyer and seller entities, we know their registered name, their ownership/listing status (privately held, publicly listed, or held by an institution such as a sovereign wealth fund or a pension fund), their type (real estate developer, owner, operator, equity fund, Real Estate Investment Trusts, REITs etc.), and the buyer's stated objective for the property purchase (i.e., investment, occupancy, redevelopment, or renovation).

Importantly, for each transaction, RCA identify the country in which both buying and selling counterparties are incorporated. This information is what we term the "nationality" of buyers and sellers. When counterparties are multinational entities, the data record whether the property was bought by the holding company itself, or by a local branch of the holding company. We use the country of incorporation of the actual entity involved in the transaction—e.g., if the local branch was involved, we would use its nationality rather than the nationality of incorporation of the holding company.

Table 1 summarizes the main features of the data. Panel A shows that the average property transacted in the data was built in 1984. The average size of transacted properties is 186,631 ft^2, and the average price is US$ 39 million. Per square foot, properties transacted at an average price of US$ 294. Panel B of the table shows that 32.6% of the transactions are for office buildings, 23.4% for retail outlets, 21.1% for rental apartments, and the remaining transactions involve industrial facilities and hotels.

The data cover transactions in 434 metropolitan areas in 70 countries; the online appendix shows a world map with the locations of all transactions in the data. In our empirical work, as we explain below, we employ a narrower geographic classification of these metropolitan areas into sub-markets—these 925 sub-markets are defined by RCA, and generally correspond to districts or boroughs of each metropolitan area (e.g., neighbourhoods such as the West End in London, or the Upper East Side in New York). Roughly a fifth of the sample comprises properties in the Central Business District (CBD) of each city in the data, with the remainder located outside the CBD. Panel B also shows that a majority of the deals (53.7%) involve the sale of a single property, but 46.3% of the deals involve selling multiple properties simultaneously. Our results are robust to using either transaction- or deal-level data; we use deals in our primary analysis and refer to transactions and deals interchangeably in what follows.[13]

[13]To assign a deal to a specific sub-market within a city, whenever there is more than one property in the portfolio that is being traded, we use the location of the property with the highest value. All our results are robust to working with individual transactions rather than deals, both in terms of magnitude and statistical significance.

2.2 Buyers and Sellers, at Home and Abroad

Panel C of Table 1 shows that buyers and sellers are of a number of different corporate types, with a slight dominance of unlisted private companies (42.1% of buyers and 43.1% of sellers). A majority of these entities can be broadly classified as real estate developers, owners, or operators (37.0% of buyers and 40.2% of sellers), but there are also large fractions of investment funds, foundations and endowments (Other), and REITs.

The top panel of Figure 1 classifies all transactions in the data by the physical locations in which they occur. The figure shows that more than half of the transactions in the sample take place in properties that are located in the United States. Outside of the US, the top five markets are Japan, Germany, the United Kingdom, Sweden, and Australia. The bottom panel reclassifies all transactions, and ranks them by the most frequently represented buyer (left) and seller (right) nationalities. The top three buyer nationalities represented in the data are the US, Japan, and Canada, and the top three seller nationalities are the US, Japan, and the UK.

The lighter shaded portion of each bar in all of the panels indicates the fraction of transactions in that occur between counterparties with *different* nationalities, while darker shades indicate transactions between counterparties that share the *same* nationality. The top panel of the figure shows that the US is a highly local market, with most counterparties transacting there sharing the same nationality (i.e., US buyers matching with US sellers in the US). In contrast, properties located in most other countries have far larger shares of transactions involving counterparties of different nationalities—because of the greater prevalence of foreign investment in commercial real estate in these countries.

The bottom panel of the figure reveals that most buyer and seller countries appear to show a high share of transactions with counterparties hailing from their own country, though this fraction varies across countries. It is worth noting that the "Other" countries in which counterparties in the sample are domiciled undertake fewer than 7,000 transactions on either buy or sell sides. This means that offshore jurisdictions such as the Cayman Islands are barely represented in the data, which is reassuring, as such transactions would be difficult to trace back to the true national origin of investment flows.

As in these figures, in our main analysis, we distinguish counterparties hail from different countries (e.g., a French company purchases a property from a German company) and those involving same-nationality counterparties (e.g., French buyers transacting with French sellers). We further distinguish between transactions occurring "at home" (e.g., a Chinese buyer purchasing a property located in China from a Chinese seller) and " abroad" (e.g., a Chinese company purchasing a German property from another Chinese company).

2.3 Company Characteristics

To ensure that the transactions that we identify are arms-length, we collect information on the shareholder structure from Bureau van Dijk's Orbis database and hand-collect evidence, primarily from news media, of M&A activity between buyer and seller companies.[14] This process allows us to eliminate 4,082 transactions that happen within the same group, or for which there is a shareholder relationship between the buyer and the seller. The final number of transactions that we employ in our analysis (123,648) is net of this data cleaning process.

2.4 Trade and Distance Data

We find that a gravity equation is well able to explain the cross-border patterns of commercial real estate investment in the data. To better understand the source of these patterns as well as the source of the high rates of same-nationality matching in the data, we acquire data that has been useful in the empirical trade literature. We describe these datasets below, as well as some new datasets (such as historic shipping routes) that has hitherto been unutilized in the empirical trade literature.

[14]We restrict our hand-collecting exercise to transactions occurring between same-nationality counterparties, both to reduce the amount of manual effort involved, and to ensure that any biases resulting from this process of manual data collection work against finding our main result of high rates of same-nationality matching.

2.4.1 Physical and Cultural Distance

We use the GeoDist/CEPII database of Mayer and Zignago (2011) to identify country pairs with a common official language, common colonial history, and a common border, as well as to measure the physical distance between the countries in our sample.

In the baseline results that we report in the paper, we opt to measure the physical distance between the most populated cities of a country. Since such measurement is important for our analysis, we also test robustness to using alternative measures, including the physical distance between capital cities, or distances to countries' population-weighted centres.

2.4.2 Historical Trading Routes

To capture the likely historical determinants of preferred foreign investment destinations and current counterparty locations, we use a newly available database of historical shipping routes, obtained from the Climatological Database for the World's Oceans. For each buyer-location country pair in our sample, we calculate the number of trips that were carried out during the period 1750-1850 between ports located within the current geographical borders of these countries.

2.4.3 Trade Agreements and Trade Volumes

We use the Regional Agreements database of the World Trade Organization (WTO) to check for the existence of free trade agreements between the country of origin of buyers and the location country of the investment property.

Finally, we use the TradHist/CEPII database of Fouquin et al. (2016) to measure bilateral trade volumes over the period 1827-2014. For each country pair, we calculate the average trade volume over the period and use it to predict the location-specific density of same-nationality sellers, as we describe in detail in Section 4.

We now turn to a deeper understanding of the patterns in counterparty matching.

3 Counterparty Matching: Nationality Bias

In this section, we investigate the patterns detected in 1, and set up a simple benchmark distribution of counterparty matching rates based on an assumption of random matching. We then compare the observed distribution of the matching rates between counterparties of same and different nationalities to this simple benchmark to derive a new measure, which we term *nationality bias*. This new measure is very similar to previous measures proposed in the home bias literature (see, for example, Coval and Moskowitz (1999)).

We estimate the magnitude of nationality bias using all transactions in the data, and describe a range of robustness checks that we conduct to verify our results. We then investigate the drivers of nationality bias by correlating the observed empirical estimates with a set of country-specific measures.

Using these observed counterparty matching patterns as a guide, the subsequent section estimates a set of gravity equations that leads to a deeper understanding of cross-border investment flows.

3.1 Patterns of Buyer-Seller Matching

Figure 2 illustrates how we estimate nationality bias in three locations around the world, corresponding to Panels A, B, and C. Panel A of the figure focuses on the 636 transactions that take place over our sample period in properties physically located in the West End of London. The top bar in this panel shows that 72% of these properties are sold by UK-incorporated entities, 7% by US-incorporated sellers, and 11% by sellers from other countries. The bottom bar in this panel focuses on the 52 transactions in the West End in which *the buyer is incorporated in the US*. The bar shows that 21% of the sellers in these transactions are from the US. The difference between the conditional (21%) and unconditional (7%) shares of US sellers gives us the measure of nationality bias (i.e., preferential matching between same-nationality counterparties) for the US in the West End, namely, $21\% - 7\% = 14\%$.

Similarly, Panel B looks at the 82 transactions occurring in the Central Business District in Sydney over the sample period. 5% of these transactions involve Chinese sellers. The corresponding fraction of Chinese sellers in the set of transactions involving Chinese buyers is 22%, resulting in a bias estimate of 22% − 5% = 17%. And Panel C shows that the same phenomenon shows up in the Quartier Central des Affaires in Paris, where 4% of all the 367 transactions involve Spanish sellers, but Spanish sellers comprise a far larger 33% share of all transactions involving a Spanish buyer.

3.2 Measurement

Consider a specific location (such as the West End of London, or the Upper East Side of Manhattan) in which commercial property is physically located. In this location, let N_{ij} be the total number of property transactions in which the buyer is from country $i = 1, ..., I$ and the seller from country $j = 1, ..., J$.

The total number of transactions involving sellers from country j is then:

$$\sum_{i=1}^{I} N_{ij}. \tag{1}$$

We can represent this as a fraction of all transactions in the location, i.e.,

$$m_j = \frac{\sum_{i=1}^{I} N_{ij}}{\sum_{j=1}^{J} \sum_{i=1}^{I} N_{ij}}. \tag{2}$$

Equation (2) is simply the "unconditional" or "benchmark" fraction outlined in the simple example at the end of the previous section.

The fraction of all transactions involving sellers from country j and buyers from country i is:

$$h_{ij} = \frac{N_{ij}}{\sum_{j=1}^{J} N_{ij}}. \tag{3}$$

A simple null hypothesis here is that $E[h_{ij}] = m_j$, i.e., that there is no systematic preferential matching for any given (i, j) pair. This null is motivated by random matching—

in which buyers arrive in a location and randomly match to available sellers in the location.[15]

We note here that we carefully consider the possibility that common preferences for particular locations or property characteristics can drive observed matching patterns in the robustness section, alongside a range of other potential issues. For now, we proceed with this simple null.

A pair of special interest here is h_{ii}, i.e., transactions involving buyers and sellers from the same country, as in the examples considered above. This allows us to define the absolute measure of *nationality bias* for buyers from countries i:

$$Bias(N_{ii}) = h_{ii} - m_i. \tag{4}$$

We can then generalize this reasoning to any location in which transactions occur, and write down a testable null hypothesis, averaged across all buyer nationalities and locations of transactions:

$$H_0 : E[Bias(N_{ii})] = 0. \tag{5}$$

[15]Violations of the null suggest either non-random patterns of matching between particular types of counterparties, or non-random patterns of arrival of buyers and sellers. We subsequently evaluate both possibilities in our empirical analysis, as well as using our model.

3.3 Estimation

Panel A of Table 2 reports average estimates of nationality bias $E[Bias(N_{ii})]$ from this exercise, as well as separate results "at home," and "abroad."[16]

In the full sample of transactions, the equal-weighted average of m_i across all buyer nationalities and locations is 24.6%, and the equal-weighted average of h_{ii} across all locations in the data is 26.6%. Using these averages, our estimate of nationality bias is a statistically significant $26.6\% - 24.6\% = 2\%$,which is the "Average effect" for the "Absolute measure" in Panel A.

When we restrict our focus to transactions that occur "at home," (i.e., the buyer is incorporated in the same country as the location of the traded property), the average market share of sellers belonging to the home country is 78.3%. The average market share of sellers in home transactions by a buyer from the same nationality is 79.5%. This leads to a relatively modest (albeit statistically significant) estimate of 1.2% for the bias at home.

However, a substantially larger bias is evident when buyers transact in foreign countries. The equal weighted average m_i for foreign sellers is 5.23%, i.e., this is the unconditional fraction of foreign sellers present in any given location. Conditional on the buyer being from the same foreign country, h_{ii} is a much larger 7.51%. The difference of 2.32% between these two numbers is substantial, since it is almost 50% of the unconditional fraction of 5.23%. We discuss the economic importance of this finding in detail in subsequent sections.

We describe the economic forces behind this phenomenon in the next section, but first quickly describe a set of robustness tests that we conduct on these estimates.

[16] The standard errors are computed using a two-stage bootstrap procedure, designed to correct for clustering at the sub-market level. First, we run $n = 1,000$ iterations of random draws of bootstrap samples. In each iteration, we draw with replacement from the set of 925 sub-markets, including all transactions observed in a given sub-market if it is drawn. We then use equations (2) and (3) to compute the sets of conditional (h_{ii}) and unconditional (m_i) market shares, and then compute the bootstrapped bias measures.

3.4 Robustness

3.4.1 Relative Measure

For comparability with previous research on systematic biases in international investments, we consider a "Relative measure," which slightly modifies equation (5) by increasing the weights in the grand average for nationalities that account for a larger share of the seller pool in each location.[17] The right-hand side of Panel A of Table 2 shows that the estimated effects are strong and highly statistically significantly different from zero with this modification, and are similarly sized across estimation methods and transaction types.

3.4.2 Placebo Simulations

We check whether the null of random matching indeed delivers estimated nationality bias of zero, or whether rejections of the null can arise spuriously from the structure of the dataset. We conduct a placebo test that imposes the null hypothesis $E[Bias(N_{ii})] = 0$, by reconstructing the sample in each of $n = 1,000$ simulation rounds, randomly matching counterparties within locations in the data. We relegate the detailed description of this exercise to the online appendix. In short, we find that both at home and abroad, and using both weighted and unweighted measures, the point estimate of nationality bias lies well outside the resulting placebo distribution, strongly rejecting the possibility of spurious rejections of the null.

[17] In this case, we calculate the bias measure as equal to $\frac{h_{ii}-m_i}{1-m_i}$, which is essentially identical to the local bias measure of Coval and Moskowitz (1999), for the simple quantification of their distance measure as equal to zero when buyers trade with sellers domiciled in the same country, and equal to one otherwise.

3.4.3 Base Effects

Estimates of nationality bias can be affected by the fact that seller fractions are calculated using a common base for each nationality, and within each location. The decisions of investors from a given country i can therefore affect the set of available transactions for investors from all other countries.

As a result, nationality bias can be mistakenly attributed to multiple countries even if it is a phenomenon restricted to a few countries. We note that this issue can also affect estimates of gravity equations in cross-border capital flows, as well as standard estimates of home bias. We relegate the description of simulation experiments and associated figures that we use to check whether this is an issue to the online appendix, but note here that the results reinforce the robustness of our estimates, and suggest that such base effects play a negligible role.[18]

3.4.4 Do Nationalities Match to Characteristics?

An important question is whether seller market shares in the full set of transactions m_i are the correct counterfactual distribution of seller nationalities for buyers from country i. One objection to this benchmark is that deviations from it could reflect the unobserved preferences of seller nationalities.

To take an example, assortative matching could drive the observed result. Say that Chinese investors have a preference for properties in a particular location, or for properties with specific characteristics. If so, their purchasing decisions would cluster around specific areas or property types. Such clustering would naturally lead to more frequent transactions between Chinese counterparties, since they will have a higher ownership share in the locations that they prefer.

[18] We also note that any adverse effects of this issue on the variance of the estimator are mitigated by our clustering of the bootstrapped standard errors at the level of sub-markets.

However, this observation would not have anything to do with a preference for transacting with other Chinese investors.[19] In such a case, the density of sellers of the same nationality would be an informative measure of "distance" between buyers and specific property locations or characteristics. This is not merely a robustness check—it is of economic interest to check for this source, as it raises the possibility that counterparty density is a potentially useful source of information about "distance" when estimating gravity effects.

To check whether assortative matching to locations or characteristics is at work, we first adopt a parametric (logit) propensity-score approach, changing the calculation of the counterfactual seller shares m_i to account for the preference of specific nationalities for particular transaction- and property-level characteristics. We estimate a logit propensity score for each transaction to involve a buyer from country i, running regressions for each buyer nationality available in the data, conditioning on a set of characteristics such as the year during which the transaction took place, the type of property, and an indicator of price quintile—using the distribution of prices within each country in every given year.[20] For each location, we then apply the Logit propensity scores as weights, to compute a conditional version of m_i, which translates into a conditional bias measure. Despite the propensity score capturing heterogeneity in preferences across buyer countries, this change in the computation of $m_i^{matched}$ results in the bias estimates falling only slightly. Panel B of Table 2 shows that the estimated overall average nationality bias effect decreases from 1 percentage point to 0.8 percentage points, and the high level of statistical significance is preserved.

We also use a non-parametric K-means clustering approach to isolate clusters of N observations within each location. We cluster along alternative dimensions, by location alone, as well as by location, transaction, and property characteristics. Panel B of Table 2 shows that even if we zoom in enough to identify nationality bias effects within small clusters of 20 transactions, the average magnitude of nationality bias is barely affected by

[19] Badarinza and Ramadorai (2018) document significant within-city variation in geographical segmentation of people from different countries in the residential real estate market, suggesting that this may be an issue.

[20] In practice, we restrict this analysis to all nationalities with a sufficient number (25 in our empirical analysis) of transactions, and use the unweighted estimates for the nationalities with small numbers of transactions.

this clustering procedure; indeed it occasionally increases the estimate of the nationality bias abroad, where same-nationality counterparty matches often occur in properties which are atypical for particular nationalities.

In the online appendix, we show the correlation between the propensity score adjusted benchmark and the baseline fractions of same-nationality sellers, and describe the clustering procedure in detail.

3.4.5 Subsample Analysis

Finally, we note that nationality bias is strong and robust across a wide variety of subsample analyses, showing up for virtually all the countries in the sample, in all time periods in the data, in a wide range of location countries, and for virtually all corporate types of counterparties. These results are available in the online appendix.

3.5 The Drivers of Nationality Bias

To understand what drives nationality bias, we compute all bilateral bias measures $Bias(N_{ij,k})$ for buyers hailing from country i transacting with sellers from country j, averaged across all locations in the countries k represented in the data. We then explore the relationship between these measures and a range of controls.

The leftmost column of Table 3 provides a point of reference. It reports the estimated magnitude of nationality bias from a regression of $Bias(N_{ij,k})$ on a dummy variable that indicates when $i = j$. The magnitude of this coefficient differs slightly from that in Table 2, because it is estimated using a sub-sample of locations, corresponding to the set of countries for which our conditioning variables are available.

In the second column of the table, we explore the hypothesis that buyers have a more general preference to trade with sellers from countries that are proximate to them in a number of ways—measured using both physical and cultural distance metrics. The data robustly reject this hypothesis. First of all, the data show that the matching bias that we discover is strictly confined to same-nationality counterparties. Counterparties from countries that are physically proximate, share a common language, colonial history, common border, or a trade agreement with one another are no more likely to be selected than a randomly selected nationality. These results make it far less likely that the matching bias that we detect is related to issues of cultural affinity.

Is nationality bias linked to the attributes of the country in which properties are physically located? The rightmost columns of Table 3 explore this possibility further. To quantify the contractual environment prevailing in different location countries, we use three measures: a measure of limited rule of law (from the Worldwide Governance Indicators of the World Bank); a measure of real estate market opacity (from the Jones-Lang-Lasalle (JLL) Real Estate Transparency Index); and a measure of limited economic development (the Log of the inverse GDP level). The data reveal that nationality bias is most pronounced in countries with limited rule of law. There is no residual effect for those in low-GDP countries or with opaque real estate markets.

Taken together, this evidence suggests that the market friction driving the pronounced preference for same-nationality counterparty matching is inadequacies in the legal and contractual environment in the destination market. To surmount this friction, foreign investors appear to rely on pre-existing networks of business relationships, in which trust may be greater, or which hold a greater possibility of alternative mechanisms of recourse. This is consistent with similar evidence (e.g., Chaney (2014), Nunn (2007)) on cross-border contracting frictions, and the role of networks in explaining the exporting behaviour of multinational firms. It is also connected with the role of trust in financial markets (e.g., Guiso et al., 2008).

3.5.1 Brokers

To explore this phenomenon further, we consider the role of brokers. We relegate this analysis to the online appendix, but briefly describe it here. For a sub-sample of 8,077 deals in our sample, we obtain information on whether they are intermediated by a broker or not. Overall, we find that nationality bias is not materially affected by the presence of a broker when transactions take place in the home country of the buyer. However, when buyers are trading abroad, the presence of a broker is associated with a significant reduction in nationality bias. This further supports our interpretation of nationality bias as being driven by underlying contracting frictions, which can be at least partially overcome by intermediaries who might be able to certify and vet counterparties.[21]

We now move to analyzing how the availability of same-nationality seller counterparties is related to the emergence of gravity effects first using a reduced form approach, and then setting up and estimating a structural model.

4 Gravity and Counterparties: Reduced-form evidence

In the previous section, we identified that buyers preferentially match with sellers hailing from the same country. The tendency is statistically significant, robust, and economically large. If buyers rationally anticipate that they will use such preferential matching to surmount local contractual frictions, it may be that the cross-border flow of investment to particular locations is in part determined by the availability of same-nationality counterparties in those locations. The benchmark empirical model for bilateral cross-border investment flows is the gravity model, so we begin by estimating this model on the cross-border commercial real-estate investment flow data.

We first estimate a reduced-form "naïve" gravity equation (see Tinbergen (1962)), which conditions the gross investment flow from country i to country k on the physical distance $D_{i,k}$ between them. Letting $N_{i,k,t}$ represent the number of transactions in the data involving buyers from country i and properties located in country k in year t:

$$\log N_{i,k,t} = \mu_i + \mu_k + \mu_t + \beta_0 \log D_{i,k} + \varepsilon_{i,k,t}. \tag{6}$$

[21] In the online appendix, we report subsample analyses of nationality bias. We find strong effects for developers and institutional investors, and insignificant effects for real estate investment trusts (REITs), both when they trade at home and abroad. Since REITs are highly specialized in trading commercial real estate, we regard them as a useful placebo test. Given their business model, we expect REITs to be least affected by issues of trust, search costs, contracting frictions, or information asymmetries, further supporting contract-based explanations.

he coefficient β_0 captures the effect of distance on the magnitude of the cross-border capital flow in commercial real estate between countries i and k.

Equation (6) includes buyer country (μ_i) and location country (μ_k) fixed effects. Head and Mayer (2014) show that the inclusion of these fixed effects makes it less likely that more general buyer and location country determinants of inbound and outbound investment flows affect estimated gravity.[22] Anticipating the inclusion of additional time-varying variables on the right-hand-side of this equation, we also include time fixed effects (μ_t) in the regression.

The leftmost column of Panel A in Table 4 confirms the presence of a very strong negative effect of distance between origin and location countries on cross-border investment flows in the data—perhaps surprisingly, the naïve gravity equation shows a strong role for distance, similar to results from standard trade and investment settings analyzed in many previous papers.[23]

Next, let $N^S_{i,k,t}$ denote the number of transactions involving sellers from country i in properties located in country k. We add this variable to the right-hand-side of equation (6). This gives us a reduced-form estimate of how the density of sellers from the same country in location k affects estimated gravity:[24]

$$\log N_{i,k,t} = \mu_i + \mu_k + \mu_t + \beta_0 \log D_{i,k} + \beta_1 \log N^S_{i,k,t} + \varepsilon_{i,k,t}. \tag{7}$$

Equation (7) looks strange at first glance, as it is obvious that every transaction involving a buyer will also involve a seller. However, the important point to note here is that $N^S_{i,k}$ for each i is the number of sellers present in each location k from the *same* country as the

[22] Using simulated data generating processes consistent with theoretical models including monopolistic competition, heterogeneous consumers, firms or industries, Head and Mayer (2014) also show that fixed effects estimates consistently generate cleaner estimates of gravity.

[23] We note here that the immobility of commercial real estate means that transportation-cost-based explanations for the success of this gravity equation are ruled out. This result is similar to Blum and Goldfarb (2006).

[24] The inclusion of the time fixed effects ensures that any estimated effects don't arise from common time-variation in buying and selling activity.

buyer.[25] The inclusion of this variable is motivated by the finding of nationality bias in the previous section.

When we estimate this equation, the second column of Panel A in Table 4 shows that the density of same-nationality sellers is a strong and statistically significant determinant of the number of buyers from the same country in the same location. Moreover, the inclusion of this variable substantially reduces the estimated coefficient on distance (β_0).

We note that buyers that purchase a property in location k might generate future follow-on purchases by the same buyer in the same location in the future, or there may be unobserved reasons for buyers from location i to be persistently attracted to location k. We account for this possibility by controlling for past buying patterns ($\log N_{i,k,t-1}$) in equation (7):

$$\log N_{i,k,t} = \mu_i + \mu_k + \mu_t + \beta_0 \log D_{i,k} + \beta_1 \log N^S_{i,k,t} + \beta_2 \log N_{i,k,t-1} + \varepsilon_{i,k,t}. \tag{8}$$

The third column of Panel A in Table 4 shows that when we control for the persistence of investment flows by buyer countries into location countries, the current availability of same-nationality sellers remains strong and statistically significant, and the coefficient on physical distance (β_0) shrinks even further.

These results suggest that the availability of same-nationality counterparties in an investment destination is strongly associated with buying activity in that location—this is interesting, and is consistent with our finding of nationality bias in the previous section. However, what is perhaps most striking and puzzling about these results is that there is a corresponding attenuation in the size and significance of the distance coefficient in these estimated equations. To better understand this result, we conduct further empirical tests below.

[25] We demonstrate using placebo simulations in the appendix that this relationship is not mechanical, and to a first approximation, $\beta_1 = 0$ is a good null hypothesis.

4.1 Explaining Counterparty Location Choices

What explains the puzzling finding that including of the density of sellers significantly reduces the size and explanatory power of distance in the gravity model? Figure 3 shows the observed densities of same-nationality counterparties in the data, averaged across all nationalities. The figure clearly shows that same nationality counterparties are distributed log-linearly by the inverse of geographical distance from the origin country of transactions—a "gravity-like" pattern. The combination of this spatial distribution of counterparties and nationality bias is responsible for the set of regression coefficients that we observe in Panel A of Table 4.

To further explore this issue, we regress the location-specific density of same-nationality counterparties on a set of explanatory variables used to capture ties between countries. Several of these variables have frequently been employed in the empirical trade literature, and others are new:

$$\log N^S_{i,k,t} = \alpha + \gamma_1 F^1_{i,k} + \gamma_2 F^2_{i,k} + \cdots + \nu_{i,k,t}. \tag{9}$$

Four of the six variables F indicate (i) whether the two countries share an official language, (ii) whether they have a common border, (iii) share a common colonial history, and (iv) have a currently active free trade agreement. The remaining two F variables capture (v) the intensity of historical shipping traffic between countries i and k, and (vi) average trade flows between countries i and k over the past decade.

Panel B of Table 4 shows the results of estimating (9). We find evidence that both historical factors and current trade flows have a strong role in determining the current density of available same-nationality counterparties. This finding suggests that the physical "beachheads" in which sellers are located is well-predicted by current and historical patterns of trade.

Are investment flows directed towards such historically established beachheads? To ascertain this, we include both fitted and residual components of $\log N^S_{i,k,t}$ from equation (9) separately in the gravity equation:

$$\log N_{i,k,t} = \mu_i + \mu_k + \mu_t + \beta_0 \log D_{i,k} + \beta_1^{fit} \log \widehat{N_{i,k,t}^S} + \beta_1^{res} \widehat{\nu_{i,k,t}} + \beta_2 \log N_{i,k,t-1} + \varepsilon_{i,k,t}. \quad (10)$$

The coefficient β_1^{fit} captures the degree to which current investment flows are directed towards locations with high predicted densities of same-nationality counterparties. Beyond the role of historical ties in determining physical beachheads, the current allocation of investment capital may also be linked to random shocks that change the density of sellers in a particular location. This residual effect is captured by the coefficient β_1^{res}.[26]

When we estimate this equation, the rightmost column of Panel A of Table 4 shows that both β_1^{fit} and β_1^{res} are strong and statistically significant determinants of bilateral cross-border investment flows in the global commercial property market. Moreover, allowing for separate coefficients on the two components of $N_{i,k,t}^S$ makes the estimated coefficient on distance, β_0, statistically indistinguishable from zero.

To explain these results, in the next section we build and structurally estimate a matching model. The model shows that these empirical patterns arise endogenously in a situation in which buyers seek to minimize contractual frictions when trading across borders.

Table 4 shows that investment flows are correlated both with the predicted spatial distribution of same-nationality counterparties as well as with random shocks to this spatial distribution. Our model attempts to rationalize these effects in equilibrium—the phenomenon is similar to Admati and Pfleiderer (1991), where directed search can lead to sunspot equilibria where initial small shock leads to herding and liquidity "spikes".

[26]Interestingly, in addition to the strong contemporaneous role of counterparty availability, in the online appendix, we also document a weak impact of the distribution of same-nationality sellers during the *previous* year. One possibility is that this reflects buyers pre-filtering the space of available locations based on the realized distribution of desirable counterparties in the preceding period.

4.2 Robustness

A note on robustness is in order here before we move to the model: in the online appendix, we confirm that these empirical results hold when we use dollar transaction volumes instead of the number of transactions on both left- and right-hand sides of the equations estimated in this section. In that case, the dependent variable is the log total USD volume invested in country k by buyers that hail from country i, and the counterparty effect is captured by the log total USD amount of proceeds from property sales in country k by sellers that originate from country i.

Our results also remain robust when we consider the entire set of bilateral flows, including any incidence of zero investment flows in the data between pairs of countries.[27] This suggests that there may be a role for the density of same nationality sellers, i.e., potential counterparties, in also determining the locations of international investment, i.e., the extensive margin of foreign investment.

We use an additional placebo approach to verify that our gravity specifications are not picking up a mechanical effect, and to better understand how matching with same-nationality counterparties can be separately identified from the role of pre-existing beachheads. We find that our results only emerge when two conditions hold. The first is nationality bias in matching, and the second is the observed spatial distribution of counterparties. Nationality bias on its own is not sufficient to generate the observed impact of the inclusion of N^S on the distance coefficient in the gravity equation. The spatial distribution of counterparties is necessary but also not sufficient on its own to generate the observed role of distance in the gravity equation.

We now move to the model. We use it both to evaluate the counterfactual gains that can be generated from eliminating this market friction, and more importantly, to better understand the economic forces that generate the observed gravity effects.

[27]Specifically, we employ the Pseudo Maximum-Likelihood (PPML) estimator of Santos Silva and Tenreyro (2006) with several different normalizations of the data. Our results are robust across these alternative estimation methods.

5 Equilibrium Matching Model

In our model, buyers and sellers are randomly matched, conditional on arriving at particular locations. If buyers and sellers are of different nationalities, buyers face a friction which acts as a reduction in the expected value that they can expect to receive from the transaction. The friction doesn't exist if they match with a seller of the same nationality. Anticipating that this will be the case, buyers scale their desired investment in particular locations, to the extent that they can predict the location-specific densities of same- and different-nationality counterparties.

Once we solve the model, we structurally estimate the deep parameters, namely, the size of the friction required to rationalize the observed patterns of matching and prices, and the cost associated with buyers tilting their investments towards particular locations and away from others. We then conduct counterfactual analysis to estimate economic magnitudes of prices and transaction volumes under different scenarios.

Our setup bears some resemblance to Piazzesi et al. (2017), although it differs markedly in several respects. First, our model features a generic market friction which maps to the underlying driver of the observed nationality bias. We think of this friction as a cost of contract enforcement or mistrust, which is more acute between counterparties of different nationalities. Second, we explicitly model heterogeneity in buyer valuations. We do so to capture distortions introduced by the friction—which may inhibit buyers with a sufficiently high valuation from accepting sellers' offers. When evaluating counterfactuals, this explicit modelling of buyer heterogeneity allows us to understand the impact of such distortions better than the more common approach in the search literature, which models random shocks to inventory to move matching rates away from 0 or 1. Finally, we introduce an element of pre-filtering into the buyer's problem, which is a simple approach to mapping the endogenous equilibrium relationships in the model back to gravity equations.

To be clear, the model is not a classic search model, in that we do not explicitly model dynamic decisions. This is because we do not have data on time on market and/or the listing process for the transactions in the data. Instead, we build a random matching model, and later introduce a pre-filtering step in which buyers anticipate the matching equilibrium and adjust their desired portfolio allocations to different local markets accordingly. We model buyers' valuation shocks as independent of their portfolio allocation choices, which allows us to model the pre-filtering step in this fashion.

5.1 The Buyer's Problem

We begin with the buyer's optimization problem associated with the matching equilibrium, conditional on their presence in a given location. We later discuss how the buyer's desired investment into particular locations is affected by their anticipation of this local equilibrium.

When they arrive in a given location, buyers randomly encounter sellers of different types (i.e., same or different nationality).[28] The probability of a successful transaction/match is altered by a friction which reduces the expected value that buyers can realize in the event of a different-nationality counterparty matches. In these encounters, sellers make take-it-or-leave-it offers that buyers can either accept or reject.

The objective function of a representative buyer, conditional on the realization of their private valuation V^B, and on receiving a take-it-or-leave-it offer P from a seller is:

$$U^B = \max\{\underbrace{(1-\lambda)V^B - P}_{\text{accept offer}}, \underbrace{0}_{\text{reject offer}}\}. \tag{11}$$

We assume that the outside option of the buyer is a profit of 0.[29] The parameter λ is the market friction, which captures the fact that the buyer perceives an altered valuation depending on their own type/nationality, and the type/nationality of the seller.

[28]For the purposes of this paper we think of these types as capturing buyer and seller nationality, but our setup is generalizable to any other classification of types.

[29]In the online appendix we discuss normalizing the outside option to $\underline{u}$. We explain in the online appendix how such a normalization is convenient to obtain a log-linear closed-form solution of the model when the pre-filtering step is included.

Before pledging capital to a given market, no buyer has any informational advantage over another. However, after having decided to enter a market, the buyer experiences a private valuation shock drawn from a uniform distribution:

$$V \sim \text{Uniform}(V_{\min}, V_{\max}). \tag{12}$$

Conditional on the private valuation shock, the matching-specific level of the contracting friction, and the quoted price P (which will be endogenously determined in equilibrium), the acceptance probability f characterizes the acceptance decision of the buyer:

$$f = \begin{cases} 1, & U^B > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

f is a key quantity in the model, as it determines both the seller's expected profits, as well as the degree to which the buyer's pre-filtering process (which we describe at the end) results in a successful transaction.

5.2 The Seller's Problem

The seller observes the bilateral friction λ, but not the buyer's private valuation V. They therefore need to form expectations about the likely probability that the buyer will accept their offer, i.e., $E[f] = \int_{V^B} f dV^B$. The seller's asking price is the result of an optimal decision, given the expected acceptance probability, and the seller's private valuation:

$$\max_P E[f] \left(P - V^S\right). \tag{14}$$

Analogous to the information structure on the buyer side, we assume that the private valuation of the seller V^S is realized after a random draw from a uniform distribution:[30]

$$V^S \sim \text{Uniform}(V^S_{\min}, V^S_{\max}). \tag{15}$$

The seller sets the price to maximize the profitability of the transaction, but adjusts the price in order to ensure that the probability that the transaction goes through is sufficiently high. The optimal asking price is therefore achieved when the increase in profit arising from marginally raising the price exactly offsets the effect of a marginal reduction in the price on the expected buyer acceptance rate.

5.3 Equilibrium

Equilibrium in each local market is defined by the acceptance rate f and asking price P such that:

- The acceptance decision of the buyer f is optimal, given the buyer's valuation V and the asking price P.
- The quoted asking price P is optimal given the seller's valuation V^S and the expected acceptance rate $E[f]$.

In the online appendix, we solve the model and show how endogenous quantities respond to variation in the magnitude of the market friction, in particular, how the endogenous response of prices ameliorates the slope of the buyer acceptance rate with respect to the friction λ in equilibrium.

[30] As we describe below, this assumption is not necessary for the solution of the model, but we add it for completeness.

We also show that, subject to a set of weak regularity conditions, the model equilibrium depends only on the average seller valuation $\overline{V^S}$, which we recover as a structural parameter from the data.

The equilibrium solution can therefore be represented by the following non-linear mapping function:

$$[f, P] = M\left(\left[\lambda, V^B_{\min}, V^B_{\max}, \overline{V}^S\right]\right) \tag{16}$$

In the model, volume and prices are tightly related. Under the assumption of rational expectations, seller pricing is match-specific: all else equal, sellers post higher prices when they meet a buyer with their own nationality, and lower prices otherwise. Moreover, sellers adjust their decisions in response to different average valuations. Reductions in average valuation lead them to post lower prices, which in turn generate higher probabilities of matching, and therefore higher expected profits.

5.4 Pre-filtering, Cross-border Investment, and Gravity

Anticipating the distortions they will face following entry into local markets, buyers have the option to pre-filter the space of available investment locations as well as their desired exposure to any given location.

We operationalize this idea in a very simple fashion, by allowing buyers to adjust the number of *desired* transactions $\overline{N}$ that they aim for in each foreign market, given the expected utility level $E[U^B] = \int_\lambda \int_{V^B} U^B d\lambda dV^B$ that is achievable for each individual transaction in that particular location. $\overline{N}$ can also be thought of as a target level of investment. The degree to which this quantity differs from the *realized* number of transactions $N = f\overline{N}$, depends on the degree to which pre-filtering results in successful matches. In equilibrium, $\overline{N}$ in any given market will depend on the matching frictions buyers expect to face in that market. These matching frictions, in turn, depend on the expected location-specific density of same-nationality counterparties $\overline{N}^S$.

Re-adjustment of desired transactions in any location is not completely free; it comes at a cost which we model as $C(\overline{N}) = \frac{\phi}{1+\gamma}(\overline{N})^{1+\gamma}$. We view this cost function as capturing the fact that the benefits of scaling investment in response to anticipated frictions are associated with costs of portfolio concentration which resemble the effect of risk aversion. The parameter γ captures the curvature of the cost function and is informative as to the underlying economics of commercial property capital allocation. The case $\gamma < 1$ is one in which there are positive returns to scale or scope, i.e., as capital allocation to a given market rises, it either makes pre-existing investments more profitable, or generates additional favorable investment opportunities that provide the investor with positive risk-adjusted compensation. In contrast, values of $\gamma > 1$ indicate a market in which there are convex adjustment costs which are more familiar from the investment literature; alternatively one in which the risks of portfolio concentration increase faster than the rate of return.

In the online appendix, we show that these assumptions together deliver an equilibrium in which buyers' optimal capital allocation is well-approximated by the following log-linear relationship, for a general class of utility specifications:

$$\log \overline{N} = \overline{\phi} + \frac{1}{\gamma} \log \overline{N}^S, \tag{17}$$

where $\overline{\phi}$ is a function of deep model parameters and equilibrium values of variables capturing the matching process for each buyer type in each location, and γ is the cost-parameter described earlier.

Equation (17) structurally links the conditions that buyers expect to face in the local market (i.e., the density of same-nationality counterparties) to the global allocation of investment flows. To the extent that the density of same-nationality sellers $\overline{N}^S$ has a distribution that is well predicted by (the inverse of) distance—as Figure 3 convincingly demonstrates, global investment flows will consequently also exhibit gravity effects.

We note that in the absence of the market friction λ, buyers' decisions are independent of seller nationalities. This also means that the pre-filtering problem is trivial—investment flows are just allocated equally across available locations, and there is no tilt towards particular locations as in equation (17). Absent the market friction, the model also implies the absence of gravity effects, since there is no other force in the model that pushes or pulls capital to particular countries.

Put differently, the model shows that contracting frictions that lead to the emergence of a nationality bias are what perpetuates gravity effects in international investment flows. Anticipating the fact that they will find it easier to trade with same-nationality counterparties, investors direct capital towards parts of the world where their compatriots have established beachheads. The spatial distribution of these beachheads, which shows a strong inverse relationship with physical distance, arises from early trading patterns and inherited historical links.

Next, we map our structural model to the data.

6 Structural Estimation of the Model

To begin with, we estimate the four deep model parameters $\left[\lambda, V_{\min}^{B}, V_{\max}^{B}, \overline{V}^{S}\right]$ by matching model-implied quantities with four empirical moments: (a) the average level of nationality bias, (b) the unit normalization of the price level, (c) differential pricing by match type, and (d) the average property valuation uncertainty. In a second step, we compute implied acceptance probabilities f and numbers of meetings $\overline{N}$, which we use to estimate the parameter γ, i.e. the argument in the cost function. We describe this process in detail below.

6.1 Magnitude of the Contracting Friction

In the version of the model presented above, we suppressed all notation identifying buyer countries i, seller countries j, and location countries k. However, when structurally estimating the parameters of the model, we work with observed quantities in the actual data. As a result, our notation must of necessity become richer, and we re-attach the appropriate indexes i, j, and k to the parameters and quantities in the model when describing our structural estimation below.

Nationality bias is what we term the observed tendency of buyers to transact with same-nationality counterparties at a higher rate than if they were simply randomly matching with potential sellers that they encounter. In the model, this phenomenon arises endogenously because of the market friction λ that makes cross-nationality transactions more costly. Put differently, λ drives a wedge between the number of random *meetings* $\overline{N}$ between potential counterparties, and the number of actual *transactions* N that are observed in the data. In other words, in the absence of the friction, buyers would be as likely to accept otherwise equivalent offers from different-nationality counterparties.

The key equilibrium quantity that links the market friction to the emergence of nationality bias is the set of equilibrium acceptance probabilities f_{ij}. Not all meetings lead to a transaction, and the equilibrium acceptance probabilities f_{ij} determine the rates at which meetings and transactions diverge, i.e.:

$$\overline{N}_{ij} = \frac{N_{ij}}{f_{ij}} \tag{18}$$

To structurally estimate a single parameter that captures the average magnitude of the market friction, we need to introduce simplifying notation, and assume that λ_{ij} depends on the nationality i of the buyer and the nationality j of the seller in the following way:

$$\lambda_{ij} = \begin{cases} 0, & \text{if } i = j \\ \lambda > 0, & \text{otherwise.} \end{cases} \tag{19}$$

Using equation (16) to obtain a direct mapping between f and λ for each buyer-seller-location pair, we can then estimate λ such that:

$$E[Bias(\overline{N}_{ii})] = 0, \tag{20}$$

i.e., by imposing the null of no average nationality bias in the rates at which buyers and sellers randomly meet, for all nationalities i, in all locations in the data. This is the first—and most important—moment condition that we use in our structural estimation. We recover the three additional moment conditions from the observed price variation across counterparty match types, as we describe next.

6.2 Prices and Valuations

We normalize the price in the group of transactions that involve buyers and sellers with different nationalities as $\overline{P} = 1$, which implies the following patterns of prices across match types:

$$P_{ij} = \begin{cases} 1, & \text{if } i \neq j \\ 1+\pi & \text{if } i = j \end{cases}, \tag{21}$$

where the marginal price difference π is estimated from a hedonic regression.[31] To quantify this variation of prices across match types, we propose the following standard specification:

$$\ln PSF_q = \alpha + \mu_l + \delta_t + \beta \mathbf{X}_i + \pi 1_{\{\text{same nationality}\}} + \varepsilon_q, \tag{22}$$

where PSF_q is the realized price per square foot for property q in period t and sub-market location l, and γ is a dummy variable that captures the price differential occurring for any transactions between buyers and sellers of the same nationality. Since we are interested in price variation by match type, net of any confounding factors, the fixed effects μ_l and δ_t eliminate the regional and time components of price dynamics, while the property- and transaction-specific control variables $\mathbf{X}_i$ control for other sources of cross-sectional heterogeneity. Table 5 Panel A reports the estimated π coefficient. On average, relative to a match between two parties of different nationalities, when a buyer and seller with the same nationality meet anywhere, the π coefficient shows that there is an increase in the price on average, of 7.36%. This is consistent with the predictions of the model—given that buyers experience the valuation distortion, sellers can set prices higher for same nationality buyers

[31]Note that this normalization of prices also determines the units of measurement for the buyer and seller valuations.

up until their point of indifference—and the empirical results on pricing suggests that this is indeed what occurs.

With the valuation model in equation (22) at hand, we use the estimated residuals to calculate a proxy for within-location valuation heterogeneity:

$$\widehat{\sigma} = \sqrt{E[Var_k(\varepsilon_{i,k,t})]} = 0.318. \tag{23}$$

To calculate the limits of the uniform distribution, we use the estimated standard deviation $\widehat{\sigma} = 0.318$ of residual price shocks, based on the hedonic regression in equation (22). Assuming that the residual valuation uncertainty is exactly mirrored in the cross-sectional heterogeneity of buyer valuations, we impose $Var(V_i) = \overline{V}_B^2 \widehat{\sigma}^2$, which allows us to calculate the lower and upper limits of the uniform distribution $V_{\min} = \overline{V}_B(1 - \widehat{\sigma}\sqrt{3}) = 0.458$ and $V_{\max} = \overline{V}_B(1 + \widehat{\sigma}\sqrt{3}) = 1.512$.[32]

6.3 Estimated Structural Parameters

Table 5 Panel B reports the estimated structural parameters. We find that on average, the friction λ amounts to 9.4% of average prices. Our estimation also suggests that sellers absorb the lion's share of the surplus in this setup—their average valuation is 0.579, relative to an average buyer valuation of 1.022, for a price level normalized at 1.0. This is because sellers form rational expectations about the distribution of buyer valuations and set prices to attract the marginal buyer.

Given the estimated values of matching rates f and the observed distributions of actual transactions N by pairs of buyers and sellers (i, j), equation (18) allows us to estimate counterfactual values of $\overline{N}$ and $\overline{N}^S$.

[32] This result is implied by the expression for the variance of the uniform distribution, i.e., $\sigma^2 = \frac{(V_{\max} - V_{\min})^2}{12}$.

Since the parameter γ conveniently pre-multiplies the counterfactual log density of same-nationality sellers in equation (17), we can estimate it directly from a simple fixed effects specification, analogous to our modeling choices in the context of gravity effects:

$$\log \overline{N}_{i,k,t} = \mu_i + \mu_k + \mu_t + \frac{1}{\gamma} \log \overline{N}^S_{i,k,t} + \varepsilon_{i,k,t}. \tag{24}$$

As described earlier, the parameter γ captures the curvature of the cost function and is informative as to the underlying economics of commercial property capital allocation. Our structural estimation indicates a value of $\gamma = 1.258$,[33] which we interpret as evidence for modest convexity of the search cost function for international commercial property investment, resembling traditional adjustment costs in the literature on investment (see, for e.g., Cooper and Haltiwanger (2006)).

6.4 Counterfactual Prices and Volumes

Having structurally estimated the parameters of the model, we can use it to evaluate counterfactual changes in the number of transactions once we eliminate the friction, i.e., by assuming that $\lambda_{ij} = 0$ for all matches, including those that involve different nationalities $(i \neq j)$. We can do this by assuming that the probability of offer acceptance is always f^{high}, including for cross-nationality meetings:

$$\frac{\Delta N_{i\neq j}}{N_{i\neq j}} = \frac{\sum_{i\neq j}(f^{high} - f^{low})\overline{N}_{ij}}{\sum_{i\neq j}(f^{low})\overline{N}_{ij}}. \tag{25}$$

We can also estimate counterfactual changes in prices:

$$\Delta P_{i\neq j} = \frac{\overline{V^S} + V^B_{\max}}{2} - \overline{P}. \tag{26}$$

[33] The coefficient pre-multiplying $\overline{N}^S_{i,k,t}$ in the estimated version of equation (24) is equal to 0.795, and it is statistically significant for a 1% confidence level, based on two-way standard errors clustered at the buyer and location country level.

Here, $\overline{V^S}$ is the estimated seller valuation and $\overline{P} = 1$ is the (normalized) average price for transactions involving match types where investors have different nationalities, as described above.

Equation (25) shows that the effect of the elimination of the market friction on the number of transactions directly results from the increase in the matching rate between buyers and sellers. Given the particular structure of this model, it is immediate to interpret the increases in transactions as gains in market liquidity. Inventory, i.e., the fraction of initiated sales that do not go through because the buyer does not accept the seller's offer, is simply given by $(1 - f)$, implying that under the counterfactual scenario in which the friction is eliminated, a larger fraction of the market clears.

In Table 5 Panel B, we show that the increase in aggregate transaction volumes when the friction is eliminated is equal to $\frac{\Delta N_{i \neq j}}{N_{i \neq j}} = 6.5\%$ and $\Delta P_{i \neq j} = 7.4\%$. Using global aggregate transaction volumes in 2016 as a reference, the corresponding total increase in volume is US\$ 36.36BN, US\$ 19.43BN which can be attributed to the increase in the number of transactions, and the remaining US\$ 16.93BN to the net price appreciation in the counterfactual equilibrium.

7 Conclusions

Gravity models have served as an empirical workhorse for modelling the behaviour of international trade and investment flows at least since Tinbergen (1962). Yet the underlying reasons for their success have proven elusive.

We use the global commercial real estate market, an important venue for foreign direct investment, as a laboratory to better understand the drivers of gravity. In this market, we document a new "nationality bias," which is the tendency for counterparties of the same nationality to preferentially transact with one another. We find that preferential matching is restricted to same-nationality matches, is unaffected by cultural or linguistic links between nations, and is unaffected by the physical distance between countries. However, we find that nationality bias is far stronger in locations in which the rule of law is weak, suggesting that contracting frictions and trust are its primary drivers.

We connect this new finding of nationality bias to the puzzling role of distance in empirical gravity equations. We first find that reduced-form gravity equations are well-able to explain foreign investment flows between origin and destination countries in this market. We then discover that the inclusion of the destination-specific density of counterparties of the same nationality as the origin country absorbs the role of distance in the gravity equation, and renders it insignificant. We find that this spatial density of same-nationality counterparties has a strong log-linear relationship with the inverse of distance from origin countries, and show that this spatial density is well-predicted by historic shipping routes, common languages, and shared colonial history among other variables. This is consistent with the historical establishment of physical "beachheads" associated with current and historical patterns of trade.

To better understand the underlying economic forces at play, we build an equilibrium matching model of the market. In the model, counterparties are more comfortable trading with others of the same nationality transactions for reasons of ease of contracting and trust. In terms of quantitative magnitudes, we find that the estimated contracting/trust friction is substantial, and conclude that under the counterfactual scenario in which the friction is eliminated, market liquidity and prices in this important market would greatly increase.

Importantly, the model is able to rationalize the puzzling role of distance in the empirically estimated gravity equations in these data. In the model, anticipating that they will face the contracting friction, buyers direct capital towards areas in which they expect to find high densities of same-nationality counterparties. The combination of the historical establishment of beachheads and the contracting friction delivers a new explanation for the ongoing empirical success of gravity equations, a longstanding puzzle in the international trade and investment literatures.

References

Admati, A. R. and P. Pfleiderer (1991). Sunshine trading and financial market equilibrium. *Review of Financial Studies 4*(3), 443–481.

Agarwal, S., T. F. Sing, and L. Wang (2017). Information immobility and learning in commercial real estate markets.

Ahearne, A. G., W. L. Griever, and F. E. Warnock (2004). Information costs and home bias: An analysis of us holdings of foreign equities. *Journal of International Economics 62*(2), 313–336.

Anderson, J. (2011). Gravity model. *Annual Review of Economics 3*(1), 133–160.

Anderson, J. E. and E. Van Wincoop (2003). Gravity with gravitas: a solution to the border puzzle. *American Economic Review 93*(1), 170–192.

Antràs, P. (2003). Firms, contracts, and trade structure. *Quarterly Journal of Economics 118*(4), 1375–1418.

Badarinza, C. and T. Ramadorai (2018). Home away from home? Foreign demand and London house prices. *Journal of Financial Economics 30*(3), 532–555.

Bailey, M., R. Cao, T. Kuchler, and J. Stroebel (2018). economic effects of social networks: Evidence from the housing market. *Journal of Political Economy (forthcoming)*.

Blum, B. S. and A. Goldfarb (2006). Does the internet defy the law of gravity? *Journal of International Economics 70*(2), 384–405.

Branikas, I., H. Hong, and J. Xu (2017). Location choice, portfolio choice. Technical report, NBER Working Paper No. 23040.

Burchardi, K., T. Chaney, and T. A. Hassan (2017). Migrants, ancestors, and foreign investments.

Burchardi, K. B. and T. A. Hassan (2013). economic impact of social ties: Evidence from german reunification. *Quarterly Journal of Economics 128*(3), 1219–1271.

Chaney, T. (2014). Network structure of international trade. *American Economic Review 104*(11), 3600.

Chinco, A. and C. Mayer (2016). Misinformed speculators and mispricing in the housing market. *Review of Financial Studies 29*(2).

Coeurdacier, N. and H. Rey (2013). Home bias in open economy financial macroeconomics. *Journal of Economic Literature 51*(1), 63.

Combes, P.-P., M. Lafourcade, and T. Mayer (2005). trade-creating effects of business and social networks: evidence from france. *Journal of International Economics 66*(1), 1–29.

Cooper, R. W. and J. C. Haltiwanger (2006). On the nature of capital adjustment costs. *Review of Economic Studies 73*(3), 611–633.

Coval, J. D. and T. J. Moskowitz (1999). Home bias at home: Local equity preference in domestic portfolios. *Journal of Finance 54*(6), 2045–2073.

Cvijanovic, D. and C. Spaenjers (2015). Real estate as a luxury good: Non-resident demand and property prices in Paris.

Fouquin, M., J. Hugot, et al. (2016). Two centuries of bilateral trade and gravity data: 1827-2014.

French, K. R. and J. M. Poterba (1991). Investor diversification and international equity markets. *American Economic Review 81*(2), 222.

Garmaise, M. J. and T. J. Moskowitz (2004). Confronting information asymmetries: Evidence from real estate markets. *Review of Financial Studies 17*(2), 405–437.

Garmendia, A., C. Llano, A. Minondo, and F. Requena (2012). Networks and the disappearance of the intranational home bias. *Economics Letters 116*(2), 178–182.

Guiso, L., P. Sapienza, and L. Zingales (2009). Cultural biases in economic exchange? *Quarterly Journal of Economics 124*(3), 1095–1131.

Han, L. and W. C. Strange (2015). Microstructure of housing markets. *Handbook of Regional and Urban Economics 5*, 813–886.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of International Economics*, Volume 4, pp. 131–195. Elsevier.

Huberman, G. (2001). Familiarity breeds investment. *Review of Financial Studies*, 659–680.

Kurlat, P. and J. Stroebel (2015). Testing for information asymmetries in real estate markets. *Review of Financial Studies 28*(8), 2429–2461.

Landvoigt, T., M. Piazzesi, and M. Schneider (2015). Housing market (s) of San Diego. *American Economic Review*, 1371–1407.

Levitt, S. D. and C. Syverson (2008). Market distortions when agents are better informed: The value of information in real estate transactions. *Review of Economics and Statistics 90*(4), 599–611.

Mayer, T. and S. Zignago (2011). Notes on CEPII's distances measures: The geodist database.

Miyakawa, D., C. Shimizu, and I. Uesugi (2016). Geography and realty prices: Evidence from international transaction-level data.

Nieuwerburgh, S. V. and L. Veldkamp (2009). Information immobility and the home bias puzzle. *Journal of Finance 64*(3), 1187–1215.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *Quarterly Journal of Economics 122*(2), 569–600.

Piazzesi, M., M. Schneider, and J. Stroebel (2017). Segmented housing search.

Portes, R. and H. Rey (2005). determinants of cross-border equity flows. *Journal of International Economics 65*(2), 269–296.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics 48*(1), 7–35.

Rauch, J. E. and V. Trindade (2002). Ethnic chinese networks in international trade. *Review of Economics and Statistics 84*(1), 116–130.

Sa, F. (2015). Immigration and house prices in the UK. *The Economic Journal 125*(587), 1393–1424.

Santos Silva, J. and S. Tenreyro (2006). log of gravity. *Review of Economics and Statistics 88*(4), 641–658.

Tesar, L. L., I. M. Werner, and I. M. Werner (1995). Home bias and high turnover. *Journal of International Money and Finance 14*(4), 467–492.

Tinbergen, J. (1962). An analysis of world trade flows. *Shaping the World Economy 3*, 1–117.

Van Nieuwerburgh, S. and J. Favilukis (2017). Out-of-town home buyers and city welfare. In *2017 Meeting Papers*, Number 486. Society for Economic Dynamics.

Table 1
Summary statistics

Panel A reports averages and cross-sectional distributions of selected property-specific variables, for the full sample of 123,648 transactions over the period between January 2007 and October 2017. Panel B reports the composition of the sample by property type, the types of deals, and the fraction of the sample for which the underlying property is located in the Central Business District. Panel C summarizes the information that we have about the buyer and seller types active in the market, by the listing status (i.e. the main source of capital), and the type of operational focus of the company (i.e. the corporate type).

Panel A

	Average	1%	25%	50%	75%	99%
Construction year	1984	1890	1975	1990	2003	2016
Total floor area (in ft^2)	186,631	5,283	51,215	113,845	232,676	1,150,000
Property price (in 2017 USD)	\$39 mil	\$1 mil	\$10 mil	\$18 mil	\$38 mil	\$337 mil
Price per square foot (in 2017 USD)	\$294.4	\$22.2	\$93.1	\$175.7	\$342.2	\$1,984.6

Panel B

Property type	No.	Freq.
Office	40,296	32.6%
Retail	28,875	23.4%
Apartment	26,063	21.1%
Industrial	23,022	18.6%
Hospitality	5,392	4.4%
Location within metropolitan area	**No.**	**Freq.**
Central Business District (CBD)	28,274	22.9%
Outside Central Business District	95,374	77.1%

Deal type	No.	Freq.
Single property	66,371	53.7%
Portfolio of properties	57,277	46.3%
Buyer objective	**No.**	**Freq.**
Investment	109,037	88.2%
Occupancy	3,467	2.8%
Renovation	6,877	5.6%
Redevelopment	4,263	3.4%

Panel C

Source of capital	Buyer		Seller	
	No.	Freq.	No.	Freq.
Private	52,106	42.1%	53,101	43.1%
Institutional	40,917	33.1%	36,611	29.7%
Public	25,055	20.3%	24,489	19.9%
Others	5,570	4.5%	9,114	7.4%
Corporate type	**Buyer**		**Seller**	
	No.	Freq.	No.	Freq.
Developer/owner/operator	45,766	37.0%	49,631	40.2%
Equity fund/investment manager	30,627	24.8%	24,930	20.2%
REIT	17,957	14.5%	16,189	13.1%
Others	29,286	23.7%	32,563	26.4%

Table 2
Nationality Bias

This table reports estimated average nationality bias effects. In the first two columns of Panel B, we use propensity score adjusted fractions of seller nationalities. The transaction characteristics include the transaction year, the property type, and an indicator of property price category, proxied by the within-country within-year price quintile. In the latter columns we calculate nationality bias effects within clusters of $N = 20$ observations, defined by the property location, and by the property location and transaction characteristics, respectively. In parentheses, we report two-stage bootstrap standard errors. *, ** and *** denote statistical significance for 10%, 5% and 1% confidence levels.

Panel A
Estimation results

	Absolute measure		Relative measure	
Average effect	0.020***		0.028***	
	(0.004)		(0.006)	
Nationality bias at home		0.012***		0.038***
		(0.001)		(0.014)
Nationality bias abroad		0.023***		0.025***
		(0.004)		(0.005)
Number of locations	925	925	925	925
Number of countries	70	70	70	70
Number of transactions	87,679	87,679	87,679	87,679

Panel B
Robustness

	Propensity-score adjusted		Clustering by location		Clustering by location and characteristics	
Average effect	0.008***		0.009***		0.023***	
	(0.001)		(0.002)		(0.004)	
Nationality bias at home		0.005***		0.010***		0.011***
		(0.001)		(0.004)		(0.004)
Nationality bias abroad		0.026***		0.009***		0.035***
		(0.003)		(0.003)		(0.004)
Number of locations	925	925	925	925	925	925
Number of countries	70	70	70	70	70	70
Number of transactions	87,679	87,679	87,679	87,679	87,679	87,679

Table 3
Understanding nationality bias

The table reports estimated coefficients from the following specification:

$$Bias(N_{ij,k}) = \mu_i + \mu_j + \mu_k + (\beta_0 + \boldsymbol{\beta_1}\mathbf{F_k})\, 1_{i=j} + \boldsymbol{\gamma}\mathbf{D_{ij}} + \varepsilon_{ij,k},$$

where $Bias(N_{ij,k})$ is the bias measure between buyers from country i and sellers from country j, when transacting in location country k. The variable set **F** contains three location k-specific factors, and **D** contains five measures of geographic and cultural distance between country i and country j. In parentheses, we report standard errors two-way clustered at the buyer and location country level. *, ** and *** denote statistical significance at the 10%, 5% and 1% confidence levels, respectively.

Same nationality	0.031***	0.031***	−0.004
	(0.008)	(0.012)	(0.016)
× Limited rule of law			0.028***
			(0.011)
× Market opacity			−0.011
			(0.014)
× Log GDP^{-1} level			0.015
			(0.011)
Buyer-seller distance		0.005	0.004
		(0.008)	(0.008)
Common border		0.001	0.001
		(0.005)	(0.005)
Common language		0.009	0.010
		(0.009)	(0.009)
Common colonial history		0.002	0.003
		(0.005)	(0.005)
Trade agreement		−0.001	−0.001
		(0.003)	(0.003)
Location country fixed effects	Yes	Yes	Yes
Buyer country fixed effects	Yes	Yes	Yes
Seller country fixed effects	Yes	Yes	Yes
Number of obs.	11,236	11,236	11,236
R^2	0.012	0.013	0.014

Table 4
Estimation of gravity model

This table reports estimated coefficients from different variants of the following estimated specifications:

$$\log N_{i,k,t} = \mu_i + \mu_k + \mu_t + \beta_0 \log D_{i,k} + \beta_1 \log N^S_{i,k,t} + \beta_2 \log N_{i,k,t-1} + \varepsilon_{i,k,t},$$

reported in Panel A, as well as the results of a two-stage procedure where we first regress the density of same-nationality counterparties on a number of exogenous explanatory variables, reported in Panel B, and then use both the fitted and the residual values in the gravity equation:

$$\log N^S_{i,k,t} = \alpha + \tau_i + \tau_k + \tau_t + \gamma_1 F^1_{i,k} + \gamma_2 F^2_{i,k} + \cdots + \nu_{i,k,t}, \tag{27}$$

$$\begin{aligned}\log N_{i,k,t} =& \mu_i + \mu_k + \mu_t + \beta_0 \log D_{i,k} + \beta_1^{fit} \log \widehat{N^S_{i,k,t}} \\ &+ \beta_1^{res} \widehat{\nu_{i,k,t}} + \beta_2 \log N_{i,k,t-1} + \varepsilon_{i,k,t}.\end{aligned} \tag{28}$$

In parentheses, we report standard errors two-way clustered at the buyer and location country level. *, ** and *** denote statistical significance for 10%, 5% and 1% confidence levels.

Panel A
Estimating gravity and counterparty effects

Log Distance	-0.445^{***}	-0.196^{**}	-0.112^{**}	-0.081
	(0.084)	(0.086)	(0.056)	(0.064)
Same-nationality sellers		0.482^{***}	0.276^{***}	
		(0.072)	(0.055)	
Lagged dep. variable			0.444^{***}	0.442^{***}
			(0.037)	(0.036)
Same-nationality sellers				
Fitted value				0.383^{***}
				(0.107)
Residual value				0.274^{***}
				(0.055)
Location country fixed effects	Yes	Yes	Yes	Yes
Buyer country fixed effects	Yes	Yes	Yes	Yes
Transaction year fixed effects	Yes	Yes	Yes	Yes
Number of obs.	1208	1208	1208	1208
R^2	0.501	0.653	0.680	0.680

Panel B

Explaining the density of same-nationality sellers

Common language	0.192***
	(0.052)
Common border	0.245***
	(0.053)
Common colonial history	0.012
	(0.054)
Free trade agreement	0.173***
	(0.053)
Historical shipping route	0.003***
	(0.001)
Log Trade flow volume	0.204***
	(0.024)
Location country fixed effects	Yes
Buyer country fixed effects	Yes
Transaction year fixed effects	Yes
Number of obs.	2789
R^2	0.504

Table 5

Structural estimation of the model

Panel A reports the estimated coefficient γ and the estimated average standard deviation of residuals across locations σ, based on the following hedonic regression specification:

$$\ln PSF_q = \alpha + \mu_k + \delta_t + \beta \mathbf{X}_i + \gamma 1_{\{\text{same nationality}\}} + \varepsilon_q,$$

where PSF_q is the realized price per square feet for property q in period t and location k. μ_k and δ_t are location and time fixed effects, and $\mathbf{X}_q$ are a set of property- and transaction-specific control variables: construction date, functional use, deal type, buyer corporate type, and buyer listing status. The dummy variable $1_{\{\text{same nationality}\}}$ takes the value of one if the buyer and the seller have the same nationality, and zero otherwise. In parentheses, we report standard errors clustered at the level of sub-markets. *, ** and *** denote statistical significance for 10%, 5% and 1% confidence levels. Panel B reports the value of the structural parameters λ, $\overline{V}^B, V^B_{\min}, V^B_{\max}$ and $\overline{V}^S$, as implied by the structural model. The quantitative results are obtained under the assumptions that $\overline{P} = 1$ for matches between buyers and sellers with different nationalities.

Panel A

Hedonic regression

Relative price for same-nationality transactions	γ :	0.0736***
		(0.0088)
Estimated residual price dispersion	$\widehat{\sigma}$:	0.3188
Hedonic control variables		Yes
Location fixed effects		Yes
Year fixed effects		Yes
Number of obs.		123,648
R^2		0.6250

Panel B

Estimated structural parameters

Model parameters	Size of market friction	λ:	0.094
	Convexity of search function	γ:	1.258
	Distribution of buyer valuations	$\overline{V}^B$:	1.022
		$V^B_{\min}$:	0.474
		$V^B_{\max}$:	1.566
	Average seller valuation	$\overline{V}^S$:	0.579
Counterfactual aggregate effects	Number of transactions		0.065
(assuming $\lambda = 0$)	Average price level		0.074

Figure 1
Geographical coverage of the sample

This figure shows the composition of our data set of global commercial property transactions, by the location country of the property, the nationality of the buyer, and the nationality of the seller. We distinguish between transactions for which the buyer and the seller have different nationalities (darker shading), and those for which the buyer and the seller have the same nationality (lighter shading). The transaction-level dataset covers the period between January 2007 and October 2017.

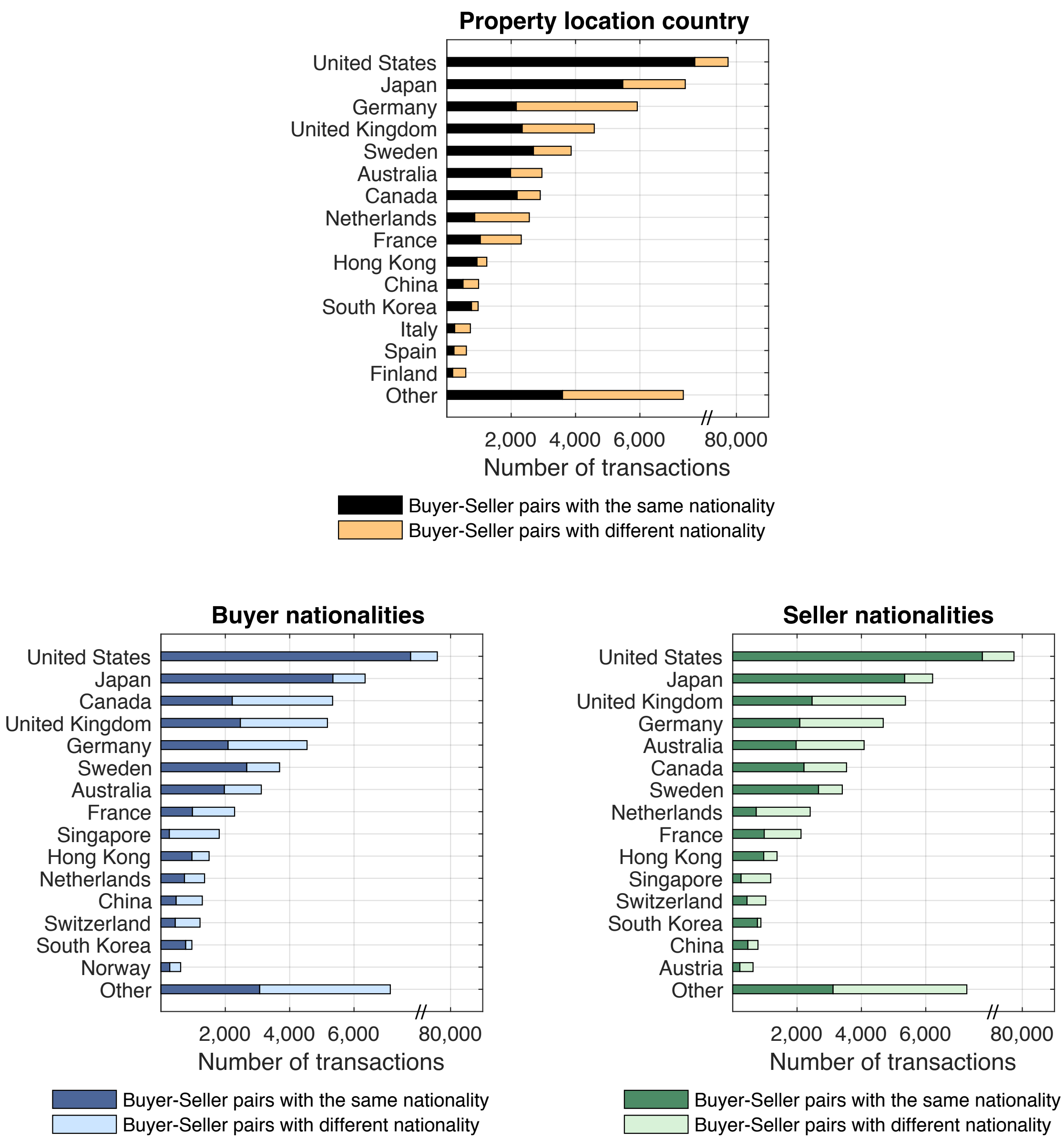
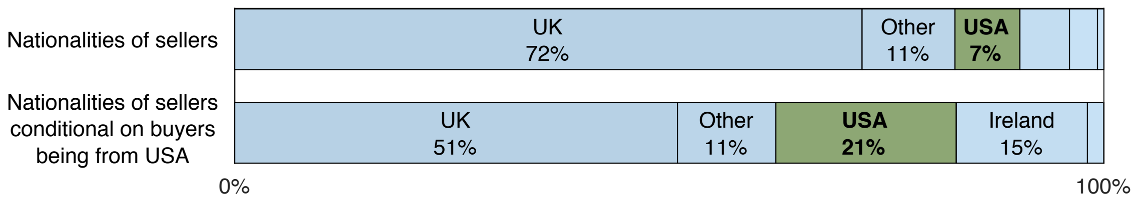


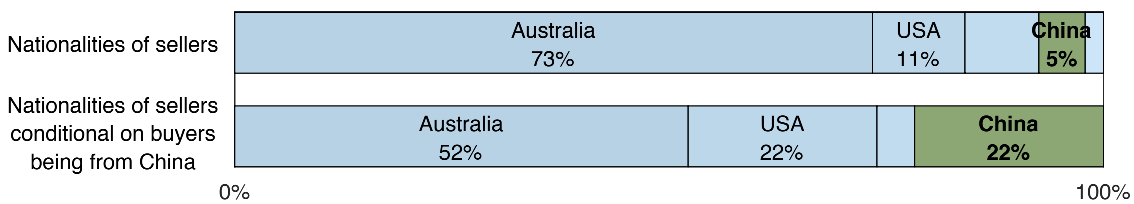
Figure 2
Illustration of the identification method

This figure reports the fractions of transactions for which the sellers have particular nationalities, both unconditionally (top bar) and conditional on the buyer being from a specific country (lower bar). The fractions are calculated within each location separately. For illustration purposes, we report results for three locations (districts/boroughs) in three different countries.

Panel A
West End, London, UK



Panel B
Central Business District (CBD) Midtown, Sydney, Australia



Panel C
Quartier Central des Affaires, Paris, France

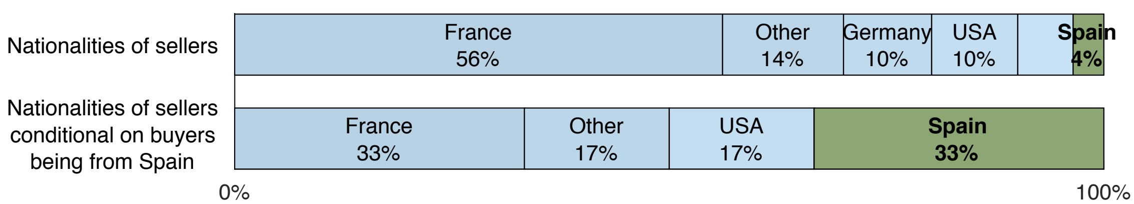
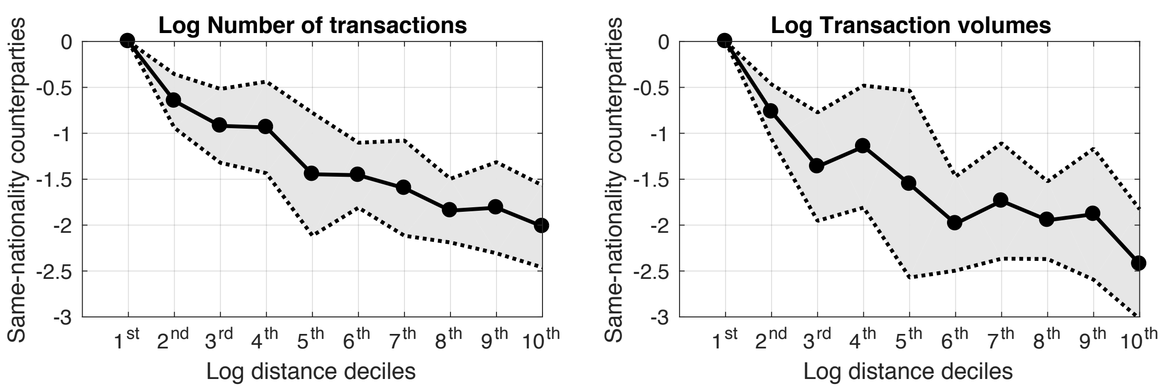


Figure 3
Gravity effects in counterparty availability

This figure reports estimated coefficients δ from the following empirical specification:

$$\log N_{i,k}^{S} = \mu_i + \mu_k + \sum_{q=2}^{10} \delta_q Decile_q(\log D_{i,k}) + \varepsilon_{i,k},$$

where $N_{i,k}^{S}$ is the number of transactions involving sellers from country i and properties located in country k. The plot on the right side repeats the estimation for the case of the corresponding total USD amount. The shaded areas indicate 95% confidence intervals based on standard errors clustered at the buyer and location country level.



Online Appendix for

Gravity, Counterparties and Foreign Investment

Cristian Badarinza and Tarun Ramadorai*

Contents

*Badarinza: Department of Real Estate, Institute of Real Estate Studies, 4 Architecture Drive, Singapore 117566, National University of Singapore, and CEPR. Email cristian.badarinza@nus.edu.sg. Ramadorai (corresponding author): Imperial College London, Tanaka Building, South Kensington Campus, London SW7 2AZ, and CEPR. Tel.: +44 207 594 99 10. Email: t.ramadorai@imperial.ac.uk.

List of Tables

List of Figures

I Robustness checks

I.1 Nationality bias: Placebo simulations

To check whether estimated nationality bias is simply a statistical artefact resulting from the structure of the dataset, arising from spurious rejections of the null, we conduct a placebo test that imposes the null hypothesis $E[Bias(N_i)] = 0$, by reconstructing the sample in each of $n = 1,000$ simulation rounds.[1] In each round, we replace the actually observed seller nationality for each transaction with one drawn at random from the pool of nationalities operating in the respective sub-market. Effectively, this procedure approximates a situation in which counterparties are matched randomly within the sub-market in which they transact. In each simulated sample, we re-compute conditional market shares $\widetilde{h}_{ii} = \frac{\widetilde{N}_{ii}}{\sum_{j=1}^{J} \widetilde{N}_{ij}}$ based on the resulting counterfactually matched transactions $\widetilde{N}_{ij}$. Since the re-sorting is implemented *within* each location, unconditional market shares m_i are unaffected, and we estimate $Bias(N_i) = \widetilde{h}_{ii} - m_i$ when the null is imposed for each nationality i and location k.[2] The results are summarized in Panel A of Figure A.3. We note that in all cases, both at home and abroad, and using both weighted and unweighted measures, the point estimate of nationality bias lies well outside the resulting placebo distribution, strongly rejecting that our estimates arise from a spurious rejection of the null.

I.2 Nationality bias: Base effects

To check for bias in the point estimates arising from the base effect described in the main body of the paper, we run a two-stage placebo test. In this test, we impose the null of random matching between buyers and sellers, but *excluding one buyer nationality at a time*. We then re-estimate nationality bias using the remaining set of nationalities in each placebo simulation round. In this way, we avoid any possible false attribution of nationality bias effects from particular countries to the remaining sample. The results reported in Panel B of Figure A.3 reinforce the robustness of our estimates, and suggest that these base effects play a negligible role. The point estimates of nationality bias lie well outside the resulting placebo distributions, across all simulated scenarios and all levels of aggregation.

[1] It is worth noting that we could still obtain nationality bias in this setup if arrival rates of counterparties into sub-markets were non-random (along a dimension other than nationality), even if matching rates were truly random. The null of no nationality bias essentially assumes this condition is true, which we verify during the simulations.

[2] Note that the counterfactual matches to different seller countries will generate a different partition of the total transactions within each sub-market, so N_{ij} assignments will change, though the total number of transactions in each sub-market location will not.

I.3 Gravity equation: Placebo simulations

In the first placebo simulation, we break any correlation in the data between buyer origin countries and investment location countries, but leave the observed tendency of buyers to preferentially match with sellers of their own countries intact. In the second simulation, we randomly match buyers with available sellers in the data, thus breaking the preferential matching tendency, but leave the correlation between buyer origin countries and investment location countries intact.

Concretely, we construct two sets of $n = 1,000$ simulated samples. In the first, we randomly assign each transaction to a location country that is drawn without replacement from the full set of location countries. This permits the observed preferential matching with same-nationality counterparties, but breaks any tendency for buyers to preferentially allocate capital to particular location countries. In the second sample, we randomly assign to each transaction a seller nationality that is drawn without replacement from the full set of seller nationalities in the original sample, but leave the allocations of capital by buyers to location countries untouched. In each trial, we re-compute the numbers of transactions $N_{i,k}$ and $N^S_{i,k}$ and the distance $D_{i,k}$. We then obtain a distribution of estimated gravity effects using these simulated samples.

Panel A of Figure A.4 reports the simulated distributions of estimated coefficients for the first placebo simulation. The two leftmost plots show that when breaking the observed spatial correlation of investment flows from buyer countries, the gravity effect vanishes, but it does so in all cases. The respective red lines in each plot show the mean of the simulated distributions of coefficients, which are both indistinguishable from zero. Dotted green lines indicate the point estimates from the true data, both of which lie well below the end of the left tail of these distributions. Interestingly, the rightmost plot suggests that in this case the estimated magnitude of the same-nationality effect comes out higher than in the original estimation. This is not surprising, since the placebo imposes random allocation of investment flows across countries, but permits buyers to match preferentially with sellers of the same nationality. Any tilt towards or away from specific countries arising from the availability of same-country counterparties, therefore, is no longer available to explain this preferential matching tendency, leading to all of the weight of preferential matching being absorbed by this coefficient.

Panel B reports simulated distributions from the second placebo trial, which breaks any preferential matching between buyers and sellers of the same nationality. In this case, by construction, the unconditional gravity effect remains unaffected, because buyers continue to invest in the same way in each destination country as in the original dataset. More importantly, the role of same-nationality counterparties is greatly reduced, and the likelihood of observing the original point estimate in a placebo sample is below 2%. This raises our confidence that the estimation of counterparty effects is not a mechanical result of the structure of the data, but rather, driven by the observed pattern of buyer-seller matches. Additionally, we note that in this case the unconditional and the conditional estimates of gravity effects are almost identical, i.e. the inclusion of the variable $N^S_{i,k}$ which measures the availability of sellers from the same country leaves the initial gravity estimate unaffected, unlike our point estimates from the original dataset — the likelihood of observing a decrease of estimated gravity effects of a similar magnitude as in our actual estimation is below 1%.

I.4 Assortative matching by location

In our main results, our approach is to calculate benchmarks m_i at a very granular scale, i.e., the chosen locations are "small" sub-markets within a city, such as districts or boroughs.[3] In Figure A.2, we present the results of an analysis that checks whether this level of granularity is sufficient to eliminate the effect of any spatial clustering by nationalities on our results. We compute Euclidean distances between each commercial property transaction in our dataset and the "central" property transaction in each location. This central transaction occurs in a fictitious location which is the average latitude and longitude across all transactions within the location. When we set locations to be "large,"i.e., countries, these estimated distances to the central transaction are indeed statistically significant for some nationalities. However, when these distances are computed to the "central" transaction in each of the 925 sub-markets that we employ in our main analysis, none of the estimated distances for any country is statistically significant at any conventional level. Put differently, any "between" variation in buyers' preferences for specific areas in a country that are correlated with their nationality is no longer relevant for our estimates, which rely on "within" variation inside narrow sub-markets of cities.

[3] As mentioned in the main body of the paper, we consider locations such as the "West End" borough (London, UK), the "Upper East Side" (New York, USA), the "Quartier Central des Affairs" (Paris, France), "CBD Midtown" (Sydney, Australia), and " Kowloon CBD Core" (Hong Kong) separately, and compute market shares m_i for each such location.

I.5 Assortative matching by characteristics

We estimate a logit propensity score for transaction q to involve a buyer from country i, running regressions for each buyer nationality available in the data:[4]

$$p_{qi} = \Pr(\text{buyer country} = i | X_q).$$

The characteristics X_q that we consider are the year during which transaction took place, the type of property (Office, Retail Apartment, Industrial, Hospitality), and an indicator of price quintile – using the distribution of prices within each country in every given year.

For each location, we apply the Logit propensity scores as weights, to compute a conditional version of m_i:

$$m_i^{matched} = \frac{\sum_{q=1}^{N} \widehat{p}_{qi} 1_{\{\text{seller country}=i|q\}}}{\sum_{q=1}^{N} \widehat{p}_{qi}},$$

which translates into a conditional bias measure:

$$h_{ii} - m_i^{matched}.$$

We report the results of this procedure in Table 2 in the paper.

I.6 Subsample analysis

To better understand how the estimated nationality bias varies across time periods, property types, or buyer objectives, we re-estimate the effects in specific narrow subsamples.[5] In Figure A.5, we show that nationality bias is detectable even when we zoom into these much smaller segments of the market, constructing unconditional market shares m_i in segments defined by specific property and transaction characteristics *within* each location.

First, the results suggest that nationality bias has been a consistent feature of the global commercial real estate market, at least over the last decade. For example, when we restrict the sample to the year 2007 (and therefore also calculate unconditional market shares m_i using only contemporaneous transactions within each location in this year), the average level of nationality bias is 6%, roughly double the level observed after 2010. This pattern

[4]In practice, we restrict this analysis to all nationalities with a sufficient number (25 in our empirical analysis) of transactions, and use the unweighted benchmark estimates for the nationalities with small numbers of transactions.

[5]Importantly, we note that the effects by segment do *not* need to sum up to the average effect. On the contrary, the average effect is filtered out by this procedure, and reference market shares m_i are recalculated in each case using the distribution of seller nationalities within each location × subsample that we consider.

is intriguing. It suggests that during and in the aftermath of the global financial crisis, the underlying drivers of the bias phenomenon have been more pronounced. This is consistent with the breakdown of trust or ease of contracting between counterparties, which suggests that these are possible drivers of nationality bias, as we discuss further below.

Second, we note that nationality bias effects are robust to further conditioning on the buyer's objective. This serves as a way to check that we aren't mistakenly classifying the specialization of companies originating from particular countries in particular types of transactions as a form of nationality bias. Both the magnitudes and the statistical significance are consistent across the two buyer objectives (Investment and Occupancy) that cover around 90% of the sample. The effects are more muted for properties meant for redevelopment or renovation, which is not surprising, given that the purchasing decision is much more property-specific in this case, and less likely to be influenced by considerations relating to the counterparty.

Turning to property-specific robustness, we find that nationality bias effects in central business districts (CBDs) are indistinguishable from those estimated outside the CBDs. Since the within-city location is one of the most important features of commercial property, we view this result as an important further validation of the absence of contamination arising from any spatial clustering. Similarly, we isolate different segments of the market along the property price dimension, distinguishing between relatively low-stakes transactions (below USD 14 million, in the lowest quintile), and high-stakes transactions (above USD 65 million, in the highest quintile). Nationality bias effects are less present at the bottom of the price distribution, but they are much more pronounced at the top. This suggests that frictions affecting different counterparty matches have a larger impact on higher-stakes deals.

In Figure A.6, we explicitly isolate effects for a set of three world regions – distinguishing between the United States, Developed and Developing countries according to the standard IMF classification of economic development levels. The results show a very pronounced pattern of increasing nationality bias between counterparties transacting in countries at the lower levels of development, especially when these counterparties are foreign. In the main body of the paper, we show that this tendency is not driven by the overall level of development of a country, but can be attributed to the particular law enforcement regime in each location.

II Derivation of the model equilibrium solution

II.1 Acceptance rates and optimal pricing

The decision of the buyer, as shown in equation (11) in the paper, depends on the quoted price P, which is endogenously determined in equilibrium. The acceptance probability f characterizes the buyer's optimal decision:

$$f = \begin{cases} 1, & V^B \geq \frac{P}{(1-\lambda)} \\ 0, & \text{otherwise.} \end{cases} \tag{A.1}$$

To understand the main mechanisms operating in the model, it is useful to first consider the following comparative statics:

$$\frac{\partial f}{\partial \lambda} < 0 \text{ and } \frac{\partial f}{\partial P} < 0. \tag{A.2}$$

The first of these derivatives shows that the more intense the friction (i.e., the larger is λ), the lower the probability of acceptance. The second shows that the higher the asking price that the buyer is offered, the less likely they are to accept the seller's offer.

The sellers need to set optimal prices to maximize the profitability of the transaction, but will need to adjust the price in order to ensure that the probability that the transaction goes through is sufficiently high.

The first-order condition for the seller's optimization problem, shown in equation (14) in the paper, implies the following pricing decision:

$$P = V^S + E[f] \underbrace{\left(-\frac{dE[f]}{dP}\right)^{-1}}_{>0}. \tag{A.3}$$

The optimal asking price is therefore achieved when the increase in profit arising from marginally raising the price exactly offsets the effect of a marginal reduction in the price on the expected buyer acceptance rate. As equation (A.2) shows, the derivative in the final parenthesis in equation (A.3) is positively signed. The price therefore depends positively on the seller valuation V^S (as a result of profit-maximizing behavior), as well as on the buyer's expected acceptance rate.

Integrating equation (A.1), we can derive an expression for the acceptance probability as a function of the price:

$$f = \frac{V^B_{\max}}{V^B_{\max} - V^B_{\min}} - \frac{1}{(1-\lambda)(V^B_{\max} - V^B_{\min})} P. \tag{A.4}$$

Substituting equation (A.4) into (A.3) delivers an expression for the pricing equation:

$$P = \frac{\overline{V}^S + V^B_{\max}(1-\lambda)}{2}, \tag{A.5}$$

where $\overline{V}^S \equiv E[V^S] = \frac{V^S_{min}+V^S_{max}}{2}$. The model equilibrium therefore depends only on the average seller valuation $\overline{V^S}$, which we recover as a structural parameter from the data.[6]

Finally, substituting equation (A.5) into (A.4), we obtain the equilibrium acceptance probability for a generic meeting between type-i buyers and type-j sellers:

$$f = \frac{V^B_{\max}}{2\,(V^B_{\max} - V^B_{\min})} \left(1 - \frac{\overline{V}^S}{(1-\lambda)V^B_{\max}}\right). \tag{A.6}$$

II.2 Expected utility

To obtain a log-linear closed-form solution of the model when the pre-filtering step is included, it is convenient to normalize the buyer's utility to a level of $\underline{u}$, which will be a function of equilibrium model parameters, as described below.

To calculate the buyer's equilibrium expected utility $E[U^B] = \int_\lambda \int_{V^B} U^B d\lambda dV^B$, we start by first integrating across the distribution of buyer valuations V^B:

$$E[U^B|\lambda] = \underline{u} + \underbrace{\text{Prob.}(u^B > 0|\lambda)}_{f} \cdot E[U^B|u^B > 0, \lambda] \tag{A.7}$$

Note that the first term in the multiplication is simply the equilibrium acceptance probability f. The second term can be calculated explicitly as a function of the equilibrium optimal

[6]To exclude degenerate corner solutions $f < 0$ and $f > 1$,we need to impose the following regularity conditions on seller valuations: $V^S_{min} > (1-\lambda)(2V^B_{min} - V^B_{max})$ and $V^S_{max} < (1-\lambda)V^B_{max}$. This regularity condition implies that the heterogeneity of seller valuations is slightly lower than the heterogeneity of buyer valuations, which is equivalent to assuming a moderate degree of asymmetric information between buyers and sellers. Since equilibrium only depends on the average seller valuation, we test that $\overline{V^S} \in [(1-\lambda)(2V^B_{min} - V^B_{max}), (1-\lambda)V^B_{max}]$ for the structurally estimated value of $\overline{V^S}$.

price schedule P:

$$E[U^B|u^B > 0, \lambda] = \frac{1}{2}\left(V^B_{max} - \frac{1}{1-\lambda}\underbrace{\frac{(1-\lambda)V^B_{max} + \overline{V}^S}{2}}_{P}\right) \tag{A.8}$$

Conditional on the level of the contracting friction λ, the buyer's expected utility is then given by the following equation:

$$E[U^B|\lambda] = \underline{u} + \frac{1}{8}\frac{\left(V^B_{max} - \frac{\overline{V}^S}{1-\lambda}\right)^2}{V^B_{max} - V^B_{min}}. \tag{A.9}$$

Next, integrating over the density of λ is simple, because of the particular matching structure described in equation (19) in the paper, i.e. normalizing the level of the friction to be equal to zero whenever a buyer meets a same-nationality seller, and equal to $\lambda > 0$ otherwise.

In this case, the likelihood that the buyer will meet a same-nationality counterparty is equal to $\frac{\overline{N}^S}{\overline{N}^{total}}$, where $\overline{N}^{total}$ is the total number of sellers with listed properties in the market.

We then have:

$$\begin{aligned}
E[U^B] &= \underline{u} + \frac{\overline{N}^S}{\overline{N}^{tot}}\frac{1}{8}\frac{\left(V^B_{max} - \overline{V}^S\right)^2}{V^B_{max} - V^B_{min}} + \left(1 - \frac{\overline{N}^S}{\overline{N}^{tot}}\right)\frac{1}{8}\frac{\left(V^B_{max} - \frac{\overline{V}^S}{1-\lambda}\right)^2}{V^B_{max} - V^B_{min}} \\
&= \underline{u} + \underbrace{\frac{1}{8}\frac{\left(V^B_{max} - \frac{\overline{V}^S}{1-\lambda}\right)^2}{V^B_{max} - V^B_{min}}}_{a} + \underbrace{\frac{1}{8}\frac{\left(V^B_{max} - \overline{V}^S\right)^2 - \left(V^B_{max} - \frac{\overline{V}^S}{1-\lambda}\right)^2}{\overline{N}^{total}\left(V^B_{max} - V^B_{min}\right)}}_{b}\overline{N}^S \\
&= \underline{u} + a + b\overline{N}^S.
\end{aligned} \tag{A.10}$$

Note that in the absence of the market friction λ, buyers' expected utility for transacting in a given location is independent of the distribution of seller nationalities in that location ($b = 0$). This means that the pre-filtering problem becomes trivial — investment flows are just allocated equally across available locations. To the contrary, when $\lambda = 0$, we have $b > 0$. In that case, the likelihood of meeting more desirable counterparties increases the buyers' expected utility per transaction, and the buyers' investment flows are consequently tilted towards particular preferred destinations. We explore this phenomenon next.

II.3 Pre-filtering

As discussed in the main body of the paper, we assume that buyers adjust their *desired* number of transactions $\overline{N}$ in response to the opportunities offered by a given market, according to the following optimization problem:

$$\max_{\overline{N}} \overline{N} E[U^B] - C(\overline{N}). \tag{A.11}$$

Here, $E[U^B]$ is the expected utility that is achievable for each individual transactions in that particular location.[7]

Given our functional form assumption for the investment adjustment cost $C(\overline{N}) = \frac{\phi}{1+\gamma}(\overline{N})^{1+\gamma}$, the first-order condition of this simple problem amounts to:

$$\phi \overline{N}^{\gamma} = E[U^B]. \tag{A.12}$$

In the previous section, we have calculated the equilibrium level of this expected utility level, which implies that:

$$\overline{N} = \frac{1}{\phi^{\frac{1}{\gamma}}} \left(\underline{u} + a + b\overline{N}^{S} \right)^{\frac{1}{\gamma}} \tag{A.13}$$

Assuming that $\underline{u} = -a$ and taking logarithms on both sides of equation (A.13), we have:

$$\log \overline{N} = \underbrace{\frac{\log b - \log \phi}{\gamma}}_{\overline{\phi}} + \frac{1}{\gamma} \log \overline{N}^{S}, \tag{A.14}$$

which is equation (17) in the main body of the paper.

[7]This specification of unconditional buyer utility implicitly assumes that eventual buyer-seller matches, as well as the realizations of private valuation shocks are independent across the set of desired transactions. While it may of course be that buyers are acquiring multiple properties at the same time, for which valuation shocks are correlated, our empirical analysis is carried out at the level of deals, so each modeled *transaction* can actually accommodate an arbitrary number of underlying properties.

Tables and Figures

Table A.1
Country-by-country effects

This table reports estimated average relative nationality bias effects $Bias(N_i)$for the countries in our sample that have the highest overall numbers of transactions. We compute weighted averages using country-specific weights in each sub-market. The weights are given by the total number of transactions for which the seller is from country i. The 'Nationality bias at home' and 'Nationality bias abroad' samples capture the cases where the buyer's country of origin is either the same or different from the country where the transacted property is located. We report bootstrap standard errors in parentheses. *, ** and *** denote statistical significance for 10%, 5% and 1% confidence levels.

			Nationality bias				Obs.
	Aggregate effect		At home		Abroad		
United States	0.015***	(0.004)	0.015**	(0.006)	0.014***	(0.005)	54,304
Japan	0.093***	(0.009)	0.098***	(0.016)	0.063***	(0.014)	7,123
United Kingdom	0.045***	(0.007)	0.049***	(0.017)	0.030***	(0.008)	6,453
Australia	0.051***	(0.016)	0.052	(0.034)	0.044**	(0.018)	4,217
Germany	0.031***	(0.007)	0.042**	(0.017)	0.008	(0.007)	3,579
France	0.081***	(0.012)	0.113***	(0.027)	0.021	(0.015)	1,606
Canada	0.025*	(0.013)	0.022	(0.018)	0.048	(0.031)	1,273
Sweden	0.073***	(0.011)	0.080***	(0.021)	0.042**	(0.020)	1,114
China	0.062**	(0.030)	0.079	(0.050)	0.023	(0.031)	1,012
Netherlands	0.122***	(0.020)	0.163***	(0.049)	0.011	(0.012)	959
Hong Kong	0.074***	(0.016)	0.079**	(0.036)	0.019	(0.019)	756
Other	0.079**	(0.031)	0.082***	(0.019)	0.090***	(0.010)	5,285

Table A.2
Estimation of gravity model: Poisson Pseudo ML

In Panel A, we report estimated coefficients from the following estimated specification:

$$I_{ik}^{b} = \beta_0 e^{\mu_i+\mu_k}(D_{ik})^{\beta_3}(I_{ik}^{s})^{\beta_4}(I_{ik}^{b,Lag})^{\beta_5}\varepsilon_{ik},$$

where I_{ik}^{b} is an indicator variable which takes the value of 1 if the number of transactions where the buyer is from country i and the property is located in country k is positive. I_{ik}^{s} is an indicator variable which takes the value of 1 if the number of transactions where the seller is from country i and the property is located in country k is positive. In Panel B, we report estimated coefficients from the following estimated specifications:

$$n_{ik}^{b} = \beta_0 e^{\mu_i+\mu_k}(D_{ik})^{\beta_3}(n_{ik}^{s})^{\beta_4}(n_{ik}^{b,Lag})^{\beta_5}\varepsilon_{ik},$$

where n_{ik}^{b} is the share of transactions where the buyer is from country i and the properties are located in country k, relative to the total number of transactions in country k. n_{ik}^{s} is the share of transactions where the seller is from country i and the properties are located in country k, relative to the total number of transactions in country k. We estimate the models using a Poisson Pseudo-Maximum Likelihood procedure, following Santos Silva and Tenreyro (2006). In parentheses, we report robust standard errors, clustered at the location and buyer country level.

Panel A

	(1)	(2)	(3)
Distance (level term)	-0.010***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
Same-nationality sellers		1.137***	0.903***
		(0.116)	(0.127)
Same-nationality buyers (Lag)			0.502***
			(0.107)
Observations	5340	5340	5340
R^2	0.333	0.406	0.418

Estimation of gravity model: Poisson Pseudo ML
(continued)

Panel B

	(1)	(2)	(3)
Distance (level term)	-0.031***	-0.026***	-0.026***
	(0.005)	(0.004)	(0.004)
Same-nationality sellers		3.421***	3.402***
		(1.037)	(1.041)
Same-nationality buyers (Lag)			0.594
			(0.658)
Observations	5340	5340	5340
R^2	0.289	0.359	0.358

Table A.3
Estimation of gravity model: Extensive margin

This table reports estimated coefficients from different variants of the following estimated specifications:

$$N_{i,k} = \mu_i + \mu_k + \beta_0 \log D_{i,k} + \beta_1 N^S_{i,k} + \beta_2 N_{i,k,Lag} + \varepsilon_{i,k},$$

where $N_{i,k}$ is the number of transactions where the buyer is from country i and the properties are located in country k. $V_{i,k}$ is the respective total USD transaction volume. $N^S_{i,k}$ is the number of transactions where the seller is from country i and the properties are located in country k. Once again, $V^S_{i,k}$ is the respective total USD volume. We compute all variables for the post-2013 period and $N_{i,k,Lag}$ and $V_{i,k,Lag}$ for the pre-2013 period respectively. We extend the coverage of the bilateral investment matrix to include buyer country $\times$ location pairs for which the transaction volume is equal to zero. In parentheses, we report robust standard errors, two-way clustered at the location country and buyer country level.

	Number of transactions			Volume of transactions		
Log Distance	-0.252**	-0.008	0.006	-0.298***	-0.022	-0.002
	(0.085)	(0.024)	(0.018)	(0.086)	(0.048)	(0.043)
Same-nationality sellers		0.188***	0.127**		0.915***	0.558***
		(0.024)	(0.041)		(0.094)	(0.152)
Same-nationality buyers (Lag)			0.498			0.467***
			(0.259)			(0.127)
Observations	6889	6889	6889	6889	6889	6889
R^2	0.114	0.860	0.869	0.144	0.782	0.799

Table A.4
Understanding nationality bias: The role of brokers

Panel A demonstrates that the sub-sample for which we have broker information is representative for the full sample of global commercial property transactions. Panel B reports estimated values of nationality bias in the sub-sample for which broker data is available, distinguish between the situation where the respective transaction is intermediated by a broker (two leftmost columns), and the situation where the transaction is not intermediated by a broker (two rightmost columns).

Panel A

	No of obs.	Price per ft^2	Same nationality
Reference sample	79,603.00	\$361.0	0.771
		(70.6)	(0.085)
Broker sub-sample	8,077.00	\$333.74	0.784
		(24.1)	(0.011)

Panel B

	With Broker		No Broker	
Average effect	0.018		0.013	
	(0.026)		(0.012)	
Nationality bias at home		0.021		-0.003
		(0.032)		(0.015)
Nationality bias abroad		0.013		0.043**
		(0.044)		(0.021)
Number of locations	96	96	300	300
Number of countries	20	20	41	41
Number of transactions	1,698	1,698	6,379	6,379

Table A.5
Understanding nationality bias

The table reports estimated coefficients from the following specification:

$$Bias(N_{ij,k}) = \mu_i + \mu_j + \mu_k + (\beta_0 + \boldsymbol{\beta_1}\mathbf{F_k} + \boldsymbol{\beta_2}\mathbf{F_i})\, 1_{i=j} + \boldsymbol{\gamma}\mathbf{D_{ij}} + \varepsilon_{ij,k},$$

where $Bias(N_{ij,k})$ is the bias measure between buyers from country i and sellers from country j, when transacting in location country k. The variable set **F** contains three country-specific factors, measured both in the location country k and the buyer's country of origin i. The variable set **D** contains five measures of geographic and cultural distance between country i and country j. In parentheses, we report standard errors clustered at the country level. *, ** and *** denote statistical significance at the 10%, 5% and 1% confidence levels, respectively.

Same nationality	0.031***	0.031**	0.019
	(0.009)	(0.013)	(0.103)
× Limited rule of law in location country			0.031**
			(0.013)
× Limited rule of law in buyer's country			−0.014
			(0.012)
× Market opacity in location country			−0.023
			(0.028)
× Market opacity in buyer's country			0.015
			(0.024)
× Log GDP^{-1} level in location country			0.016
			(0.011)
× Log GDP^{-1} level in buyer's country			−0.003
			(0.007)
Log Buyer-seller distance		0.005	0.005
		(0.008)	(0.008)
Common border		0.001	0.001
		(0.005)	(0.005)
Common language		0.009	0.010
		(0.007)	(0.007)
Common colonial history		0.002	0.003
		(0.005)	(0.005)
Trade agreement		−0.001	−0.001
		(0.003)	(0.003)
Location country fixed effects	Yes	Yes	Yes
Buyer country fixed effects	Yes	Yes	Yes
Seller country fixed effects	Yes	Yes	Yes
Number of obs.	11,236	11,236	11,236
R^2	0.012	0.013	0.014

Figure A.1
Location of transactions in the data

In this figure, the red marks indicate the locations of commercial property included in our transaction-level dataset. The source of the data is Real Capital Analytics.

Figure A.2
Spatial clustering of commercial property transactions

This figure demonstrates that aggregating at the sub-market level is sufficient to eliminate spatial clustering of commercial property transactions by the buyers' nationalities. We report T-statistics for each of the country-specific coefficients γ_i, from the following estimated specification:

$$D_q = \alpha + \sum_{i=1}^{I} \gamma_i + \varepsilon_q,$$

where D_q is the Euclidean distance between property q and the center location of properties in a given location. In the left panel, we calculate the distance to the average location of transactions occurring in the same country. In the right panel, we calculate the distance to the average location of properties occurring in the same sub-market within a city. To isolate the country-specific clustering for buyers originating from country i, we restrict the set of transactions to the cases where the buyer is a foreigner. The red lines indicate critical values for 90% (dotted line) and 95% (continuous line) confidence levels.

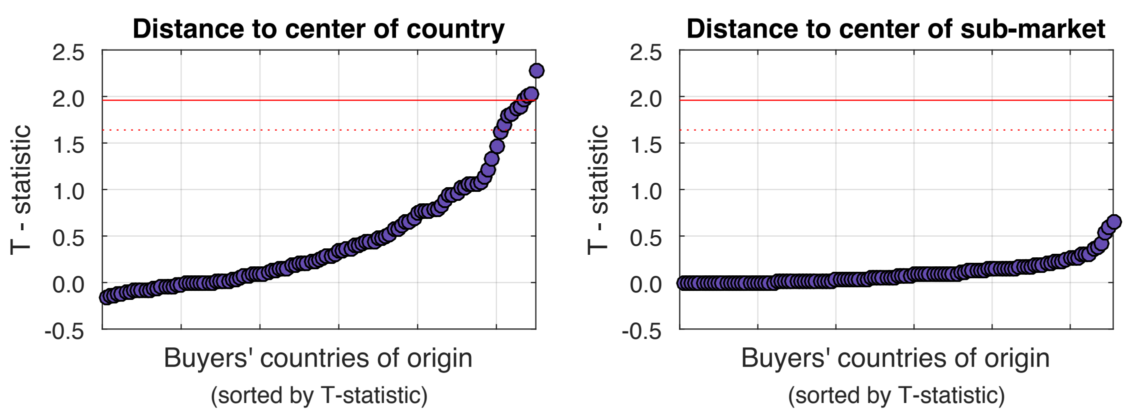


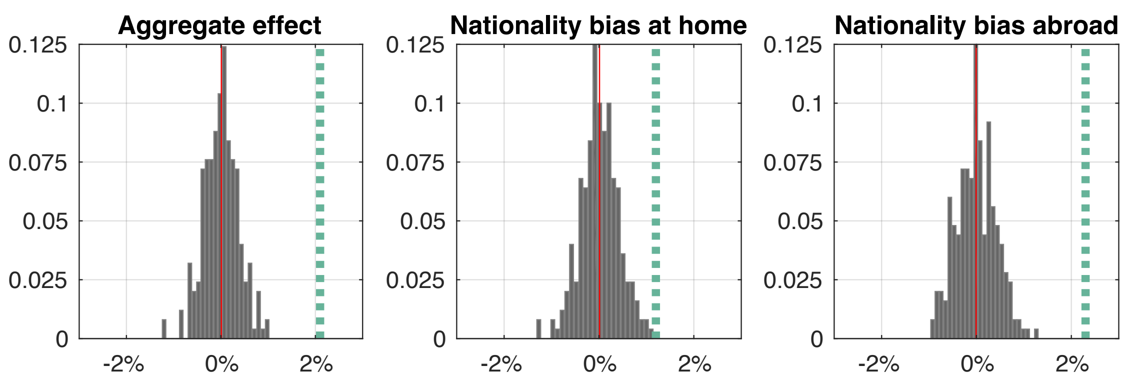
Figure A.3

Placebo tests

This figure reports the distribution of estimated average nationality bias abroad effects across a set of placebo samples, where we randomly re-assign the countries of origin of sellers (Panel A). We consider $n = 1,000$ iterations. In Panel B, we implement a two-stage placebo test where we impose the Null hypothesis of random matching between buyers and sellers, excluding one buyer nationality at a time and estimating nationality bias on the remaining set of nationalities. The dotted green lines indicate point estimates of nationality bias measures, computed using equal-weighted averages.

Panel A

Standard method



Panel B

Accounting for base effect

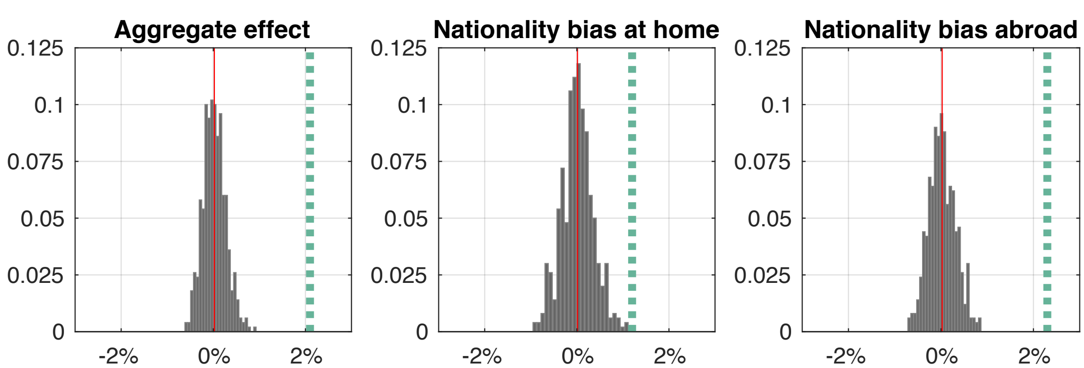


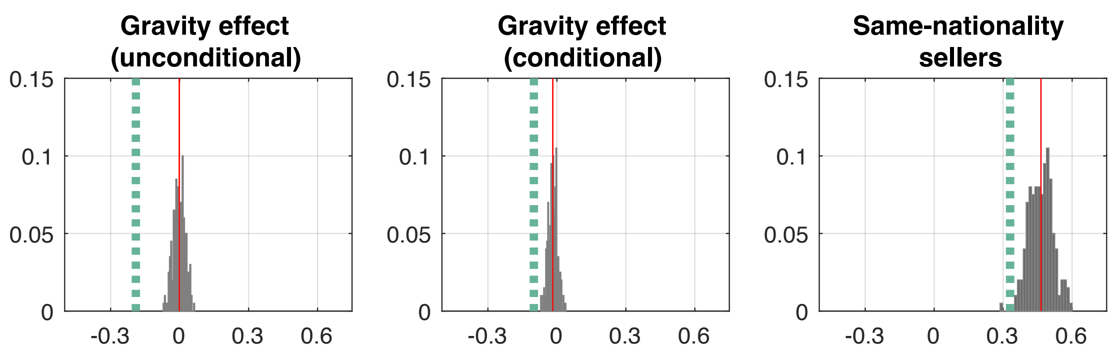
Figure A.4

Gravity effects: Placebo tests

This figure reports the distribution of estimated gravity and same-nationality counterparty effects across a set of placebo samples, where we randomly re-assign location countries (Panel A) and countries of origin of sellers (Panel B). We consider $n = 1,000$ iterations. The dotted green lines indicate point estimates from our benchmark setup with buyer country and location country fixed effects, controlling for the distribution of past transactions. The red lines indicate means of the respective placebo distributions.

Panel A

Random assignment of location country



Panel B

Random assignment of counterparty

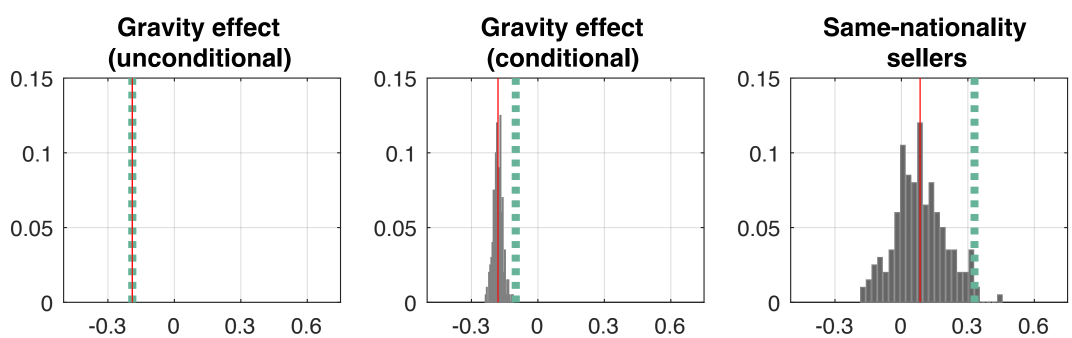


Figure A.5

Subsample analysis

This figure reports estimated average relative nationality bias effects across sub-market segments and countries, constructed within samples defined by each of the variables on the left-hand side of the graphs. Error bars indicate statistical significance for a 10% confidence level.

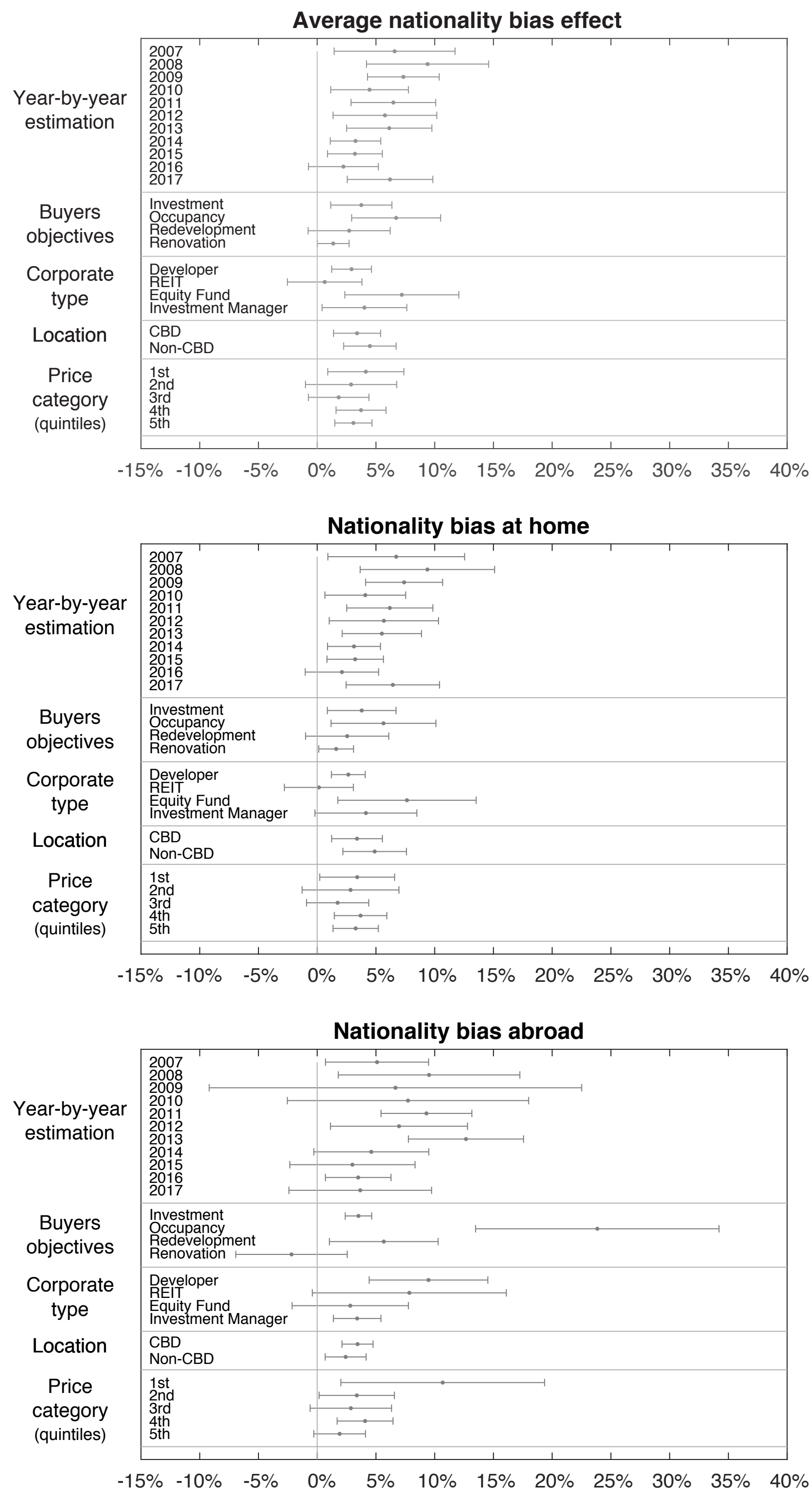


Figure A.6
Nationality bias: Effects across world regions

This figure reports average relative nationality bias effects, for three groups of location countries: the United States (USA), developed countries, and developing countries, using the classification of the International Monetary Fund (IMF). We compute weighted averages using country-specific weights in each sub-market. The weights are given by the total number of transactions for which the seller is from country i. The 'Nationality bias at home' and 'Nationality bias abroad' samples capture the cases where the buyer's country of origin is either the same or different from the country where the transacted property is located. Error bars indicate statistical significance for a 10% confidence level.

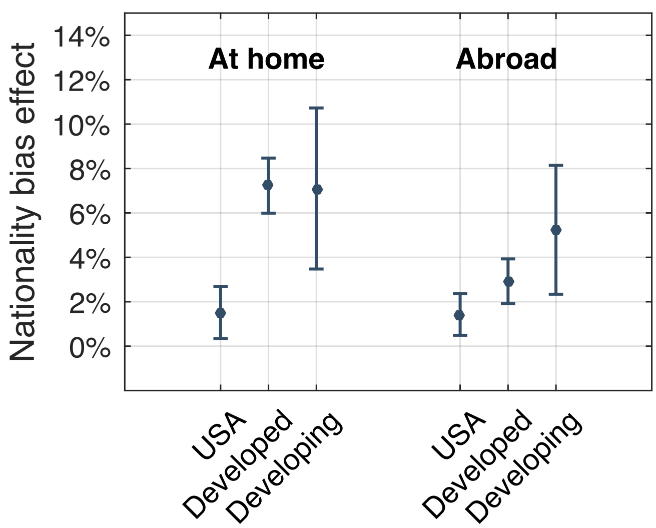


Figure A.7
Adjustment of seller fractions using propensity score matching

This figure illustrates the adjustment of fractions of seller nationalities, controlling for possible assortative matching between buyers and sellers. For each transaction, we compute the likelihood that the transaction involves a buyer from country i, and use the resulting propensity scores as matching weights, to compute adjusted fractions of seller nationalities ($m_i^{matched}$). The set of conditioning variables includes the year during which the transaction took place, the type of property (Office, Retail etc.), and an indicator of the price quintile, calculated using the distribution of prices within each country in any given year.

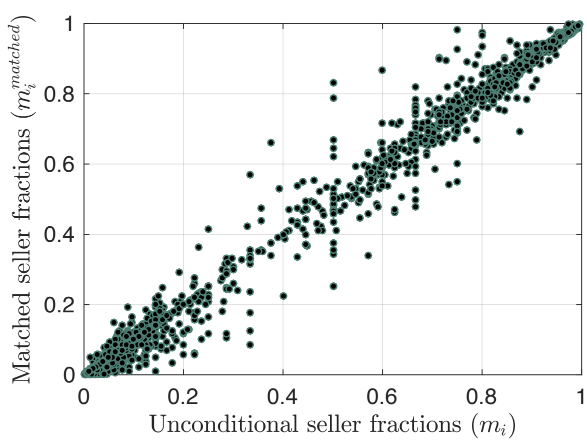


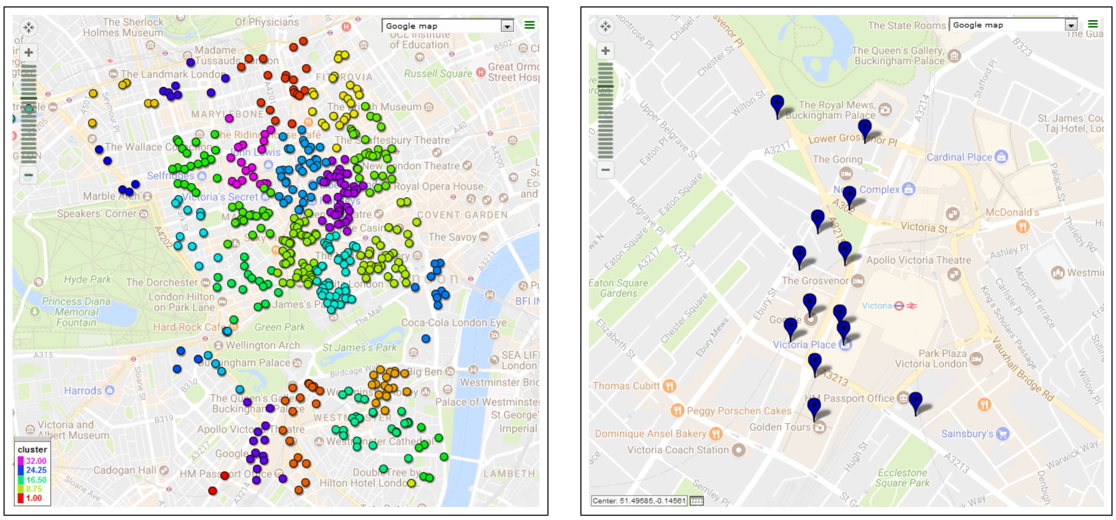
Figure A.8

Illustration of the K-means clustering approach

In Panel A, we determine cluster allocations based on the geographical location of the property. The left sub-panel shows a map of the entire sub-market. The left sub-panel restricts the view to a typical within the sub-market. In Panel B, we use the geographical location of the property together with other transaction characteristics (the year during which the transaction took place, the property type, and the property price category, proxied by the within-country within-year price quintile). We indicate individual properties with a colorized solid circle. The color of the circle indicates the cluster to which the respective property has been allocated.

Panel A

Clustering by location



Panel B

Clustering by location and property characteristics

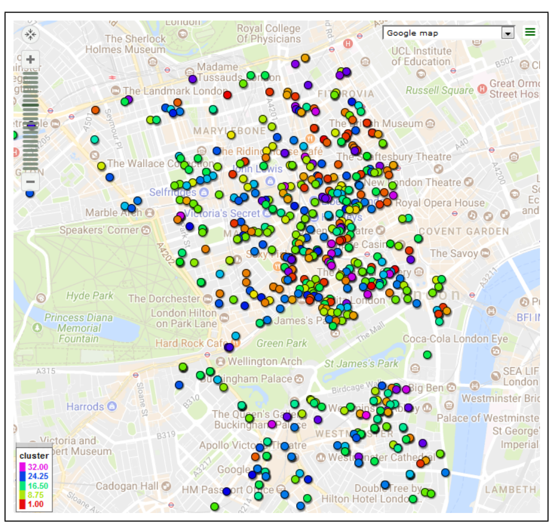


FIGURE A.9

Illustrating the endogenous response of volumes and prices

This figure reports the adjustment of model quantities in response to changes in the market friction. The quantitative results are obtained under the assumption that $\overline{P} = 1$ for matches between buyers and sellers with different nationalities, and for the estimated values of the structural parameters.