Gravity, Counterparties, and Foreign Investment

Cristian Badarinza, Tarun Ramadorai, and Chihiro Shimizu†

October 25, 2018

Abstract

The importance of distance in gravity models of international trade and investment poses a continuing puzzle. Using comprehensive data which record buyer and seller nationalities and investment destinations in cross-border high-value commercial real estate investments, we find a robust tendency of counterparties to preferentially match with others from their own country, especially in poorly-governed locations. We also find that the role of distance in a naïve gravity equation is well captured by the location-specific density of same-nationality counterparties. To explain this reduced-form evidence, we build and structurally estimate an equilibrium matching model with a contracting/trust friction that affects cross-nationality transactions.

*We thank seminar participants at the National University of Singapore, Warwick Business School, London Business School, Imperial College Business School, Tinbergen Institute, and John Campbell, Darrell Duffie, 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: Nihon University, Tokyo, and Center for Real Estate, MIT, Cambridge, Massachusetts. Email: shimizu.chihiro@nihon-u.ac.jp.
1 Introduction

Gravity models are very successful at explaining international trade and investment flows, though the underlying reasons for their success are the subject of active investigation.¹ A continuing puzzle about these models is why trade and foreign investment flows decline substantially with the physical distance between origin and destination countries. A promising class of explanations highlights the role of informational and contracting frictions between counterparties, and analyzes the formation of networks across borders that facilitate transactions.²

We provide new evidence on these issues, and structurally estimate a model to explain the new facts that we uncover. The evidence comes from comprehensive data which cover all high-value transactions in over 70 countries in the global commercial real estate market, an important venue for cross-border investment. Uniquely, these data identify both counterparties in each transaction, as well as their countries of incorporation. In the data, the nationalities of buyers and sellers often differ from one another, as well as from the destination country of the investment. This allows us to assess how matching between counterparties from different nations contributes to the empirical success of gravity equations—which explain bilateral flows between origin and destination countries, but do not include information about the provenance of the seller.

The data reveal that buyers have an unusually strong tendency to transact with sellers who are from the same country as themselves, especially when these transactions occur abroad, i.e., in destinations other than counterparties’ common country of origin.³ We term this pronounced preference to transact with counterparties from the same country

---

¹Anderson (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.

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

³We use the domicile status of firms interchangeably with the term “nationality” in what follows.
nationality bias. This tendency is widespread, showing up for different nationalities and in different locations, and it is economically large and statistically robust.

To assess how nationality bias affects estimated gravity, we begin by estimating a standard naïve gravity equation, which explains the log volume of investment between pairs of buyer origin countries and investment destination countries using the log of physical distance between origin and destination. As with most estimated gravity equations in the literature, we find that in this setting, the coefficient on distance is also significantly negative. We then add to this equation the transaction volume in the destination country generated by sellers who hail from the buyer’s country of origin, to capture the location-specific density of potential same-nationality counterparties. This variable contributes substantial explanatory power, and more importantly, its addition renders the estimated coefficient on log distance in the augmented gravity equation indistinguishable from zero. This result holds true controlling for the past volume in the location country generated by buyers from the origin country, and is not mechanical, as we confirm using simulations. We view this striking observation as reduced-form evidence of an underlying equilibrium relationship, and it motivates a deeper investigation of the drivers of nationality bias.

We more precisely measure nationality bias using the simple null hypothesis of no systematic preferential matching between buyers and sellers based on their country of origin. If this null were true, then the fraction of all observed transactions that involve sellers from a particular country (we call this the “benchmark seller fraction”) would be exactly the same as the fraction of sellers from that country in the subsample of transactions that involve buyers from the same country as the seller. Alternatively, if buyers from a particular

---

4These equations are estimated with log GDPs of the origin and destination countries, as well as with origin and destination country fixed effects as recommended by Head and Mayer (2014), who 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. The fixed effects model is increasingly preferred because it allows for drivers of “gravitational” forces other than the overall GDP level of a country.
country prefer same-nationality counterparties, we would, as we do, observe a systematic bias, with the fraction of sellers in transactions with same-country buyers being far higher.\footnote{As an illustrative example, consider a specific region of the world, say the UK, in which transactions occur, and assume that Indian sellers account for one-tenth of all transactions in this location, regardless of buyer nationality. Under the null of no nationality bias, Indian sellers should also constitute a tenth of all transactions conducted by Indian buyers. If the fraction of Indian sellers in transactions involving Indian buyers is greater than a tenth, we conclude that there is a systematic preference for Indian buyers to transact with Indian sellers.}

The magnitudes that we estimate are economically large. When buyers transact in properties located in their country of origin, they are on average 2\% more likely to match with sellers of their own country relative to the benchmark seller fraction. However, when buyers transact in foreign locations, the corresponding increase in their propensity to match with same-nationality sellers is a substantial 44\% of the unconditional fraction of sellers from foreign countries. Put differently, nationality bias is far stronger abroad than it is at home. We also find that the prices of properties involving buyers and sellers from the same country are higher, on average, by 7.36\% controlling for a range of hedonic characteristics, time, and region effects. This is substantial, considering that the average transaction in the data is US$ 39MM.

A possibility that we evaluate carefully is that particular nationalities might have preferences for particular property characteristics or locations, which we attribute instead to a counterparty preference.\footnote{This is an interesting possibility, and if true, would change the interpretation of our findings. Under this interpretation, the role of counterparty nationality in the gravity equations is different—it would, in this case, simply be a better measure of the “distance” or affinity between origin countries and specific attributes of the investment.} To check whether this is the case, we conduct a number of additional tests. First, we adjust the benchmark seller fractions using a propensity score match to correct for any preference of individual nationalities for particular property characteristics. Second, we adopt a more non-parametric approach by clustering properties on the basis of location and other property characteristics into very small groups (of 20 properties) using a k-means clustering algorithm and re-estimate nationality bias within these clusters. Third, we conduct a placebo analysis in which we randomly reassign nationalities to sellers,
and strongly reject the possibility that our estimates arise from spurious rejections of the null. None of these tests greatly affects our point estimates of this bias, suggesting that matching to characteristics isn’t the main driver of the counterparty matching patterns in the data.

We find that nationality bias is strongest when the GDP and transparency of the market in the investment destination country are low. We also find that there is no gravity in nationality bias, i.e., preferential counterparty matching is restricted to narrow same-nationality matches, and there is no increased tendency over and above this for buyers to transact with counterparties from physically proximate countries. These and other facts that we uncover suggest either that contracting frictions across national boundaries are severe (see e.g., Nunn (2007)), or that trust and network formation is restricted to narrow domains within national boundaries, at least for counterparties in the commercial real estate market.

We also conduct a range of subsample analyses, and show that nationality bias robustly shows up in all time periods in the data, in a majority of the investment locations across the world, for a wide range of nationalities of buyers and sellers, across different corporate objectives of buyers, and is significantly larger during the financial crisis, as well as for the largest dollar value transactions in the data. In addition to checking the robustness of nationality bias, these facts further support the possibility that trust and contracting frictions explain the patterns of preferential matching that we find.

Finally, to better understand the underlying economics driving these results, we set up an equilibrium matching model with heterogeneous buyers and sellers, random matching, and endogenous determination of volumes and prices in a rational expectations equilibrium. The key assumption in the model is that transactions with some counterparties are subject

---

7We measure this using the Jones-Lang-Lasalle (JLL) index of commercial real estate market transparency.
to a friction which affects their expected value.\textsuperscript{8} This permits same-nationality matches to be preferred to other-nationality matches. In the model, 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 it to find the size of the friction that is consistent with the data. On average, the friction amounts to an expected value reduction of 9.4\%, meaning that buyers would be willing to pay this high a premium to avoid transacting with counterparties with different nationalities. We evaluate the counterfactual by eliminating this friction from the model. In the resulting 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 per transaction of 7.4\%. The combination of these two changes generates a substantial increase in the counterfactual volume of cross-country investment.

We then simulate data from the model using the estimated 9.4\% valuation reduction friction, and find that distance shows up in a naïve reduced-form gravity equation estimated using the simulated data. The magnitude of the friction is sufficient to generate a distance coefficient which explains 8\% of the size of the observed distance coefficient in the reduced-form gravity equation estimated in the data.

The underlying economic mechanism generating gravity in the matching model is nationality bias—buyers match with sellers relatively more often in countries with higher densities of same-nationality counterparties. The spatial distribution of densities of same-nationality counterparties is exogenous to our model; we estimate it using the distribution of observed

\textsuperscript{8}We interpret this friction as a generic representation of difficulties in contracting, or a lack of trust that affects transactions with different nationality counterparties, especially in poorly governed investment destinations, as suggested by the correlations between estimated nationality bias and variables capturing the quality of local governance.
transactions involving same-nationality sellers across investment locations. We find that this measure of same-nationality counterparty density declines log-linearly with physical distance from buyer countries.

Put differently, the model can explain the persistence of gravity using nationality bias, but not the origins of gravity. If gravity determined the locations of outbound investment at some point in history, matching frictions generating nationality bias are a force which perpetuate the role of distance in estimated gravity equations. To understand this, consider a situation in which buyers located in a specific country initially invest in destination countries in a manner that declines with distance between origin and destination countries. Given this initial distribution, and the strong tendency of future buyers to match with same-nationality sellers, the model generates a continuing role for distance in the gravity equation.\(^9\)

In addition to the large literature on gravity models mentioned earlier,\(^10\) our work is related to the growing literature on the role of networks, affinity, and trust in international trade and finance.\(^11\) It is also related to the literature on home bias at home and abroad.\(^12\) Our use of commercial real estate market data connects the paper to the growing literature on information asymmetries and social networks in real estate markets\(^13\). Our theoretical work builds on frameworks developed by Han et al. (2015), Landvoigt et al. (2015), and Piazzesi et al. (2017) on segmented housing search, but extends this literature in two ways, introducing a new matching friction to capture nationality bias in the model, and explicitly

\(^9\)It is important to recall here that we find no gravity in nationality bias itself. The decline in density of preferred counterparties with log distance is thus the only force in the model that can deliver the observed distance coefficient in the naïve gravity equation.

\(^10\)Other important papers in this literature include Anderson and Van Wincoop (2003), and Antràs (2003).

\(^11\)See, for example, Combes et al. (2005), Guiso et al. (2009), Garmendia et al. (2012), Burchardi and Hassan (2013), and Burchardi et al. (2017).

\(^12\)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).

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.\textsuperscript{14}

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 and reports estimates of this bias. Section 4 estimates gravity equations, and connects nationality bias with gravity. Section 5 investigates the drivers of nationality bias. Section 6 introduces the equilibrium matching model, and Section 7 describes how we structurally estimate the model and use it to evaluate counterfactuals. Section 8 concludes.

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 MM.

\textsuperscript{14}See also Sa (2015), Cvijanovic and Spaenjers (2015), Miyakawa et al. (2016), and Agarwal et al. (2017).
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 or REIT etc.), and the stated objective (of the buyer) for the property purchase (investment, occupancy, redevelopment, or renovation).

The most important piece of information for the purposes of this paper is the country in which each entity is incorporated; this information is also explicitly captured by RCA for all buyers and sellers, and is what we use to determine the location/nationality of the buyers and sellers. If buyers or sellers are multinational entities, we also know whether the property was bought by the holding company itself, or by a local branch of the holding company. When classifying the nationality of buyers and sellers, we use the country of incorporation of the actual party that was involved in the transaction (for example, the local branch), regardless of the location 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 MM. 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 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 of each metropolitan area (e.g., boroughs 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 outside the CBD. Panel B also shows that a majority of the deals (53.7%) involve the transaction of a single property, but 46.3% of the deals involve multiple properties. We check robustness to this, but simply refer to transactions and deals interchangeably in what follows.\(^{15}\)

### 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.

Figure 1 shows the main countries in the data in which properties are located, as well as the principal nationalities of buyers and sellers. The top panel shows that more than half of the transactions in the sample take place in the US. Outside of the US, the largest markets are Japan, Germany, the United Kingdom, and Sweden.

In the figure, the lighter portion of each bar indicates the fraction of the transactions in each country (or involving specific buyer or seller nationalities) between counterparties with different nationalities, while darker shades indicate matches between counterparties with the same nationality.

\(^{15}\)We implement our analysis at the level of deals. 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 consider the location of the property with the highest value. The estimation of nationality bias is robust to working with individual transactions rather than deals - the estimated average relative bias is equal to 0.082 in that case, and statistically significant at the 1% confidence level.
The figure shows that the United States appears to be a highly local market, with most buyer-seller pairs sharing the same nationality (this happens to be US buyers matching with US sellers in the US). In contrast, properties located in most other countries have far larger shares of transactions which involve buyers and sellers of different nationalities—which is associated with the greater prevalence of foreign investment in commercial real estate in these countries.

From this simple look at the data, most buyer and seller countries also 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 important as those would be difficult to trace back to the true origin country of the investment flows.

In our main analysis, we classify each transaction on the basis of buyer and seller nationalities, distinguishing between situations in which counterparties from different countries transact with one another (e.g., a French company purchases the property from a German company) and situations in which counterparties are incorporated in the same country (e.g., French buyers transacting with French sellers). We further distinguish between transactions occurring “at home” (e.g., a Chinese buyer purchasing from a Chinese seller in China) and “abroad” (e.g., a Chinese company trading with another Chinese company in Germany).

### 2.3 Company Characteristics

We collect ownership information from Bureau van Dijk’s Orbis database and check media reports for evidence of M&A activity between buyer and seller companies. To reduce the amount of manual matching of the company details to Orbis records, we restrict this procedure to the transactions where the buyer and the seller are incorporated in the same country. This 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 total number of observations mentioned above (123,648) is net of this data cleaning and is the number of data points in our final sample.

2.4 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 in properties located in the West End of London that take place over our sample period. 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 and unconditional shares of US sellers, i.e., 21% and 7%, gives us the measure of nationality bias 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 our 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%. 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% of all transactions involving a Spanish buyer.

3 Counterparty Matching: Nationality Bias

In this section, we more formally define nationality bias – the measure is very similar to previous measures proposed in the home bias literature (see, for example, Coval and Moskowitz (1999)) – and estimate it using the transactions in the data. We then link this
measure to estimated gravity in the subsequent section. We consider a range of checks to verify that nationality bias is robust in the section thereafter.

3.1 Measurement

Consider a specific location (such as the Upper East Side) in which companies of different nationalities meet and trade commercial property. In this location, let $N_{ij}$ be the total number of transactions in which the buyer is from country $i = 1, \ldots, I$ and the seller from country $j = 1, \ldots, J$.

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

$$\sum_{i=1}^{I} N_{ij}. \quad (1)$$

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

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

Equation (2) is simply the “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:

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

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.\(^{16}\) A pair of special interest here is $h_{ii}$,

\(^{16}\)We carefully consider the possibility that common preferences for particular location-specific characteristics drive observed biases in the robustness section, alongside a range of other potential issues. For now, we simply define the null in this manner.
transactions involving buyers and sellers from the same country, as in the examples considered above.

We can then generalize this reasoning to any location $k$ in which transactions occur. We have:

$$h_{ii}^k = \frac{N_{ii}^k}{\sum_{j=1}^{I_k} N_{ij}^k} \quad \text{and} \quad m_j^k = \frac{\sum_{i=1}^{I_k} N_{ij}^k}{\sum_{j=1}^{I_k} \sum_{i=1}^{I_k} N_{ij}^k},$$

which allows us to define the absolute measure of bias for buyers from countries $i$ transacting in locations $k$:

$$\text{Bias}_{k}^{i} = h_{ii}^k - m_i^k,$$

and the testable null hypothesis, averaged across all buyer countries and locations of transactions:

$$H_0 : E[\text{Bias}_{k}^{i}] = 0.$$

### 3.2 Nationality Bias in the Data

The leftmost panel, “full sample” of Figure 3 shows basic results from testing equation (6). In the full sample of transactions, the equal-weighted average across all locations $k$ and countries $i$ of $m_i^k$ is 24.6%, and the equal-weighted average of $h_{ii}^k$ is 26.6%. Using these averages, our estimate of $E[\text{Bias}_{k}^{i}]$ is a statistically significant $26.6\% - 24.6\% = 2\%$.

In the second panel from the left, we estimate $E[\text{Bias}_{k}^{i}]$ only for transactions that occur “at home,” i.e., when $i = k$, and in the rightmost panel, we do so only using transactions occurring “abroad,” i.e., when $i \neq k$. At home, the average market share of sellers belonging to the home country is 78.31%. In turn, the average market share of sellers in all transactions in the home market that involve a buyer from the same nationality is 79.55%. This leads to a relatively modest $1.44\%$ estimate of the bias. However, a far bigger bias is evident when buyers transact in countries that are not their own. On average, the equal weighted average of $m_i^k$ when $i \neq k$ is 5.23%. However, when buyers transact abroad, they match with sellers from the same country at a higher rate. Here, the estimate of $h_{ii}^k$ is 7.51%. The difference
between these two numbers $E[Bias^k_i|i \neq k] = 2.32\%$, which is substantial, since it is almost 50% of the unconditional fraction (i.e., $m^k_i$ when $i \neq k = 5.23\%$).

For comparability with previous research on systematic biases in international investments,\textsuperscript{17} we also consider a relative measure, which slightly modifies equation (6) by increasing the weights in the grand average for nationalities that account for a larger share of the seller pool in each location $k$:

$$\overline{Bias}^k_i = \frac{h^k_{ii} - m^k_i}{1 - m^k_i},$$ \hspace{1cm} (7)

with the associated testable Null hypothesis:

$$H_0: E[\overline{Bias}^k_i] = 0.$$ \hspace{1cm} (8)

As before, we use equations (2) and (3) to compute the sets of conditional ($h^k_{ii}$) and unconditional ($m^k_i$) market shares. We then compute $Bias^k_i$ and $\overline{Bias}^k_i$ across all pairs of locations and nationalities. Table 2 then presents the average $E[Bias^k_i]$ and $E[\overline{Bias}^k_i]$ from this exercise, which average 1% and 3.7% across all locations, respectively.\textsuperscript{18} We can also further separate this result into nationality bias at home ($i = j = k$) and abroad ($i = j \neq k$), and in both sets of locations, the effects are strong and highly statistically significantly different from zero, and similarly sized across locations.\textsuperscript{19}

\textsuperscript{17}Equation (7) 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.

\textsuperscript{18}The numbers in this table differ slightly from those in Figure 3, because we weight locations $k$ by the total number of transactions involving buyers from country $i$ in an attempt to reduce noise. This makes no material difference to the results.

\textsuperscript{19}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 ($h^k_{ii}$) and unconditional ($m^k_i$) market shares, and then compute the bootstrapped bias measures.
We note that nationality bias is strong and robust in subsample analysis, shows up for virtually all the countries in the sample, and in a wide range of location countries. We describe these robustness checks later in the paper, but now turn to understanding how nationality bias affects estimated gravity.

4 Gravity and Counterparties: Reduced-form evidence

We begin with the reduced-form “naïve” gravity equation for trade and investment (see Tinbergen (1962)), which conditions the gross investment flow from a country $i$ to country $k$ on the product of the two countries’ GDP levels, and varies inversely with the distance $D_{ik}$ between them. Letting $N_{ik}^b$ represent the number of transactions involving buyers from country $i$ and properties located in country $k$:

$$\log N_{ik}^b = \beta_0 + \beta_1 \log GDP_i + \beta_2 \log GDP_k + \beta_3 \log D_{ik} + \epsilon_{ik}. \quad (9)$$

The coefficient $\beta_3$ captures the effect of distance on the magnitude of the cross-border capital flow in commercial real estate between countries $i$ and $k$.

Next, let $N_{ik}^s$ denote the number of transactions involving sellers from country $i$ and properties located in country $k$. We add this variable to the above regression to obtain a simple reduced-form estimate of how the density of sellers from the same country in location $k$ affects estimated gravity:

$$\log N_{ik}^b = \beta_0 + \beta_1 \log GDP_i + \beta_2 \log GDP_k + \beta_3 \log D_{ik} + \beta_4 \log N_{ik}^b_{Past} + \beta_5 \log N_{ik}^s + \epsilon_{ij}. \quad (10)$$

Equation (10) 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 $N_{ik}^s$ for each $i$ is the number of sellers present in each location $k$ from the same country as the buyer. As we demonstrate using placebo simulations below, the relationship is not
mechanical, and to a first approximation, $\beta_5 = 0$ is a good null hypothesis.

It is true, however, that a buyer having purchased a property in the past in location $k$ might generate follow-on purchases by the same buyer in the same location in the future, or there may other reasons driving an unobserved propensity for buyers from location $i$ to be attracted to location $k$. To control for this possibility, we therefore split the sample into two equal parts and estimate equation (10) in a post-2013 sample, which also allows us to control for past buying patterns ($\log N_{ik}^{b,Past}$) in the pre-2013 period.

In addition to the number of transactions, we also consider variants of equations (9) and (10) which explain dollar cross-border investment volume. In this case, the dependent variable is the log total USD volume $\log V_{ik}^b$ invested in country $k$ by buyers that hail from country $i$, and the counterparty effect is captured by the log total USD amounts of proceeds $\log V_{ik}^s$ from property sales in country $k$ by sellers that originate from country $i$.

We also consider variants of the above equations that follow recent best practice in the estimation of gravity models, and include buyer country ($\beta_i$) and location country ($\beta_k$) 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.\(^{20}\)

As before, we repeat the estimation of both the naïve gravity equation:

$$\log N_{ik}^b = \beta_0 + \beta_i + \beta_k + \beta_3 \log D_{ik} + \varepsilon_{ik},$$

as well as the one which includes past buying activity and the current density of same-nationality sellers in each location:

$$\log N_{ik}^b = \beta_0 + \beta_i + \beta_k + \beta_3 \log D_{ik} + \beta_4 \log N_{ik}^{b,Past} + \beta_5 \log N_{ik}^s + \varepsilon_{ij}. \tag{12}$$

\(^{20}\)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.
The leftmost column of Panel A in Table 3 confirms the presence of a very strong negative effect of distance between origin and location countries on cross-border investment flows in the data—the naïve gravity equation estimates a strong role for distance, similar to standard trade and investment settings analyzed in many previous papers.

The second column shows that once we add in the density of same-nationality sellers, the coefficient on distance in the resulting equation becomes statistically indistinguishable from zero, while the presence of same-nationality sellers is strong and statistically significant. Finally, when we control for the persistence of investment flows by buyer countries into location countries by including lagged buyer country flows to these destinations, the current availability of same-nationality sellers remains strong and statistically significant, with the distance coefficient still statistically indistinguishable from zero.

The rightmost columns of Panel A confirm these phenomena when dollar transaction volumes are used instead of the number of transactions. Panel B of the table confirms the same observations when we consider the entire set of bilateral country matches including the zero investment flows in the data between a large number of bilateral pairs. This suggests that there may be a role for the density of same nationality sellers, i.e., potential counterparties, in determining the locations of international investment, i.e., the extensive margin of foreign investment. Panel C shows that these results are robust to including buyer and location country fixed effects, which filter out any average variation in investment flows driven by push or pull factors from and into a given country.

4.1 Gravity Effects: Placebo simulations

As mentioned earlier, it is important that (10) is not picking up a mechanical effect. To verify that this is not the case, and to better understand how distance between origin and destination countries and the role of same-nationality counterparties are separately identified, we run two placebo simulations. In the first of these simulations, 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 $N_{ik}^b$ and $N_{ik}^s$ and the distance $D_{ik}$. We then estimate equations (11) and (12) and obtain a distribution of estimated gravity effects using these simulated samples.

Panel A of Figure 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 both estimated (11) (leftmost panel) and estimated (12) (middle panel). 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 $N_{ik}$ 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%.

We have confirmed in this section that the density of sellers from a given country in a location is highly correlated with the density of buyers from that country in the same location. Moreover, we have shown that the inclusion of the density of sellers of a given nationality in a particular location renders estimated gravity effects between buyer country and location country insignificant. Our next step is to investigate the drivers of the observed nationality bias.

5 Understanding and Explaining Nationality Bias

In this section, we first verify that estimated nationality bias is not a statistical artefact arising from the structure of the data, dig deeper with a number of robustness checks of our initial estimates of nationality bias, and finally, run simple reduced-form regressions
on classes of variables that have been used in the gravity literature to explain the role of distance.

5.1 Placebo Simulations

We first 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 do so by conducting a placebo test that imposes the null hypothesis $E[\text{Bias}_i^k] = 0$, by reconstructing the sample in each of $n = 1,000$ simulation rounds.\(^{21}\) 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 $\tilde{h}_{ii}^k = \frac{\tilde{N}_{ki}^k}{\sum_j \tilde{N}_{ij}^k}$ based on the resulting counterfactually matched transactions $\tilde{N}_{ij}^k$. Since the re-sorting is implemented within each location $k$, unconditional market shares $m_i^k$ are unaffected, and we estimate $\text{Bias}_i^k = \tilde{h}_{ii}^k - m_i^k$ when the null is imposed for each nationality $i$ and location $k$.\(^{22}\) We relegate this figure to the online appendix, but highlight here 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.

\(^{21}\)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.

\(^{22}\)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.
5.2 Base Effects

We also note that our estimates of nationality bias may be affected by the fact that seller fractions are calculated using a common base for each nationality and within each location. The decision of investors from a given country $i$ therefore affect the transaction possibilities of investors from all other countries, and nationality bias can be mistakenly attributed to multiple countries. We note that this phenomenon likely also affects estimates of gravity equations in cross-border capital flows, as well as standard estimates of home bias. 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.

To check for bias in the point estimates arising from this source, 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. We relegate these figures to the online appendix, but note here that the results 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.

5.3 Do Nationalities Match to Underlying Characteristics?

An important question when estimating nationality bias is whether seller market shares in the full set of transactions $m_i^k$ are the correct counterfactual distribution of seller nationalities for buyers from country $i$. One possible objection to the use of this benchmark is whether deviations from it could be driven by unobserved factors that are correlated with seller nationalities. This is a similar concern to those faced by previous analyses of bias in international portfolio allocations.
5.3.1 Spatial Clustering

For example, assortative matching could drive the observed result. For example, it might well be the case that Chinese investors have a preference for properties in a given location, or those with particular characteristics located in particular cities. If this were the case, their purchasing decisions may actually be unrelated to the nationality of the seller, but rather, simply clustered around specific areas or property types. This geographic clustering would lead naturally to more frequent transactions between Chinese investors, since they will have a higher ownership share in the locations that they prefer, but it might not have anything to do with a preference for transacting with other Chinese investors.\(^{23}\) In this sense, then, the availability of sellers of the same nationality would be a better measure of the “distance” between buyers and specific locations or characteristics — thus raising the possibility that it’s just a better proxy than physical distance for gravity effects.

In our main results, our approach is to calculate benchmarks \(m_{ik}\) at a very granular scale, i.e., locations \(k\) are “small” sub-markets within a city, such as districts or boroughs.\(^{24}\) In the online appendix, 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 \(k\) to be “large,” i.e., countries, these estimated distances to the central transaction are indeed statistically significant for

\(^{23}\)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.

\(^{24}\)As mentioned earlier, we consider locations such as the “West End” borough (London, UK), the “Upper East Side” (New York, USA), the “Quartier Central des Affaires” (Paris, France), “CBD Midtown” (Sydney, Australia), and “Kowloon CBD Core” (Hong Kong) separately, and compute market shares \(m_{ik}\) for each such location \(k\).
some nationalities. However, when these distances are computed to the “central” trans-
action 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 coun-
try that are correlated with their nationality is no longer relevant for our estimates, which
rely on “within” variation inside narrow sub-markets of cities.

5.3.2 Propensity-score matching approach

Matches between sellers and buyers may reflect preferences for property characteristics,
and not just specific locations. To check whether assortative matching to characteristics
drives the observed nationality bias, we first adopt a parametric 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. To do
so, 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:\(^{25}\)

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

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 \( k \), we apply the Logit propensity scores as weights, to compute a
conditional version of \( m_i \):

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

\(^{25}\)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.
which translates into a conditional bias measure:

\[ \text{Bias}^{\text{matched}}_i = h_{ii} - m_i^{\text{matched}}. \]

In the online appendix, we show results from this exercise, as well as the correlation between the propensity score adjusted benchmark and the baseline fractions of same-nationality sellers. Despite the propensity score capturing heterogeneity in preferences across buyer countries, this change in the computation of \( m_i^{\text{matched}} \) results in the bias estimates falling only slightly. For example, 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.

5.3.3 Non-parametric clustering approach

We also use a non-parametric \( K \)-means clustering approach to isolate clusters \( z \) of \( N \) observations within each location \( k \). As above, we consider clustering along alternative dimensions, by location, transaction, and property characteristics: the year during which transaction took place, the type of property, and an indicator of price quintile. We choose \( N = 20 \) to balance estimation precision (larger clusters) against tougher controls (smaller clusters), and calculate \( h^z_{ii} \) and \( m^z_i \) as before, within each cluster \( z \). The online appendix presents these results, which show that even if we zoom in enough to identify nationality bias effects within small clusters of 20 transactions (often located on opposite sides of the same street), the average magnitude of nationality bias is barely affected.
5.4 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.\footnote{Importantly, we note that the effects by segment do \textit{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 $m^k_i$ are recalculated in each case using the distribution of seller nationalities within each location $\times$ subsample that we consider.} Figure 5 shows that nationality bias is detectable even when we zoom into these much smaller segments of the market, constructing unconditional market shares $m^k_i$ in segments defined by specific property and transaction characteristics \textit{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 past decade. For example, when we restrict the sample to the year 2007 (and therefore also calculate unconditional market shares $m^k_i$ 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 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 increased difficulty 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.

Concerning the role of the corporate type, 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. Indeed, since REITs are highly specialized in trading commercial real estate, we regard them as a useful placebo test. We expect REITs to be most cushioned from issues of trust, search costs, contracting frictions, or information asymmetries.

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 US$ 14 MM, in the lowest quintile), and high-stakes transactions (above US$ 65 MM, 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.

5.5 Nationality Bias, Governance, and Development

To better understand the drivers of nationality bias, we compute the bias measure $Bias_{ij}^k$ in each location $k$ for buyers originating in country $i$ when trading with sellers from country $j$, and condition it on a range of variables. The leftmost column of Table 4 reports the estimated magnitude of nationality bias (this is the equally-weighted average of $Bias_{ij}^k$ across all $i$ and $k$) – as illustrated graphically in Figure 3, to provide a point of reference (this is just a regression of $Bias_{ij}^k$ on a dummy variable that indicates when $i = j$).

In the second column of the table, we explore the hypothesis that buyers have a more
general preference to trade with sellers that hail not necessarily from their own country of origin, but from countries that are located in their close proximity. In other words, we check if there are gravity effects in counterparty matching, over and above nationality bias, adding in a measure of distance between countries \( i \) and \( j \) to the right-hand-side.

The data robustly reject this hypothesis, i.e., the matching bias that we discover is strictly confined to same-nationality counterparties. It is therefore less likely to be related to issues of cultural affinity, and seems more likely linked to the structure of the market in which the trades take place.

The rightmost columns of Table 4 explore this possibility further. To quantify the contractual environment of different location countries, we use the composite Jones-Lang-LaSalle (JLL) Real Estate Transparency Index, which measures the availability of transparent real estate market data on price and performance; the quality of market fundamentals; the nature of corporate governance in the underlying location; measures of the quality of the legal system; and the transparency of the real estate transaction process in locations around the world. Higher values of this index indicate what we term greater “opacity” of the destination country. The estimation results show that nationality bias effects are most pronounced in countries with a low level of GDP, with an even greater effect for those low-GDP countries with opaque real estate markets.

In the online appendix, we report additional estimation results where 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.

Together with the lack of evidence on gravity effects in buyer-seller matching rates (i.e., the fact that Germans don’t seem more likely to trade with the French than the Chinese, for example), this evidence points towards the fact that the underlying fundamental market
friction that drives our results is tightly linked to the structural and legal environment in
the destination market, which leads investors to rely on pre-existing networks of business
relationships—consistent with similar evidence of Chaney (2014) on the exporting behaviour
of multinational firms. These results are also consistent with trust-based theories of market
transactions such as Guiso, Sapienza, and Zingales (2008).

In the next section, we analyze the link between the availability of same-nationality seller
counterparties and the emergence of gravity effects using a more structural approach. To do
so, we build a stylized equilibrium model of the market, in which we think of nationality bias
as arising from a contracting friction between counterparties of different nationalities. We
use the model to evaluate the counterfactual gains that can be generated from eliminating
this market friction, and more importantly, to understand the degree to which we can
rationalize the emergence of gravity effects in an equilibrium matching framework.

6 Equilibrium Matching Model

Our model takes a number of features from the competitive search model of Piazzesi et al.
(2017), and adds some new features to customize the model for our needs. First, we intro-
duce a generic market friction into the model which maps to the underlying driver of the
observed nationality bias. Second, we explicitly model heterogeneity in buyer valuations.
We do so to explicitly capture the distortions that are introduced by the friction – which
may impede buyers with a sufficiently high valuation from accepting seller offers for prop-
erties. When evaluating counterfactuals, this explicit modelling of buyer heterogeneity allows
us to better understand the impact of such distortions than the more common approach
in the search literature, which models random shocks to inventory to move matching rates
away from 0 or 1.

In our model, buyers of type $i$ randomly encounter sellers of type $j$, and matching is
driven by the friction which affects different-nationality counterparty matches.\textsuperscript{27} In these encounters, sellers make take-it-or-leave it offers that buyers can either accept or reject.

### 6.1 The Buyer’s Problem

The decision problem of the buyer conditional on receiving a take-it-or-leave-it offer from a seller is:

\[
\max \left\{ (1 - \lambda)V^B - P, \ 0 \right\}. 
\]  

(13)

We assume that the outside option of the buyer is a profit of 0. The parameter $\lambda$ is the market friction, which captures the fact that there is a distortion to the valuation perceived by the buyer, depending on their own type/nationality, and the type/nationality of the seller. As mentioned earlier, this friction can be thought of as representing difficulties in writing or enforcing contracts with different-nationality sellers, or as a valuation distortion arising on account of buyer mistrust of different-nationality sellers.

We assume that buyer valuations are uniformly distributed:

\[
V^B \sim \text{Uniform}(V^B_{\min}, V^B_{\max}).
\]  

(14)

The decision of the buyer depends on the quoted price $P$, which is endogenously determined in equilibrium. Let $f^*$ characterize the optimal decision:

\[
f^* = \begin{cases} 
1, & V^B \geq \frac{P_{ij}}{(1-\lambda)} \\
0, & \text{otherwise}.
\end{cases}
\]  

(15)

\textsuperscript{27}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.
To understand the main mechanisms operating in the model, it is useful to consider the following comparative statics:

\[
\frac{\partial f^*}{\partial \lambda} < 0 \quad \text{and} \quad \frac{\partial f^*}{\partial P} < 0.
\]  

(16)

The first of these derivatives shows that the more intense the friction (i.e., the larger is \(\lambda\)), 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.

### 6.2 The Seller’s Problem

The seller observes the bilateral friction \(\lambda\), but not the buyer’s private valuation \(V^B\). They therefore choose the asking price optimally, given their expectation about the likely probability that the buyer will accept the offer, and their own valuation \(V^S\):

\[
\max_P E[f](P - V^S).
\]  

(17)

The first-order condition for equation (17) implies the optimal pricing decision:

\[
P^* = V^S + E[f] \left( -\frac{\partial E[f]}{\partial P^*} \right)^{-1} > 0.
\]  

(18)

The seller needs to set the price 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 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 (16) shows, the derivative in the final parenthesis in equation (18) is posi-
tively signed. The price therefore depends positively on the seller valuation $V^S_j$ (as a result of profit-maximizing behavior), as well as on the buyer’s expected acceptance rate. In what follows, we assume that we do not know the seller’s valuation, but solve for it to match the data.

6.3 Equilibrium

Equilibrium in this market is defined by a set of acceptance rates $f$ and asking prices $P$ such that:

- The acceptance decision of the buyer $f^*$ is optimal, given the buyer’s valuation $V^B$ and the asking price $P$.
- The quoted asking price $P^*$ is optimal given the seller’s valuation $V^S$ and the expected acceptance rate $E[f]$.
- Sellers form rational expectations about the acceptance probability, conditional on the buyer’s type:

$$f = E[f^*] = \Pr(f^* = 1).$$

Integrating equation (15), we can derive an expression for the acceptance probability as a function of the price:

$$f = \frac{V^B_{\text{max}} - V^B_{\text{min}}}{1 - \lambda} \frac{1}{(V^B_{\text{max}} - V^B_{\text{min}})^2} P.$$  (20)

Substituting equation (20) into (18) delivers an expression for the pricing equation:

$$P = \frac{V^S + V^B_{\text{max}}(1 - \lambda)}{2}.$$  (21)

Finally, substituting equation (21) into (20), we obtain the equilibrium acceptance proba-
bility for a generic meeting between type-$i$ buyers and type-$j$ sellers:

$$f = \frac{V_B^{\text{max}}}{2(V_B^{\text{max}} - V_B^{\text{min}})} \left(1 - \frac{V_S}{(1 - \lambda)V_B^{\text{max}}}\right).$$

Equation (22)

The model described above shows that volume and price 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. What’s more, for different levels of seller valuations, sellers will also adjust their prices. Reductions in valuations lead them to post lower prices, which in turn generate higher probabilities of matching, and therefore higher expected profits. In the online appendix, we show how model quantities respond to variation in the magnitude of the market friction, in particular how in equilibrium the endogenous response of prices ameliorates the slope of the buyer acceptance rate with respect to the friction $\lambda$.

Equilibrium conditions

Equations (21) and (22) summarize the model equilibrium conditions. Note that the equilibrium solution is defined by three sets of variables that will need to be quantified:

- a set of equilibrium values of endogenous variables $\{f\}$ and $\{P\}$,
- a set of exogenous market conditions: $V_B^{\text{min}}$ and $V_B^{\text{max}}$,
- a set of deep parameters: $\{\lambda\}$ and $\{V_S\}$.

7 Structural Estimation of the Model

We now turn to the quantitative implications of the model, discussing how we recover estimates for each of these variables and parameters from the data.

In the version of the model presented above, we have 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 now 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.

### 7.1 Estimated buyer acceptance rate \(f\)

One key equilibrium quantity is the set of equilibrium acceptance probabilities \(f_{ij}^k\), which
we recover from our empirical estimates of nationality bias.

To understand this, we need to introduce simplifying assumptions and additional nota-
tion. First, we assume that the friction \(\lambda_{ij}^k\) depends on the nationality \(i\) of the buyer and
the nationality \(j\) of the seller in the following way:

\[
\lambda_{ij}^k = \begin{cases} 
0, & \text{if } i = j \\
\lambda, & \text{otherwise.}
\end{cases}
\]  

(23)

Second, we assume that seller valuations are homogeneous and equal to \(V^S\), which we
recover as a structural parameter from the data. This modelling choice correspondingly
reduces the space of \(\{f_{ij}^k\}\):

\[
f_{ij}^k = \begin{cases} 
\text{f}_\text{high}, & \text{if } i = j \\
\text{f}_\text{low}, & \text{otherwise.}
\end{cases}
\]  

(24)

Let \(N_{ij}^k\) denote the number of meetings in which the buyer is from country \(i\), the seller
from country \(j\) and the properties are located in location \(k\) (explicitly accounting for the
model assumption that not all meetings lead to a transaction) and let \(N_{ij}^k\) denote the number
of actual realized transactions. For each country pair \((i, j)\), we therefore have:

\[
N_{ii}^k = \text{f}_\text{high}N_{ii}^k, \text{ and } N_{ij}^k = \text{f}_\text{low}N_{ij}^k \text{ for } i \neq j.
\]  

(25)
We can now use equations (2) and (3) to express the empirically estimated nationality bias $Bias_i^k$ in terms of the number of meetings and realized matching rates:

$$Bias_i^k = h_i^k - m_i^k = \frac{N_{ii}^k}{\sum_j N_{ij}^k} - \frac{\sum_i N_{ij}^k}{\sum_j \sum_i N_{ij}^k}$$

$$= \frac{f_{high} N_{ii}^k}{f_{high} N_{ii}^k + f_{low} \sum_{i \neq j} N_{ij}^k} - \frac{f_{high} \sum_{i \neq j} N_{ij}^k + f_{low} \sum_{j \neq i} \sum_{i \neq j} N_{ij}^k}{2 f_{high} N_{ii}^k + f_{low} \sum_{j \neq i} \sum_{i \neq j} N_{ij}^k}$$ (26)

Finally, to back out the implied acceptance probabilities from equation (26), we need to quantify the distribution of the number of meetings $\{N_{ij}^k\}$. Analogous to equations (2) and (3), define the ratio of all meetings in which the seller is from country $j$ to all meetings between buyers and sellers regardless of nationality as:

$$\bar{m}_j^k = \frac{\sum_i N_{ij}^k}{\sum_j \sum_i N_{ij}^k}$$ (27)

and the fraction of all meetings in which the seller is from country $j$ conditional on the buyer being from country $i$ as:

$$\bar{h}_{ij}^k = \frac{N_{ij}^k}{\sum_j N_{ij}^k}$$ (28)

The number of meetings $N_{ij}^k$ can now be calculated under the identifying assumption that $E[\bar{m}_j^k] = E[\bar{h}_{ij}^k], \forall (i,j)$, i.e., by imposing the null hypothesis of no aggregate nationality bias in the rate at which buyers and sellers randomly meet. Since the remaining model parameters uniquely determine $f_{high}$ for $\lambda = 0$, equation (26) then allows for a direct mapping between the estimated level of the bias and the acceptance rate $f_{low}$.

---

28When setting $\lambda = 0$ in equation (22), the acceptance probability is pinned down uniquely by the values of the other model parameters.
7.2 Prices and Buyer Valuations

We next turn to the estimation of \((V^B_{\text{max}}, V^B_{\text{min}})\) and the equilibrium values of \(\{P\}\). To quantify the variation of prices across match types, we estimate the following standard hedonic regression specification:

\[
\ln PSF_q = \alpha + \mu_k + \delta_t + \beta x_i + \gamma \mathbf{1}_{\text{same nationality}} + \varepsilon_q, \tag{29}
\]

where \(PSF_q\) is the realized price per square foot for property \(q\) in period \(t\) and location \(k\), and \(\gamma\) 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 \(\mu_k\) and \(\delta_t\) eliminate the regional and time components of price dynamics, while the property- and transaction-specific control variables \(x_i\) control for other sources of cross-sectional heterogeneity.

Table 6 Panel A reports the estimated \(\gamma\) 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 \(\gamma\) coefficient shows that there is an increase in the price on average, of 7.36%. We use this estimate alongside the other parameters to pin down \(\lambda\) and \(V^s\).

Finally, we use the estimated residuals from equation (29) to estimate a proxy for within-location valuation heterogeneity:

\[
\hat{\sigma} = \sqrt{E[Var_k(\varepsilon^k_{i,t})]} = 0.318. \tag{30}
\]

For identification, we normalize the price in the group of transactions involving buyers and sellers with different nationalities as \(\overline{P} = 1\). The estimated \(\gamma\) then implies the following
patterns of prices across match types:

\[
P = \begin{cases} 
1, & \text{if } i \neq j \\
1 + \gamma & \text{if } i = j
\end{cases}
\] (31)

This normalization also determines the units of measurement for the seller valuation \( V^S \) and the distribution of buyer valuations \( V^B_i \in [V^B_{\min}, V^B_{\max}] \). To calculate the limits of the uniform distribution, we use the estimated standard deviation \( \hat{\sigma} = 0.318 \) of residual price shocks, based on the hedonic regression in equation (29). Assuming that the residual valuation uncertainty is exactly mirrored in the cross-sectional heterogeneity of buyer valuations, we impose \( \text{Var}(V^B_i) = \overline{V_B} \hat{\sigma}^2 \), which allows us to calculate the lower and upper limits of the uniform distribution \( V^B_{\min} = \overline{V_B}(1 - \hat{\sigma}\sqrt{3}) = 0.458 \) and \( V^B_{\max} = \overline{V_B}(1 + \hat{\sigma}\sqrt{3}) = 1.512 \), and we estimate \( \overline{V_B} \) from \( V^B_{\max} = \overline{V_B}(1 + \hat{\sigma}\sqrt{3}) \), and equation (21).

7.3 Estimates of Deep Model Parameters

Given the values for \( \{f\} \), \( \{P\} \), and \( (V^B_{\min}, V^B_{\max}) \) described above, we can use the system of equations (21) and (22) to uniquely pin down \( \lambda \) and \( V^S \). Table 6 Panel B reports the estimated structural parameters. We find that on average, the friction \( \lambda \) amounts to 9.4% of average prices.

7.4 Evaluating Counterfactuals

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 \( (i \neq j) \).

---

29This results is implied by the expression for the variance of the uniform distribution, i.e., \( \sigma^2 = \frac{(V_{\max} - V_{\min})^2}{12} \).
We can do this by assuming that the probability of offer acceptance is always $f^{high}$, including for cross-nationality meetings:

$$\frac{\Delta N_{i \neq j}}{N_{i \neq j}} = \frac{\sum_{i \neq j} (f^{high} - f^{low}) N_{ij}}{\sum_{i \neq j} (f^{low}) N_{ij}}.$$  \hspace{1cm} (32)

We can also estimate counterfactual changes in prices:

$$\Delta P_{i \neq j} = \frac{V^S + V^B_{\max}}{2} - \overline{P}.$$ \hspace{1cm} (33)

Here, $V^S$ 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 (32) 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 6 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 (US$ 660 BN) 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.
7.5 Evaluating Gravity Effects

The structural model also allows us to identify the counterfactual matching rates between transactions involving same-nationality and different-nationality counterparties.

Using the observed distribution of trades across the entire set of locations and buyer nationalities as described in equation (25), we can compute the counterfactual distribution of transactions $N_{ik}$

$$N_{ik} = \sum_j N_{ik}^j,$$  \hspace{1cm} (34)

for any given level of the market friction $\lambda$.

Analogous to the case of the empirically observed distribution of investment flows, we can use these model-implied observations to estimate a standard gravity equation:

$$\log N_{ik} = \beta_1 \log GDP_i + \beta_2 \log GDP_k + \beta_3 D_{ik} + \varepsilon_{ik}. \hspace{1cm} (35)$$

Figure 6 summarizes the results of these estimation exercises. We normalize the magnitude of the estimated $\beta_3$ coefficient by the corresponding level of its empirical counterpart obtained from equation (9).

When the market friction is eliminated, we cannot explain any of the role of distance in the estimated gravity equation in the data. For a level of the market friction that is equal to the structurally estimated value of $\lambda = 0.094$, the model is able to explain 7.5% of the actually observed estimate.

While the overall explanatory power for the distance effect is clearly limited, we still believe it is surprising that the model is able to match gravity patterns that are not in the model directly. The underlying economic mechanism that drives the explanatory power of the model is nationality bias—which has greater force in countries with higher densities of same-nationality counterparties.

Figure 7 shows the observed densities of same-nationality counterparties in the data,
averaged across all nationalities. The figure shows that same nationality counterparties are distributed log-linearly by geographical distance. The combination of this spatial distribution and the nationality bias drive the estimated model-implied gravity effects in global investment flows.\textsuperscript{30}

The way we interpret this result is that the current version of the model can explain the persistence of gravity using nationality bias, but not the origins of gravity. If gravity determined the locations of outbound investment at some point in history, matching frictions generating nationality bias are a force which will drive persistence in the role of distance, and perpetuate observed gravity.

8 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 reduced-form gravity equations help to explain foreign investment flows in this market, but the availability of same nationality counterparties appears to absorb the role of distance in the gravity equation.

\textsuperscript{30}Importantly, the location of same-nationality counterparty availability are even more visible when controlling more aggressively for country-level variation through buyer and location country fixed effects. On one side, this result suggests that in the commercial property market, aggregate capital flows are only weakly related to relative sizes of the economy — notably, this is the case just for the number of transactions, a proxy for the extensive margin of investment, and not for overall volumes, i.e. the intensive margin. On the other hand, it justifies our approach of estimating nationality bias effects within given locations, and effectively filtering out any systematic sources of variation in aggregate capital flows.
Providing further clues to the microfoundations of gravity, we find that nationality bias itself exhibits no role for distance, and is stronger in poorer and weakly-governed locations. These facts render cultural affinities a less likely explanation for the observed performance of gravity, and make it more likely that contracting frictions or trust are the underlying drivers of the phenomena we observe in the data.

To better understand the underlying economic forces at play, we build an equilibrium matching model of the market. We use the model to structurally estimate the size of the underlying friction, which we relate to greater counterparty comfort with same-nationality transactions for reasons of ease of contracting and trust. We find that the estimated 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. We also learn that nationality bias can help to explain the persistent success of gravity models, given an initial role for location in determining outbound investments.

These results are intriguing, and economically important given the high-stakes environment which we study. While we have made a start on providing evidence on the mechanisms that drive the observed phenomenon of nationality bias, further research is needed to validate the precise economic channels underpinning it. For example, the degree that sellers pre-filter their search space, and influence the matching rate towards buyers with the same nationality remains an open question. This could further exacerbate the frictions that we model. In future versions of this paper, we hope to extend our structural framework to account for and better understand such effects.
References


Electronic copy available at: https://ssrn.com/abstract=3141255
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**

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>1%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total floor area (in ft²)</td>
<td>186,631</td>
<td>5,283</td>
<td>51,215</td>
<td>113,845</td>
<td>232,676</td>
<td>1,150,000</td>
</tr>
<tr>
<td>Property price (in 2017 USD)</td>
<td>$39 mil</td>
<td>$1 mil</td>
<td>$10 mil</td>
<td>$18 mil</td>
<td>$38 mil</td>
<td>$337 mil</td>
</tr>
<tr>
<td>Price per square foot (in 2017 USD)</td>
<td>$294.4</td>
<td>$22.2</td>
<td>$93.1</td>
<td>$175.7</td>
<td>$342.2</td>
<td>$1,984.6</td>
</tr>
</tbody>
</table>

**Panel B**

<table>
<thead>
<tr>
<th>Property type</th>
<th>No.</th>
<th>Freq.</th>
<th>Deal type</th>
<th>No.</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office</td>
<td>40,296</td>
<td>32.6%</td>
<td>Single property</td>
<td>66,371</td>
<td>53.7%</td>
</tr>
<tr>
<td>Retail</td>
<td>28,875</td>
<td>23.4%</td>
<td>Portfolio of properties</td>
<td>57,277</td>
<td>46.3%</td>
</tr>
<tr>
<td>Apartment</td>
<td>26,063</td>
<td>21.1%</td>
<td>Buyer objective</td>
<td>109,037</td>
<td>88.2%</td>
</tr>
<tr>
<td>Industrial</td>
<td>23,022</td>
<td>18.6%</td>
<td>Investment</td>
<td>3,467</td>
<td>2.8%</td>
</tr>
<tr>
<td>Hospitality</td>
<td>5,392</td>
<td>4.4%</td>
<td>Occupancy</td>
<td>6,877</td>
<td>5.6%</td>
</tr>
<tr>
<td>Location within metropolitan area</td>
<td>28,274</td>
<td>22.9%</td>
<td>Renovation</td>
<td>4,263</td>
<td>3.4%</td>
</tr>
<tr>
<td>Central Business District (CBD)</td>
<td>95,374</td>
<td>77.1%</td>
<td>Redevelopment</td>
<td>95,374</td>
<td>77.1%</td>
</tr>
</tbody>
</table>

**Panel C**

<table>
<thead>
<tr>
<th>Source of capital</th>
<th>Buyer</th>
<th>Seller</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>Freq.</td>
</tr>
<tr>
<td>Private</td>
<td>52,106</td>
<td>42.1%</td>
</tr>
<tr>
<td>Institutional</td>
<td>40,917</td>
<td>33.1%</td>
</tr>
<tr>
<td>Public</td>
<td>25,055</td>
<td>20.3%</td>
</tr>
<tr>
<td>Others</td>
<td>5,570</td>
<td>4.5%</td>
</tr>
<tr>
<td>Corporate type</td>
<td>Buyer</td>
<td>Seller</td>
</tr>
<tr>
<td></td>
<td>No.</td>
<td>Freq.</td>
</tr>
<tr>
<td>Developer/owner/operator</td>
<td>45,766</td>
<td>37.0%</td>
</tr>
<tr>
<td>Equity fund/investment manager</td>
<td>30,627</td>
<td>24.8%</td>
</tr>
<tr>
<td>REIT</td>
<td>17,957</td>
<td>14.5%</td>
</tr>
<tr>
<td>Others</td>
<td>29,286</td>
<td>23.7%</td>
</tr>
</tbody>
</table>
Table 2
Estimation results

This table reports estimated average nationality bias effects. We compute weighted averages using country-specific weights in each sub-market. The weights are given by the total number of transactions in location $k$ for which the seller is from country $i$. The 'Nationality bias at home' and 'Nationality bias abroad' samples capture the cases $i = country_k$ and $i \neq country_k$, respectively. We report standard errors in parentheses. *, ** and *** denote statistical significance for 10%, 5% and 1% confidence levels, based on two-stage bootstrap standard errors.

<table>
<thead>
<tr>
<th></th>
<th>Absolute measure</th>
<th>Relative measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average effect</td>
<td>0.010***</td>
<td>0.037***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Nationality bias at home</td>
<td>0.007***</td>
<td>0.038***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Nationality bias abroad</td>
<td>0.027***</td>
<td>0.030***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Number of locations</td>
<td>925</td>
<td>925</td>
</tr>
<tr>
<td></td>
<td>925</td>
<td>925</td>
</tr>
<tr>
<td>Number of countries</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Number of transactions</td>
<td>87,679</td>
<td>87,679</td>
</tr>
<tr>
<td></td>
<td>87,679</td>
<td>87,679</td>
</tr>
</tbody>
</table>
Table 3
Estimation of gravity model

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

\[
\log N_{ik}^b = \beta_0 + \beta_1 \log GDP_i + \beta_2 \log GDP_k + \beta_3 \log D_{ik} + \beta_4 \log N_{ik}^s + \varepsilon_{ik},
\]

\[
\log V_{ik}^b = \gamma_0 + \gamma_1 \log GDP_i + \gamma_2 \log GDP_k + \gamma_3 \log D_{ik} + \gamma_4 \log V_{ik}^s + \nu_{ik},
\]

where \(N_{ik}^b\) is the number of transactions where the buyer is from country \(i\) and the properties are located in country \(k\). \(V_{ik}^b\) is the respective total USD transaction volume. \(N_{ik}^s\) is the number of transactions where the seller is from country \(i\) and the properties are located in country \(k\). Once again, \(V_{ik}^s\) is the respective total USD volume. In Panel B, we report a variant of this specification where we consider variable levels as opposed to log terms. This extends the coverage of the bilateral investment matrix to include buyer country \(\times\) location pairs for which the transaction volume is equal to zero. In Panel C we estimate equivalent specifications where we include buyer country and location fixed effects:

\[
\log N_{ik}^b = \beta_0 + \beta_i + \beta_k + \beta_3 \log D_{ik} + \beta_4 \log N_{ik}^s + \varepsilon_{ik},
\]

\[
\log V_{ik}^b = \gamma_0 + \gamma_i + \gamma_k + \gamma_3 \log D_{ik} + \gamma_4 \log V_{ik}^s + \nu_{ik},
\]

In parentheses, we report robust standard errors, clustered at the location country level.

### Panel A

<table>
<thead>
<tr>
<th></th>
<th>Log Number of transactions</th>
<th>Log Volume of transactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Distance</td>
<td>-0.34***</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Same-nationality sellers</td>
<td>0.74***</td>
<td>0.47***</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Same-nationality buyers (Past)</td>
<td>0.37***</td>
<td>0.40***</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>GDP controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>321</td>
<td>321</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.218</td>
<td>0.622</td>
</tr>
</tbody>
</table>

### Panel B

<table>
<thead>
<tr>
<th></th>
<th>Number of transactions</th>
<th>Volume of transactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Distance</td>
<td>-0.17***</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Same-nationality sellers</td>
<td>0.19***</td>
<td>0.13**</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Same-nationality buyers (Past)</td>
<td>0.47</td>
<td>0.47***</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>GDP controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>6889</td>
<td>6889</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.035</td>
<td>0.853</td>
</tr>
</tbody>
</table>
### Table 3
Estimation of gravity model
(continued)

**Panel C**

<table>
<thead>
<tr>
<th></th>
<th>Number of transactions</th>
<th>Volume of transactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Distance</td>
<td>-0.70***</td>
<td>-0.19*</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Same-nationality sellers</td>
<td>0.69***</td>
<td>0.33***</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Same-nationality buyers (Past)</td>
<td>0.49***</td>
<td>0.49***</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Buyer and location country FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>321</td>
<td>321</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.648</td>
<td>0.813</td>
</tr>
</tbody>
</table>

46
Table 4
Understanding nationality bias

The table reports estimated coefficients from the following specification:

$$Bias_{ij}^k = \alpha + \beta_0 1_{i=j} + \beta_1 D_{i,j} + \beta_2 F^k + \beta_4 1_{i=j} F^k + \varepsilon_{i,j}^k,$$

where $Bias_{ij}^k$ is the bias measure between buyers from country $i$ and sellers from country $j$ in location country $k$, $D$ are quartile dummies for the log distance between the countries of the buyer and the seller, and $F^k$ are location-specific factors such as log GDP and the JLL Transparency Indicator. We include the orthogonalized component of the JLL Transparency Indicator, controlling for log GDP in the specific location country. In parentheses, we report standard errors clustered at the country level. *, ** and *** denote statistical significance at the 10%, 5% and 1% confidence levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>0.022***</th>
<th>0.023***</th>
<th>0.016***</th>
<th>0.016***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same-nationality</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Log Buyer-Seller distance</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Same-nationality × Low GDP</td>
<td></td>
<td></td>
<td>0.019***</td>
<td>0.012*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Same-nationality × Medium GDP</td>
<td></td>
<td></td>
<td>0.003</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Same-nationality × Opacity × Low GDP</td>
<td></td>
<td></td>
<td>0.031**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Same-nationality × Opacity × Medium GDP</td>
<td></td>
<td></td>
<td>-0.024</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Same-nationality × Opacity × High GDP</td>
<td></td>
<td></td>
<td>-0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.020)</td>
<td></td>
</tr>
<tr>
<td>Number of obs.</td>
<td>40,305</td>
<td>40,305</td>
<td>40,305</td>
<td>40,305</td>
</tr>
</tbody>
</table>
Table 5
Robustness checks: Controlling for assortative matching

This table reports estimated nationality bias effects. In the first two columns we use propensity score adjusted fractions of seller nationalities. 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. 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. 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 $i = \text{country}_k$ and $i \neq \text{country}_k$, respectively. We report standard errors in parentheses. *, ** and *** denote statistical significance for 10%, 5% and 1% confidence levels.

<table>
<thead>
<tr>
<th></th>
<th>Propensity-score adjusted</th>
<th>Clustering by location</th>
<th>Clustering by location and characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average effect</td>
<td>0.008***</td>
<td>0.009***</td>
<td>0.012***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Nationality bias at home</td>
<td>0.005***</td>
<td>0.008***</td>
<td>0.009***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Nationality bias abroad</td>
<td>0.026***</td>
<td>0.017***</td>
<td>0.039***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Number of locations</td>
<td>925</td>
<td>925</td>
<td>925</td>
</tr>
<tr>
<td>Number of countries</td>
<td>70</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Number of transactions</td>
<td>87,679</td>
<td>87,679</td>
<td>87,679</td>
</tr>
</tbody>
</table>

48
Table 6
Structural estimation of the model

Panel A reports the estimated coefficient $\gamma$ and the estimated average standard deviation of residuals across locations $\sigma$, based on the following hedonic regression specification:

$$\ln PSF_q = \alpha + \mu_k + \delta_t + \beta X_i + \gamma I_{\text{same nationality}} + \varepsilon_q,$$

where $PSF_q$ is the realized price per square feet for property $q$ in period $t$ and location $k$. $\mu_k$ and $\delta_t$ are location and time fixed effects, and $X_i$ 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 $I_{\text{same nationality}}$ 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 $\lambda$, $V_B$, $V_{B\min}$, $V_{B\max}$ and $\bar{V}$, as implied by the structural model. The quantitative results are obtained under the assumptions that $P = 1$ for matches between buyers and sellers with different nationalities.

### Panel A
Hedonic regression

<table>
<thead>
<tr>
<th>Relative price for same-nationality transactions</th>
<th>$\gamma$ : 0.0736*** (0.0088)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated residual price dispersion</td>
<td>$\hat{\sigma}$ : 0.3188</td>
</tr>
<tr>
<td>Hedonic control variables</td>
<td>Yes</td>
</tr>
<tr>
<td>Location fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>123,648</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.6250</td>
</tr>
</tbody>
</table>

### Panel B
Estimated structural parameters

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>$\lambda$ : 0.094</th>
<th>$V_B$ : 1.022</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of market friction</td>
<td>$V_{B\min}$ : 0.474</td>
<td>$V_{B\max}$ : 1.566</td>
</tr>
<tr>
<td>Distribution of buyer valuations</td>
<td>$V_B$ : 1.566</td>
<td>$\bar{V}$ : 0.579</td>
</tr>
<tr>
<td>Average seller valuation</td>
<td>(assuming $\lambda = 0$)</td>
<td></td>
</tr>
<tr>
<td>Number of transactions</td>
<td>0.065</td>
<td></td>
</tr>
<tr>
<td>Average price level</td>
<td>0.074</td>
<td></td>
</tr>
</tbody>
</table>
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
Nationality bias: Preliminary analysis

This figure reports equal-weighted average fractions of sellers nationalities in their home market and in foreign markets (‘abroad’). We first report unconditional averages, taking into consideration all available deals, irrespective of the nationality of the buyer. We then restrict the view on deals where the buyer and the seller have the same nationality, distinguishing between the case when the parties trade in their joint country of origin, and the case in which they trade in a foreign market. The difference between the conditional market share and the unconditional one indicates the strength of the nationality bias phenomenon. The error bars indicate 90% confidence intervals.
Figure 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 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.

### Average nationality bias effect

<table>
<thead>
<tr>
<th>Year-by-year estimation</th>
<th>Buyers objectives</th>
<th>Corporate type</th>
<th>Location</th>
<th>Price category (quintiles)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Investment</td>
<td>Developer</td>
<td>CBD</td>
<td>1st</td>
</tr>
<tr>
<td></td>
<td>Occupancy</td>
<td>REIT</td>
<td>Non-CBD</td>
<td>2nd</td>
</tr>
<tr>
<td></td>
<td>Redevelopment</td>
<td>Equity Fund</td>
<td></td>
<td>3rd</td>
</tr>
<tr>
<td></td>
<td>Renovation</td>
<td>Investment Manager</td>
<td></td>
<td>4th</td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Nationality bias at home

<table>
<thead>
<tr>
<th>Year-by-year estimation</th>
<th>Buyers objectives</th>
<th>Corporate type</th>
<th>Location</th>
<th>Price category (quintiles)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Investment</td>
<td>Developer</td>
<td>CBD</td>
<td>1st</td>
</tr>
<tr>
<td></td>
<td>Occupancy</td>
<td>REIT</td>
<td>Non-CBD</td>
<td>2nd</td>
</tr>
<tr>
<td></td>
<td>Redevelopment</td>
<td>Equity Fund</td>
<td></td>
<td>3rd</td>
</tr>
<tr>
<td></td>
<td>Renovation</td>
<td>Investment Manager</td>
<td></td>
<td>4th</td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Nationality bias abroad

<table>
<thead>
<tr>
<th>Year-by-year estimation</th>
<th>Buyers objectives</th>
<th>Corporate type</th>
<th>Location</th>
<th>Price category (quintiles)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Investment</td>
<td>Developer</td>
<td>CBD</td>
<td>1st</td>
</tr>
<tr>
<td></td>
<td>Occupancy</td>
<td>REIT</td>
<td>Non-CBD</td>
<td>2nd</td>
</tr>
<tr>
<td></td>
<td>Redevelopment</td>
<td>Equity Fund</td>
<td></td>
<td>3rd</td>
</tr>
<tr>
<td></td>
<td>Renovation</td>
<td>Investment Manager</td>
<td></td>
<td>4th</td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 6
Endogenous gravity effects in the model

This figure reports estimated counter-factual gravity effects based on the counter-factual distribution of transactions $N_{ik}$ for a given level of the matching friction ($\lambda$), using the observed distribution of trades $N_{ik}$ and the estimated structural parameters of our benchmark search model. We estimate gravity effects with the following standard specification:

$$\log N_{ik} = \beta_0 + \beta_1 GDP_i + \beta_2 GDP_k + \beta_3 \log D_{ik} + \varepsilon_{ik}.$$ 

We report the percent fraction of the counter-factual gravity effect relative to the total magnitude estimated in the actual data.
Figure 7
Gravity effects in counterparty availability

This figure reports estimated coefficients $\delta$ from the following empirical specification:

$$\log N_{ik} = \beta_0 + \beta_1 \log GDP_i + \beta_2 \log GDP_k + \sum_{q=2}^{10} \delta_q Decile_q(\log D_{ik}) + \varepsilon_{ik},$$

where $N_{ik}$ is the number of transactions involving sellers from country $i$ and properties located in country $k$. The rightmost sub-panels repeat the estimation for the case of $V^S_{ik}$, the corresponding total USD amount. Panel B repeats the estimation including seller country and location country fixed effects. The shaded areas indicate 95% confidence intervals based on standard errors clustered at the location country level.

Panel A
GDP controls

Panel B
Location and seller country fixed effects
Online Appendix for

Gravity, Counterparties and Foreign Investment

Cristian Badarinza and Tarun Ramadorai*

List of Tables

A.1 Country-by-country effects ........................................... 1

List of Figures

A.1 Location of transactions in the data .................................. 2
A.2 Spatial clustering of commercial property transactions .......... 3
A.3 Placebo tests ................................................................ 4
A.4 Adjustment of seller fractions using propensity score matching . . 5
A.5 Illustration of the K-means clustering approach ................. 6
A.6 Nationality bias: Effects across world regions .................... 7
A.7 Illustrating the endogenous response of volumes and prices ...... 8

*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.
Table A.1
Country-by-country effects

This table reports estimated average relative nationality bias effects ($Bias_k^i$) 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 $i = country_k$ and $i \neq country_k$, respectively. We report standard errors in parentheses. *, ** and *** denote statistical significance for 10%, 5% and 1% confidence levels.

<table>
<thead>
<tr>
<th>Country</th>
<th>Aggregate effect</th>
<th>Nationality bias</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aggregate effect</td>
<td>At home</td>
<td>Abroad</td>
</tr>
<tr>
<td>United States</td>
<td>0.015*** (0.004)</td>
<td>0.015** (0.006)</td>
<td>0.014*** (0.005)</td>
</tr>
<tr>
<td>Japan</td>
<td>0.093*** (0.009)</td>
<td>0.098*** (0.016)</td>
<td>0.063*** (0.014)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.045*** (0.007)</td>
<td>0.049*** (0.017)</td>
<td>0.030*** (0.008)</td>
</tr>
<tr>
<td>Australia</td>
<td>0.051*** (0.016)</td>
<td>0.052 (0.034)</td>
<td>0.044*** (0.018)</td>
</tr>
<tr>
<td>Germany</td>
<td>0.031*** (0.007)</td>
<td>0.042** (0.017)</td>
<td>0.008 (0.007)</td>
</tr>
<tr>
<td>France</td>
<td>0.081*** (0.012)</td>
<td>0.113*** (0.027)</td>
<td>0.021 (0.015)</td>
</tr>
<tr>
<td>Canada</td>
<td>0.025* (0.013)</td>
<td>0.022 (0.018)</td>
<td>0.048 (0.031)</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.073*** (0.011)</td>
<td>0.080*** (0.021)</td>
<td>0.042** (0.020)</td>
</tr>
<tr>
<td>China</td>
<td>0.062** (0.030)</td>
<td>0.079 (0.050)</td>
<td>0.023 (0.031)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.122*** (0.020)</td>
<td>0.163*** (0.049)</td>
<td>0.011 (0.012)</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>0.074*** (0.016)</td>
<td>0.079** (0.036)</td>
<td>0.019 (0.019)</td>
</tr>
<tr>
<td>Other</td>
<td>0.079** (0.031)</td>
<td>0.082*** (0.019)</td>
<td>0.090*** (0.010)</td>
</tr>
</tbody>
</table>
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 $\gamma_i$, 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**
Figura A.4
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 ($m_{i}^{\text{matched}}$). 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.5**
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.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 $i = \text{country}_k$ and $i \neq \text{country}_k$, respectively. Error bars indicate statistical significance for a 10% confidence level.
**Figure A.7**

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 $P = 1$ for matches between buyers and sellers with different nationalities, and for the estimated values of the structural parameters.