Learning to Import from Your Peers*

Márta Bisztray
Institute of Economics of the Hungarian Academy of Sciences

Miklós Koren
Central European University,
Institute of Economics of the Hungarian Academy of Sciences and CEPR

Adam Szeidl
Central European University and CEPR

September 21, 2018

Abstract

We use firm-level data from Hungary to estimate knowledge spillovers in importing through fine spatial and managerial networks. By identifying from variation in peers' import experience across source countries, by comparing the spillover from neighboring buildings with a cross-street placebo, and by exploiting plausibly exogenous firm moves, we obtain credible estimates and establish three results. (1) There are significant knowledge spillovers in both spatial and managerial networks. Having a peer which has imported from a particular country more than doubles the probability of starting to import from that country, but the effect quickly decays with distance. (2) Spillovers are heterogeneous: they are stronger when firms or peers are larger or more productive, and exhibit complementarities in firm and peer productivity. (3) The model-implied social multiplier is highly skewed, implying that targeting an import-encouragement policy to firms with many and productive neighbors can make it 26% more effective. These results highlight the benefit of firm clusters in facilitating the diffusion of business practices.

Keywords: imports, peer effects, spatial spillovers, manager networks, social multiplier

JEL codes: F14, R32, D22

*Emails: bisztray.marta@krtk.mta.hu, korenm@ceu.edu, szeidla@ceu.edu. We thank Krisztian Fekete, Gabor Nyeki, Jeno Pal, Balint Szilagyi, and Andras Vereckei for outstanding research assistance, Christian Fons-Rosen, Sergey Lychagin and seminar audiences for comments. Koren and Szeidl thank the financial support of the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) ERC grant agreements number 313164 (Koren) and 283484 (Szeidl), and Szeidl thanks the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme grant agreement number 724501. Bisztray thanks the financial support of the Lendület Grant 'Firms, Strategy and Performance' of the Hungarian Academy of Sciences.

1 Introduction

Imports have large positive effects on firm productivity (Amiti and Konings 2007, Halpern, Koren and Szeidl 2015), yet there is much heterogeneity in similar firms' importing behavior. One explanation for this heterogeneity is the presence of informal trade barriers, when specific knowledge or a trusted partner is needed for a productive import relationship. When informal barriers are active, importing may diffuse from firm to firm through personal and business connections. Mion and Opromolla (2014), Mion, Opromolla and Sforza (2016), Fernandes and Tang (2014) and Kamal and Sundaram (2016) document such diffusion for exports, but at present we have limited evidence on the—equally important—import side of the market.[1,2] Are there knowledge spillovers in importing? If there are, what factors facilitate or limit diffusion? The answers can shed light on the puzzling cross-firm heterogeneity in importing and its productivity benefits; and can guide trade policy to exploit indirect effects.

In this paper we use firm-level data from Hungary to document and analyze knowledge diffusion in importing. In doing so, we make three main contributions. First, we develop a portfolio of empirical designs which rule out many alternative explanations and help advance the identification of trade spillovers in spatial and managerial networks. We address firm heterogeneity by identifying from source country variation, exclude spatial omitted variables by exploiting the precise neighborhood structure, and also use plausibly exogenous firm moves. We consistently find significant spillover effects. Second, we investigate the factors associated with stronger diffusion. We find that knowledge flows are stronger when firms or peers are larger or more productive. Knowledge flows also exhibit complementarities in firm and peer productivity, showing that positive sorting can increase the overall adoption of importing. Third, we demonstrate in a counterfactual analysis how network density and positive sorting combine to shape adoption patterns. We document that the model-implied social multiplier of importing is highly skewed in the number and type of peers, implying that import subsidies targeted at firms in buildings with many productive neighbors are much more effective.

[1] We review the literature on knowledge spillovers in trade in detail below.

[2] From the results on exports one cannot generalize to imports: finding a foreign supplier is probably easier than finding a foreign client, increasing the supply of, but decreasing the demand for, knowledge diffusion.

In Section 2 we present our data. We use a firm-level panel that contains rich information about Hungarian firms during 1993-2003. We combine three data sources: the Hungarian firm register, balance sheet data from the National Tax and Customs Administration, and trade data from the Hungarian Customs Statistics.[3] The firm register contains, for the full universe of Hungarian firms, the precise address of the firm, all owners with their country of origin, and all firm officials with signing rights, as well as changes over time. As a result, we can trace changes in spatial and ownership links and the moves of people. The balance sheet data include additional information on the foreign ownership share and the industry of firms. And the customs data contain annual export and import flows at the HS6 product level for each firm, separately for each destination and source country.

Section 3 presents our first main contribution: the empirical strategy and results on import spillovers. The key identification concern with estimating spillovers is one common to studies of peer effects (Manski 1993): that a firm and its peer's import choices may be correlated for reasons unrelated to learning. For example, firms in a particular industry may make correlated location and import decisions. We address this endogeneity problem using two main research designs exploiting progressively narrower sources of variation, in combination with placebo tests and sample definition choices that rule out several omitted variables.

Our first research design is a linear probability model measuring the effect of peer firms' country-specific experience on a firm's decision about starting to import from the *same country.* We implement this design by including firm-year and country-year fixed effects, effectively exploiting variation within a firm in a given year: we ask if having a peer which has past experience with a given country increases the probability of starting to import from that country, rather than from another country. To increase comparability we only look at four source countries similar in terms of imports: the Czech Republic, Slovakia, Romania, and Russia. And to ensure that all firms are the same distance from the border we only consider firms located in Budapest.

We use this research design to estimate knowledge diffusion in two networks: close spatial neighborhoods and managerial networks. Within spatial neighborhoods we consider three types of

[3] While firm register and balance sheet data cover a longer period, we do not have access to detailed trade data after 2003.

peers: firms in the same building, firms in the two neighboring buildings, and, as a placebo, firms in the two closest cross-street buildings. In managerial networks, we define peers as firms from which an official with signing rights has moved to the firm of interest. To limit confounding effects we always exclude ownership-connected firms—defined as those which share an ultimate owner with the firm of interest—from the spatial and managerial peer groups.

Our first design yields significant positive import diffusion estimates in both networks. For neighborhood networks we document highly spatially localized spillovers. Having a same-building peer with import experience from a specific country increases the probability of starting to import from the same country by 0.2 percentage points, which roughly doubles the baseline probability of starting to import from one of the four countries. The effect of a neighbor-building peer's import experience is only one-fifth as large, indicating fast decay by distance.[4] The placebo effect of a cross-street peer's import experience is insignificant and small. Finally, in managerial networks the same design yields spillover estimates which are twice as large as the same-building effect.

This design addresses several omitted variable problems which often plague estimates of knowledge diffusion. Most directly, by exploiting variation across source countries it addresses the basic concern that importers tend to be connected to other importers. Specifically, in the absence of source country variation the firm-year fixed effects would soak up all the variation in peers' import experience.[5] In addition, our controls and placebo also address more subtle country-specific omitted variables. In particular, by controlling for ownership links we remove omitted variables based on joint ownership. Results below show evidence on diffusion across industries, addressing concerns with same-industry clustering. And, most important, the neighboring building versus cross-street building comparison rules out any remaining omitted variable as long as knowledge spillovers decay faster than the spatial correlation in that variable.

One remaining concern with our first design is that it does not make explicit the source of variation in peer firms' experience, and therefore it may be subject to some unspecified—highly

[4] We also estimate a decay coefficient and find it to be similar to but somewhat higher than existing within-city spillover decay estimates. This confirms the pattern in the literature that knowledge spillovers are spatially concentrated and suggests that building boundaries may be particularly important barriers in our context.

[5] Mion et al. (2016) also use firm-year fixed effects in their study of the *export* experience of managers moving across firms.

spatially concentrated—omitted variable. In our second design we address this problem by exploiting a concrete plausibly exogenous source of variation: firm moves. We conduct an event study of the impact of firms with country-specific import experience moving into an address where no such experience was present earlier. The move is a positive shock to local country-specific knowledge. We show that firms located in such an address start to import from the country known by the mover with a higher probability than from other countries, relative to firms in addresses where the mover had no such experience. Consistent with the logic of diffusion, the response of imports to the move is gradual. The magnitude of the estimate is comparable to that of our first research design. The consistency of the results identified in different networks and from increasingly narrow sources of variation further supports the knowledge spillovers interpretation.

In Section 4 we present our second main contribution: the heterogeneity of the spillover effect. We explore heterogeneity both to internally validate our estimates and to obtain lessons about mechanisms. We measure heterogeneous effects both by the characteristics of the firm and those of the peer, as well as their interactions. Focusing on same-building peers, we find that larger, more productive and foreign-owned firms benefit more from peers' import experience. Firms also learn more from peers which are larger, more productive or foreign-owned. And spillovers are also stronger when more peers have import knowledge. These results are all consistent with the knowledge diffusion interpretation: better firms are likely to be both more receptive to information and more effective in passing it on, and multiple sources should further increase the rate of diffusion.[6]

We then document that the strength of the spillover also exhibits *complementarities* between the firm's and the peer's characteristics. We show that high-productivity firms tend to learn even more from higher-productivity peers than low-productivity firms do. Similarly, we show that the effect of peers operating in the same industry or importing the same product category is significantly larger than that of other peers. At the same time, spillovers from peers operating in different industries or importing different product types are still significant. The results on complementarities are potentially relevant because they suggest that positive sorting—even holding fixed the network structure—can generate aggregate gains in importing.

[6] The effects are broadly similar but weaker for neighbor-building and managerial connections.

In Section 5 we present our third main contribution: a counterfactual analysis to assess the policy implications of the estimated import spillover effect. Our results so far imply that spillovers should be stronger when (i) the number, and (ii) the productivity of experienced peers is higher. To quantitatively evaluate the combined impact of these forces, we compute the model-implied social multiplier effect on imports of a firm entering into an import market, which incorporates spillovers over the next five years. We calculate the multiplier using the same-building estimate which accounts for heterogeneity by the productivity of the firm and its peers, and also allows for an increase in spillovers with the number of experienced peers. Because the number and productivity of peers varies across the sample, we obtain a separate multiplier for each firm which has not imported yet from one of the four countries.

The results show substantial skewness in the social multiplier. In particular, we find that the five-year social multiplier is 1.03 for the median firm and 1.13 for the firm in the 90th percentile. Thus, while accounting for spillovers is not important for the typical firm, it is potentially quite important for a substantial share of firms. An implication is that there may be significant gains from targeting trade policies. We confirm this by showing that a targeted import subsidy policy treating firms for which spillover effects are the largest can be 26% more effective than a non-targeted one. Because finding the firms with the highest expected indirect treatment effect only requires public information on firms' balance sheet and address, this targeting is in principle directly implementable. Our result quantifies the benefit of clusters—especially of firms with high productivity—in facilitating the diffusion of good business practices.

Literature. We build on a literature on knowledge spillovers in trade, most of which studies the diffusion of exporting. An important part of the literature explores spatial spillovers. Early work focused on the diffusion of the decision to export, and obtained mixed results.[7] More recent work studies the diffusion of specific knowledge, such as export experience with a particular country or product, and generally finds evidence for spillovers (Koenig 2009, Koenig, Mayneris and Poncet 2010, Poncet and Mayneris 2013, Castillo and Silvente 2011, Ramos and Moral-

[7] For example, Aitken, Hanson and Harrison (1997), Barrios, Görg and Strobl (2003), Bernard and Jensen (2004), Lawless (2009) and Pupato (2010) found negative results, while Clerides, Lach and Tybout (1998), Lovely, Rosenthal and Sharma (2005), Greenaway and Kneller (2008) and Dumont, Merlevede, Piette and Rayp (2010) found positive effects.

Benito 2013, Mayneris and Poncet 2015). Using uniquely rich data on trade partners Kamal and Sundaram (2016) document the diffusion of concrete export partners. And Fernandes and Tang (2014) document export spillovers using for guidance a formal model that allows them to test specific predictions of the learning hypothesis.

All these papers define spatial neighborhoods to be cities or similarly large agglomerations. Our spatial spillover results improve identification by using substantially more precise measures of neighborhoods. When networking benefits decay rapidly in space (Arzaghi and Henderson 2008), spatial networks should be measured at a fine resolution to avoid confounding variation from omitted spatially correlated variables. Our results show that spillovers do decay fast, highlighting the relevance of our precise measures. More broadly, we also contribute to this literature by our focus on imports, our analysis of heterogeneous effects and the implications for targeted trade policies.

Another part of the export spillovers literature studies spillovers through managerial moves. These papers show that having a manager with prior export experience join the firm increases the likelihood that the firm starts to export (Choquette and Meinen 2015, Mion and Opromolla 2014, Mion et al. 2016, Sala and Yalcin 2015, Masso, Rõigas and Vahter 2015). We contribute to this work by focusing on import spillovers; by having a comprehensive study in which we compare spillovers in managerial networks to spillovers in spatial networks; and by our analysis of heterogeneous effects and the implications for targeted policies.

Given this existing work on export spillovers, our main focus in this paper is the more novel and equally important topic of import spillovers. There is almost no work on this topic, the sole exceptions—to our knowledge—being Harasztosi (2011) and Harasztosi (2013), which estimate import spillovers in Hungarian NUTS4 agglomeration units. Our contribution to this work is the use of more precise neighborhood definitions, a variety of empirical designs that limit confounding factors, a more comprehensive analysis of multiple networks, the results on heterogeneous effects, and the policy counterfactual analysis.

Finally, we build on a literature on firm networks and diffusion in networks. Chaney (2014) develops a model in which firms can acquire trading partners through existing contacts; Fafchamps and Quinn (2016) and Cai and Szeidl (2016) show that managerial meetings can facilitate the diffusion of business relevant information; and Banerjee, Chandrasekhar, Duflo and Jackson (2013) explore network-based targeting of microfinance in the presence of knowledge diffusion. Our study documents and analyzes these sort of network effects in the novel and important context of import spillovers.

2 Data

2.1 Data sources

We create our panel of Hungarian importers by combining data from three sources.

Firm registry 1993-2003. Data from the Hungarian Company Register contain basic information for the full universe of Hungarian firms, including the firm's name, tax identifier, and precise address: zip code, city, street, number, floor and door number. These variables have associated start and end dates, allowing us to track firm moves over time. The registry data also contain information about the firm's owners, and officials with signing rights which include directors, board members, the CEO, and some employees. As the employees with signing rights are usually at or near the top of the firm hierarchy, we sometimes—slightly imprecisely—refer to these people as managers. For firm owners the data contain the name and registry number; and for person owners and officials the name, mother's name and home address. These records also have start and end dates. We use the name, mother's name and address to create an anonymous unique identifier for each individual in the data. We use this identifier to track individuals across firms and over time. Our method allows for typos and slight variations in names, such as omitting the middle name.

Balance sheets 1993-2003. We have balance sheet data for all double-bookkeeping Hungarian firms from the National Tax and Customs Administration of Hungary. These data also include the firm's industry at the 2-digit NACE level (Revision 1.1), and the shares of its capital owned by foreign entities, domestic private entities, and the Hungarian state.

International trade 1993-2003. Detailed firm level trade data come from the Hungarian Customs Statistics. These data contain yearly exports and imports by each firm to and from each foreign country at the Harmonized System (HS) 6-digit product category. The reason that our sample period ends in 2003 is that the firm-level trade data are not available for later years.

We use unique firm identifiers to link these three datasets.

2.2 Main sample and variable definitions

Firm sample. We focus on imports from four countries that are comparable in terms of their exports to Hungarian firms: the Czech Republic, Slovakia, Romania and Russia. To avoid variation in distance from the border, we use only firms with headquarters in Budapest, which account for over 20% of all the firms in Hungary. Accordingly, when a firm moves its headquarters out of Budapest, we let it exit from our sample. These exclusions result in our *main firm sample* which contains 211,598 firms and 1,189,402 firm-year observations.

We conduct most of the analysis using our *analysis sample*, a (firm, source country, year) panel derived from our main firm sample. In this three-way panel we only include observations in which a firm in the main sample has not yet imported from the given source country up until the previous year. This sample construction allows us to estimate the probability that a firm *starts* to import from a particular country for the first time. We also make three additional exclusions. (1) We exclude firms for which the headquarters' address is missing, because for them we cannot define spatial networks. (2) We exclude firms which have more than 50 same-building peers to ensure that our results are not driven by large hubs. (3) We start the data in 1994 because separate trade data for the Czech Republic and Slovakia are only available starting 1993 and the analysis sample requires importer status of peers in the previous year.[8,9] After these exclusions, the analysis sample contains 88% of the firms in the main firm sample and has 3,778,517 firm-year-country observations. About 5% of the firms in the main sample import from at least one of the four countries at least in one year during 1993-2003.

Variable definitions. We define the firm to have import experience with a country in a year

[8] For the same reason we cannot include firms in the first year they appear in the data. We also exclude those observations—1550 firm-year pairs—in which the firm has no address data from the previous year and no spatial peers coming from the previous year can be defined.

[9] As we have no data on firms' import history before 1992/1993, we cannot rule out that firms classified as not yet importers already imported from the country before 1992. In a robustness check we use the more recent part of our analysis sample, 1998-2003, in which we can rule out that not-yet-importer firms have a recent unobserved import experience with the country.

if it has imported from that country in that year or in a previous year. This definition captures the idea that the firm has acquired import experience specific to that country by that year. We define experience with exports or with foreign owners in an analogous way. We classify a firm as foreign-owned in a year if it had majority foreign ownership that year.

We classify imported products by their purpose using the Broad Economic Categories (BEC) classification. We create four product categories: Consumer goods (BEC 1, 6), Industrial supplies (BEC 2, 3), Capital goods (BEC 41, 51, 52) and Parts and accessories (BEC 42, 53).

Using the Levinsohn and Petrin (2003) methodology we estimate from the balance sheet data total factor productivity (TFP) for each firm in each year, assuming a Cobb-Douglas revenue production function with capital and labor as factors and materials as an input, allowing coefficients to vary by two-digit industries. We normalize log productivity within each 2-digit industry to have mean zero in our main firm sample. We then assign firms to productivity quartiles in each year t, based on the average of their yearly 2-digit-industry-specific productivity percentile over the years $t-2$, $t-1$ and t. Taking the average over three years reduces noise, and results in a smooth but time-varying productivity index.

2.3 Firm networks

A key ingredient in our analysis is data on peers in firm networks. We work with three classes of peers, defined based on spatial, personal and ownership connections.

Spatial peers. We use a highly localized definition of spatial connections. We create three different spatial peer groups. (i) Same-building peers, defined as firms with the same street address up to building number. (ii) Neighbor-building peers, defined for a firm with building number n as firms in buildings in the same street with numbers $n-2$ and $n+2$.[10] (iii) Cross-street peers, defined as firms in buildings in the same street numbered $n-1$ and $n+1$. From all three peer groups we exclude firms which have an ownership link—as defined below—to the firm of interest in the given year. Because the address data has dates, all these peer groups are year specific.

Person-connected peers. We define a firm B to be a person-connected peer of firm A in year t if

[10] Streets in Budapest have an even and an odd side.

some person X is an official with signing rights of firm A in year t and was an official with signing rights of firm B at some earlier date. We will often focus on person connections that can transmit import experience with some country c, which happens when firm B had import experience with c before person X left that firm.

In all person-connected definitions we exclude people with signing rights who are liquidators—officials assigned to handle liquidation of the company—as well as people who are officials or owners of more than 15 firms in the given year. We also exclude from the set of person-connected peers firms which are likely to have shared decision makers with the firm of interest: those ever connected to the firm through ownership links (as defined below), and those that have the exact same address including floor and door number. But we do include peer firms which are located outside Budapest.

With slight imprecision, we sometimes refer to the person-connected network defined this way as the managerial network.

Ownership-connected peers. We classify firms A and B to be linked by ownership in year t if they have a common ultimate owner. This includes two types of connections: (1) when A and B have a direct or indirect common owner; (2) when one of the firms is a direct or indirect owner of the other. We also include peers located outside Budapest in the ownership-connected peer group of a firm.

2.4 Summary statistics

Table 1 presents descriptive statistics on the firms in our main sample. The first column refers to all firms in all years, the second column to firms in years in which they have already had import experience from one of our four source countries, and the remaining columns to firms with import experience from specific countries.

Comparing between columns 1 and 2 shows that importers are on average older, larger, more likely to be foreign owned, more likely to export, and have higher productivity than the industry average. These patterns are familiar (Bernard, Jensen and Schott 2009). The remaining columns show that importers from the four countries of interest are fairly similar in terms of all the variables in the table, consistent with our intuition that these source countries are roughly similar in terms

Table 1: Descriptive statistics

	All firms	All importers from any of the 4 countries	Czech Republic	Slovakia	Romania	Russia
Number of firms	211,598	10,575	5,807	4,534	3,534	2,005
Age	5.5 (3.8)	8.2 (5.7)	8.1 (5.6)	8.2 (5.8)	8.1 (5.1)	8.4 (6.7)
Number of employees	8 (229)	124 (1,515)	104 (1,258)	118 (1,422)	124 (1,637)	191 (2,073)
Log sales	9.0 (2.1)	12.1 (2.4)	12.2 (2.2)	12.1 (2.3)	11.8 (2.5)	12.1 (2.7)
Export share	0.04 (0.17)	0.14 (0.26)	0.12 (0.24)	0.12 (0.24)	0.15 (0.26)	0.21 (0.31)
Log total factor productivity	0.00 (0.97)	0.07 (1.08)	0.09 (1.04)	0.07 (1.07)	0.04 (1.08)	0.07 (1.22)
Share of foreign-owned	0.13 (0.33)	0.32 (0.47)	0.35 (0.48)	0.30 (0.46)	0.28 (0.45)	0.36 (0.48)
Share of state-owned	0.004 (0.060)	0.018 (0.134)	0.016 (0.127)	0.018 (0.133)	0.016 (0.126)	0.027 (0.162)
Number of distinct addresses	78,453	9,428	5,403	4,617	3,648	2,221

Notes: Sample includes firms with headquarters in Budapest, 1993-2003. We report log total factor productivity as the difference from the 2-digit industry average in Budapest. Standard deviations are in parentheses below the sample averages.

of their associated import barriers.

Table 2 shows the number of firms and importers over time during our sample period. The rapid increase in the number of firms is likely due to the development of the capitalist economy in the 1990s. And the increase in the number of importers is probably a consequence of several factors: more firms, lower formal trade barriers, and a country more deeply embedded in the international economy. The considerable increase in importing shown in the table is a key source of variation for our analysis below.

Table 3 reports the distribution of degree (number of peers) in the different firm networks.

Table 2: Number of firms and importers by year

year	Number of firms					
		importing from				
	total	any of the 4 countries	Czech Republic	Slovakia	Romania	Russia
1993	50,982	1,810	753	758	563	509
1994	63,592	2,702	1,175	1,225	754	675
1995	74,516	3,514	1,642	1,599	956	822
1996	86,702	4,197	2,029	1,905	1,127	937
1997	99,858	4,885	2,489	2,185	1,381	1,025
1998	113,366	5,530	2,916	2,410	1,631	1,137
1999	122,407	6,064	3,304	2,588	1,786	1,231
2000	133,031	6,578	3,683	2,784	2,018	1,292
2001	142,433	6,989	3,948	2,955	2,211	1,338
2002	148,574	7,305	4,207	3,095	2,382	1,365
2003	153,941	7,696	4,506	3,311	2,620	1,386

Notes: Sample includes firms with headquarters in Budapest, 1993-2003. A firm is defined to be importing from a country if it has imported at least once from that country by the given year.

Table 3: Number of peers in various networks

Number of peers	Percent of firms in 2003 with n peers in				
	same building	neighbor building	cross-street building	person network	ownership network
0	22.3	31.0	49.7	86.7	48.4
1	13.3	13.7	12.4	8.2	19.5
2	9.0	8.5	6.9	2.5	10.3
3	7.4	6.2	4.7	1.1	5.9
4	6.1	5.3	3.7	0.5	3.6
5 or more	41.9	35.3	22.6	1.0	12.3
Average number of peers	8.4	5.2	3.3	0.3	4.7

Notes: Same building is the building of the firm (street number denoted n). Neighbor building: buildings in the same street with numbers $n+2$ and $n-2$. Cross-street building: buildings in the same street with numbers $n+1$ and $n-1$. Person network: firms in which a current manager of the firm previously had signing right. Ownership network: firms having a common ultimate owner with the firm.

The average degree—shown in the bottom row—is the highest for the same-building network (8.4) and the lowest for the the person-connected network (0.3). The neighbor-building and cross-street networks are between these two extremes (average degrees of 5.2 and 3.3) and although the latter is more sparse, have a roughly similar degree distribution. In all networks a substantial share of firms are isolated, i.e. have zero neighbors. This heterogeneity in degree across firms is one key reason for our finding below that targeting import subsidy policies can substantially increase their effectiveness.[11]

3 Estimating import spillovers

This section presents our empirical strategy and results on the effect of peers' experience on a firm's import decision. Our main hypothesis is that importing requires source-country specific knowledge, which in turn diffuses in various firm networks. As a result, we predict that firms which—other things equal—have peers with experience importing from a particular country are more likely to start importing from that country.

We divide this section into four parts. We begin by presenting motivating evidence which highlights a key component of the logic for identification: variation in peers' import experience across different source countries. We then present two empirical designs. The first design directly exploits this source country variation, and yields spillover estimates in both spatial and managerial networks as well as placebo estimates that confirm the logic of identification. The second design further improves identification for spillovers in spatial networks by exploiting plausibly exogenous firm moves. In the final part we assess the magnitude of our spillover estimates.

3.1 Motivating evidence

Table 4 shows how we exploit source country variation in peers' import experience. The table reports the probability of a firm starting to import from a particular country in a year, conditional on it starting to import from one of the four countries that year, and conditional on different importing patterns of its peers. The four panels correspond to peers defined by the same-building,

[11] Section O1 of the Online Appendix contains additional descriptive statistics about networks and imports. Figures O1-O3 and Table O1 show that importers are fairly similar across source countries. Table O2 shows that the majority of importers imports only from one of the four countries. Table O3 shows that patterns of experienced peers have wide variation across firms.

Table 4: Share of importers with experienced peers

Share of firms starting to import	Firm has peers with import experience: only from country C	Firm has peers with import experience: from any other country
only from country C	44%	18%
from any other country	48%	78%

(a) peers in same building

Share of firms starting to import	Firm has peers with import experience: only from country C	Firm has peers with import experience: from any other country
only from country C	34%	21%
from any other country	61%	75%

(b) peers in neighbor building

Share of firms starting to import	Firm has peers with import experience: only from country C	Firm has peers with import experience: from any other country
only from country C	45%	18%
from any other country	44%	76%

(c) peers in person network

Share of firms starting to import	Firm has peers with import experience: only from country C	Firm has peers with import experience: from any other country
only from country C	55%	14%
from any other country	41%	82%

(d) peers in ownership network

Notes: Each value in each panel is the share of firms specified by the row of the panel relative to the sample of firms specified by the column of the panel. Weighted average across the four countries, with the number of observations in a country as weights. Sample contains firms starting to import from at least one of the four countries. Percentages in a column do not add up to 100% as we exclude firms which start to import both from country C and from another country.

neighbor-building, person-connected and ownership-connected networks. Within each panel, the top row shows the share of firms which start to import from a country c, while the bottom row shows the share which start to import from a different country.[12] The left column computes this share for firms with peers that have import experience with c but not the other countries; and the right column for firms with peers that have import experience with a different country but not

[12] Percentages do not add to 100 as we exclude firms which start to import from both c and another country.

c. We report the average share when c runs across the four countries, weighted by the number of observations per country.

The table shows that in each network, the share of firms starting to import from country c is always higher when peers have c experience than when peers have non-c experience. This fact suggests that peers' experience influences firms' import decisions and forms the basis for our identification strategy. We now turn to more fully develop this empirical approach and derive statistical inference, explicitly address confounds, conduct placebo analysis and incorporate additional plausibly exogenous variation.

3.2 Research design 1: Peers' country-specific import experience

Our main specification is the following linear hazard regression equation:

$$Y_{ict} = \sum_{n} \beta_n X_{ic,t-1}^{n} + \alpha_{it} + \mu_{ct} + \epsilon_{ict}. \tag{1}$$

Here i indexes firms, c indexes source countries and t indexes years, thus each observation is a (firm, source country, year) triplet. We estimate the regression in our analysis sample, which contains observations where firm i has *not yet imported from country* c before year t. The left-hand-side variable Y_{ict} is an indicator for i importing from country c in year t. Given that the sample excludes prior importers from c, Y_{ict} measures entry into importing from c. On the right-hand side we include indicators for the presence of country-specific import experience in various peer groups n. Specifically, $X_{ic,t-1}^{n}$ is and indicator which equals one if there is at least one firm in firm i's peer group n in year $t-1$ which has import experience from country c at time $t-1$, that is, which imported from c in $t-1$ or earlier.[13] We use lagged peer experience because we expect information diffusion to take time. We consider the five different peer groups (n) defined in Section 2.3 above: (1) firms in the same building; (2) firms in the two neighboring buildings, (3) firms in the two cross-street buildings; (4) person-connected peers; and (5) firms in the same ownership network. Finally, α_{it} denotes firm-year fixed effects, μ_{ct} denotes country-year fixed effects, and ϵ_{ict} represents

[13] Subsequent specifications will reuse the notation for variables Y, X, and coefficient β with slightly different meaning. To minimize the risk of confusion, we will explain the notation of each estimating equation directly after it is introduced.

other sources of variation in importing.

Our main hypothesis is that, due to knowledge spillovers, $\beta_n > 0$ for the spatial and managerial networks. We also expect $\beta_n > 0$ for the ownership network, but in that network the mechanism need not be a spillover: it is also possible that the common owner's knowledge causes firms in the network to import from the same country.

Because they play an important role in identifying our key coefficients, it is useful to discuss the fixed effects in equation (1). The firm-year fixed effects α_{it} control for any omitted variable driving import behavior which is specific to the given firm in the given year. This is a rich set of fixed effects, and the only reason it can be included is because the data have an additional panel dimension: multiple source countries. In particular, estimating equation (1) in the absence of data on source countries, or with a single source country, would not be feasible because the firm-year fixed effects would soak up all the variation in the dependent variable. In this sense the key β_n coefficients are identified from source country variation. An implication is that standard firm controls, such as sales, employment, ownership status, or other balance sheet variables need not be included in the regression, since they are already picked up by the firm-year effects. In turn, the second set of fixed effects μ_{ct} pick up country-year specific variation, for example business cycle fluctuations in a source country that might affect the supply of imports. Due to their presence, we do not need to include country-specific controls such as the exchange rate or GDP of the source country.

Beyond import spillovers, slightly modified versions of equation (1) can also be used to estimate other kinds of spillovers. We will look at cross-activity spillovers where on the right-hand side of the equation we measure peer firms' country-specific experience in a different domain, such as exporting to or having a foreign owner from the country; and (in Appendix A) we will also use a variant to present evidence on export spillovers.

Identification. Since equation (1) is essentially a peer effects regression, the main threats to identification are those highlighted by Manski (1993): endogenous peer groups and correlated omitted variables.[14] Endogenous peer groups might arise because of clustering or because of peer

[14] The reflection problem is less relevant for us because we focus on the effect of peers' *past* import experience on the firm's import decision.

choice. An example in the spatial network is when firms from one industry, or "high-type" firms, tend to both co-locate and make similar import decisions, creating spurious correlation between $X^n_{ic,t-1}$ and ϵ_{ict}. An example in the managerial network is when a firm hires a manager because of her or his import knowledge. And an example of correlated omitted variables is when particular physical locations are better for importing from a country c, perhaps because they are close to c.

Our first research design addresses these concerns in three main ways. (1) Source-country variation. By using this variation we address the basic concern that importers tend to be connected to other importers. As discussed above, if we were to estimate equation (1) ignoring the source of imports, the firm-year fixed effects α_{it} would soak up all the variation. The implication is that remaining threats to identification must be based on country variation: for example, if certain types of firms tend to import from certain countries and co-locate with each other. (2) Sample definition. We use comparable source countries; firms based in Budapest; and we omit ownership-based links from the spatial and managerial networks. Our sample choices mitigate several concerns. Because the source countries are similar, it is less likely that "high-type" firms import from one, while "low-type" firms import from another. Because all firms are in Budapest, omitted variables based on distance from a country are muted. And by removing ownership links we address the concern that correlated decisions may be driven by a common owner. In addition, by focusing on imports we limit the concern of endogenous manager choice as knowledge of importing seems less likely to be a driver of hires than for example knowledge of exporting would be. (3) Placebo spatial peers. Perhaps the most convincing component of our design is that by exploiting the fine spatial structure we can compare same-building and neighbor-building spillovers with a cross-street "placebo spillover". As long as spillovers are more spatially concentrated than the omitted variables—an assumption consistent with the results of Arzaghi and Henderson (2008)—estimating higher β coefficients for the closer spatial peers is evidence for knowledge diffusion.

For the above reasons we feel that the most plausible confounds are accounted for by our current research design. Still, a possible concern is that, because the design does not make explicit the source of variation in peer firms' experience, it may be subject to some remaining—highly spatially concentrated—omitted variable. In the next subsection we address this concern by combining the

Table 5: Effect of peer experience on same-country imports

Dependent variable: starting to import	(1)	(2)	(3)	(4)
Peers with import experience in:				
same building	0.22***			0.22***
	(0.03)			(0.03)
neighbor building	0.04**			0.04**
	(0.02)			(0.02)
cross-street building	0.03			0.03
	(0.02)			(0.02)
person network		0.43***		0.41***
		(0.09)		(0.09)
ownership network			0.53***	0.53***
			(0.05)	(0.05)
Firm-year FE	Yes	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes	Yes
Observations	3,778,517	3,778,517	3,778,517	3,778,517

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Dependent variable is an indicator for the firm starting to import from the country in the given year. Right-hand side variables are indicators for specific types of peers with prior country-specific import experience. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

current design with plausibly exogenous variation in peer firms' experience due to firm moves. Although that approach requires weaker identification assumptions, it can only be used to estimate spillovers in spatial networks. We therefore begin the analysis with the current design to demonstrate that knowledge spillovers about imports are present quite broadly across different types of networks.

Results. Table 5 reports estimates of regression (1). In this and all subsequent tables reporting regression results, coefficients are measured in percentage points. To account for spatial correlation in the error term, in all specifications we cluster standard errors by building.

Column 1 focuses on spatial spillovers. The estimated effect of having a same-building peer with country-specific import experience is a significant 0.22. Intuitively, having a peer with experience importing from a particular country, e.g., Slovakia, increases the probability that the firm starts to import from that country by 0.22 percentage points. For comparison, the baseline probability that a firm starts to import from a specific country is 0.19 percent; thus having a peer with experience importing from a country more than doubles the probability of entering that import market. Column 1 also reports that the estimated effect of having a peer with country-specific import experience in a neighboring building is a significant 0.04. This is a fifth as large as the same-building effect, and shows that while spillovers to neighboring buildings are also present, their intensity declines rapidly with distance. The cross-street spillover effect is an even smaller and insignificant 0.03. This result lends support to our identification strategy: if a spatially correlated omitted variable was driving our estimates, we would expect that variable to also affect firms in buildings across the street. Taken together, these estimates strongly support the presence of spatial spillovers in importing.

Column 2 reports the analogous estimate for the person-connected networks. Having a firm official who had prior experience importing from a country increases the probability of importing by a significant 0.43, or almost half a percentage point. This estimate is twice as large as the same-building spillover effect. The larger magnitude seems intuitive: same-building diffusion is likely to be more limited because interactions between members of different firms are probably less common and less intense. In contrast, for person-connected spillovers, interactions are almost guaranteed since the manager now works for the firm.

Column 3 shows the analogous estimates in the ownership-connected network. Here we estimate an even larger coefficient of 0.53. Importantly, this coefficient cannot be interpreted as a knowledge spillover because it is likely partly driven by a common owner making sequential import decisions for her or his firms. Indeed, the reason we include this specification is to show that controlling for the common ownership channel—which we do by excluding ownership-connected firms from the other networks—is important to convincingly document knowledge spillovers in spatial and managerial networks.

Column 4 shows that combining all three types of networks in the same specification leaves the estimates essentially unchanged, indicating that the different networks represent genuinely different spillovers. We conclude that there are significant import spillovers in both spatial and managerial networks. In Appendix A we show that these results are also robust to a range of specification changes including various subsamples (Table A1, A3), additional controls for the firms' or its peers' country-specific experiences (Table A1) and different measures of connections (Table A2).

3.3 Research design 2: Peer moves

In our second research design we exploit a specific, plausibly exogenous source of variation in peer knowledge, which is created by firm moves. Focusing on the same-building spillover, we explore the effect of a peer with particular import experience moving into the building on a firm's subsequent import decision. This design has power because moves are quite frequent, with more than 25% of the firms in our main sample moving at least once.[15] As it is unlikely that the mover would internalize the effect of its import experience on other firms in the building when it chooses its location, we can plausibly assume that country-specific experience brought by the mover is an exogenous shock for the local firms. Similarly, although the owner of the building might want to attract good firms, it is less plausible that she would want to attract firms with specific import experience.

We estimate the impact of moves using an event study, in which the event is when a firm moves from another address into a building. The sample consists of (i, c, t), that is, (firm, source country, year) observations where firm i is located in a building in some year t which is subsequent to some other firm j moving in the same building. The event is the earliest date at which another firm moves into the building of i. To limit the confounding effects of preexisting neighbors, we restrict the sample to observations in which no incumbent firm in the building had import experience with the country c prior to the event. We do not require that the mover firm j has import experience with the country c.

Using this sample, we estimate the following regression equation:

$$Y_{ict} = \sum_{\tau=1}^{5} \beta_\tau \cdot D_{it}^{\tau} \times X_{ic} + \sum_{\tau=1}^{5} \gamma_\tau \cdot D_{it}^{\tau} + \alpha_{it} + \mu_{ct} + \epsilon_{ict}. \tag{2}$$

[15] We present descriptive statistics on moves in Table A4 of the Appendix.

Here Y_{ict} is an indicator for firm i having imported from country c in some year up to and including t. D_{it}^{τ} is an event-year indicator which equals one if the mover firm came to the building of i exactly τ years before t; and the $\tau = 5$ category also includes those observations in which the move occurred more than 5 years ago. X_{ic} is an indicator for the mover firm having had import experience with country c by the time of the move. As before, α_{it} and μ_{ct} denote firm-year and country-year fixed effects and ϵ_{ict} denotes the error term.

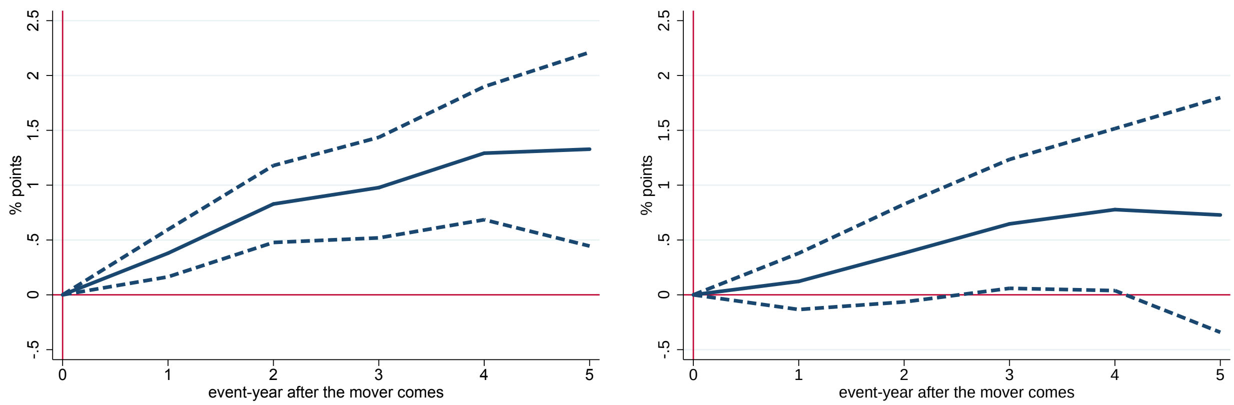
In this specification the coefficients γ_{τ} measure the baseline dynamics of importing from a country c following a move by any firm. The coefficients of interest are the β_{τ} which measure the additional gains in importing when the mover had prior experience with country c. Because of the firm-year fixed effects, similarly to the previous research design this regression is also identified from variation across source countries. Because Y_{ict} indicates if the firm has ever imported from c by t, and because the sample definition ensures that the i has not imported from c before the mover's arrival, Y_{ict} effectively measures if a firm with no prior import experience starts to import from c in the period between the arrival of the mover and t. Thus β_{τ} captures the probability that the firm learns how to import by year τ, even if that firm does not import in every subsequent year.

Figure 1 presents visual evidence from the event study by plotting the estimates of β_{τ} together with their 95% confidence intervals. Panel (a) shows the results from the specification without fixed effects, while Panel (b) from one that includes the full set of fixed effects. Although the point estimates in the second specification are somewhat lower and the standard errors wider because of the large increase in the number of controls, both specifications show the same basic pattern: a gradual and eventually significant increase in the probability of importing from a country subsequent to a new neighbor with country-specific import experience moving in. The fact that the increase is gradual is consistent with the idea of knowledge diffusion.

In the more conservative fixed effect specification of Panel (b), the insignificant first-year effect of 0.12 percentage points increases to a significant 0.78 percentage points after four years.[16] These estimates have the same order of magnitude as the estimated same-building effect of 0.22 percentage points in research design 1, but highlight the importance of explicitly considering the dynamic

[16] We report the full set of coefficients in Table A5 of Appendix A.

Figure 1: Effect of experienced peer moving into building on same-country imports



(a) without fixed effects

(b) with fixed effects

Notes: Sample includes firm-country pairs with the firm located in a building where a new mover arrives, and where no firm has imported from the country before, observed in years after the move. The solid lines show the estimated difference in the number of importers τ years after the move in buildings with movers having country-specific import experience vs with inexperienced movers. τ=5 includes 5 or more years. The dashed lines show the 95% confidence interval, with standard errors clustered by building. Panel (a) shows estimates without fixed effects and Panel (b) shows the same estimates including firm-year and country-year fixed effects in the regression.

response to moves. The pattern revealed here serves as one motivation for examining the dynamic response of further import entries to a new entry in the counterfactual analysis of Section 5 below, where we will also be able to compare explicitly the dynamics implied by research designs 1 and 2.

In summary, our research designs 1 and 2, exploiting different sources of variation, different networks, as well as a placebo design, consistently yield evidence in support of the presence and economic relevance of the knowledge diffusion hypothesis. We conclude that knowledge spillovers in spatial and managerial networks play an important role in shaping firms' import decisions.

3.4 Benchmarking magnitudes

To get a better sense of the magnitude of the spillover effect here we compare it to three sets of benchmarks. As our first benchmark we use export spillovers, the existence of which was documented by Mion and Opromolla (2014), Fernandes and Tang (2014) and Kamal and Sundaram (2016) among others. To make this comparison meaningful, we use the same data and empirical approach for both types of spillovers: we employ our identification strategy 1 to also estimate export spillovers in Hungary. Table A6 in Appendix A presents the results. Both the patterns and the magnitudes are similar to our import spillover results. For example, in the full model including other type of experience as well the same-building effect is 0.16 percentage points, the neighbor-building effect is 0.04 percentage points and the managerial peer effect is 0.37 percentage points. Relative to the baseline hazard of starting to export, 0.21, these estimates correspond to an increase in export probability of 76%, 19% and 176%, while the analogous numbers for the increase in import probability relative to its baseline of 0.19 are 116% , 21% and 216%. Export spillovers, like import spillovers, are also highly concentrated in space. We conclude that diffusion of knowledge about importing is about as strong as diffusion of knowledge about exporting.

As a second benchmark we ask what increase in firm productivity would predict the increase in the probability of importing created by knowledge spillovers. In our sample the probability of starting to import from a country is 0.19% for not-yet-importer firms in the lowest productivity quartile,[17] 0.28% in the second quartile, 0.47% in the third quartile and 0.58% in the highest quartile. Consequently, the estimated same-building import spillover effect of 0.21 percentage points is comparable to the predicted increase in the probability of starting to import as a firm moves from the second to the third productivity quartile. This result further confirms the economic significance of the estimated import spillover effect.

In our third benchmark we look not at the strength of the spillover but at its speed of decay in space. In particular, we infer a parameter of spatial decay that can be explicitly compared to similar decay parameters in the literature. Our approach is to convert the same-building and neighbor-building estimates of research design 1 to a distance-based metric. We work with the decay function $\beta_{ij} = k \cdot e^{-\delta \cdot dist_{ij}}$, where β_{ij} is the estimated spillover from firm j to i, $dist_{ij}$ is the spatial distance between the two firms, and k and δ are parameters. In the 65% of the sample which we were able to geocode, we find that the average distance between two neighboring buildings is 28.1 meters. Assuming that distance is zero if two firms are in the same building, calibrating δ and k to the specification of column (4) in Table 5, we obtain $\delta = 0.0579$/m. This implies that spillovers decline by 5.6% every meter.[18] This value is somewhat higher than other estimates of

[17]Coincidence with the overall baseline is due to many firms with missing data on productivity.

[18]The formula to calculate it is $1 - e^{-\delta \cdot dist}$ where δ is the estimated decay parameter and $dist$ denotes distance in

within-city spatial decay. Indeed, the estimates of Arzaghi and Henderson (2008) on networking benefits among advertising agencies in Manhattan imply a decay of 0.3 % per meter; those by Rossi-Hansberg, Sarte and Owens III (2010) on housing externalities in Richmond imply a decay of 0.2% per meter; and those by Ahlfeldt, Redding, Sturm and Wolf (2015) on production and residential externalities in Berlin imply decays of 0.4% respectively 1% per meter.[19] The main common feature of these results and ours is that they all represent fairly strong decay: knowledge spillovers appear to be highly spatially concentrated. And the fact that our estimate is the highest suggests that in our context building boundaries are important barriers to diffusion. Our decay parameter estimate may be useful for calibrating urban economics models that feature knowledge diffusion of business practices such as importing.

4 Heterogeneity and Mechanisms

In this section we investigate the heterogeneity of import spillovers by firm and peer characteristics. We focus on same-building spillovers because these were the strongest and most cleanly identified. We first explore heterogeneous effects separately by firm and peer characteristics, and then investigate how the interaction between these characteristics influences the strength of diffusion. This analysis yields lessons about the mechanism of spillovers, highlighting the potential benefits of clusters and targeted policies, which we then quantitatively evaluate in the counterfactual analysis of Section 5.[20]

meters.

[19] To calculate these decay parameters, we use column 3 of Table 4 in Arzaghi and Henderson (2008), the estimate on page 524 in Rossi-Hansberg et al. (2010), and column 1 of Table V in Ahlfeldt et al. (2015).

[20] Table O4 of the Online Appendix presents the relative size and the share of imports by firm groups used for the heterogeneity estimates. In Section O2.1 (Tables O5-O7) of the Online Appendix we include the corresponding heterogeneity results for spillovers in other networks.

4.1 Strength of diffusion by firm and peer characteristics

Firm heterogeneity. We estimate heterogeneous effects by firm characteristics using the following regression, which is a modification of research design 1:

$$Y_{ict} = \sum_{h=1}^{H} \beta_h \cdot X^{sb}_{ic,t-1} \times I^h_{it} + \text{controls}_{ict} + \alpha_{it} + \mu_{ct} + \epsilon_{ict}. \tag{3}$$

Here h indexes firm categories by a characteristic, such as productivity quartiles; and I^h_{it} is an indicator which equals one if firm i in period t is in the particular category h, such as the highest productivity quartile. The variable X^{sb} is an indicator for peers' import experience in the same building. Accordingly, the coefficients β_h measure the effect of experienced same-building peers for firms in category h. For completeness, the controls include the analogous interactions of the category indicators with import experience in the four other networks (neighbor building, cross-street building, managerial and owner network).[21] As usual, α_{it} and μ_{ct} denote firm-year and country-year fixed effects and ϵ_{ict} denotes the error term.

Table 6 reports the results from estimating heterogeneous effects by firm size, productivity and ownership. Column 1 focuses on size measured as employment, and categorizes firms into four groups. Group 1 includes those firms with at most 5 employees, group 2 those with 6-20 employees, group 3 those with 21-100 employees and group 4 includes firms with more than 100 employees.[22] The coefficient of 0.07 percentage points shows significant spillover effects for the smallest firms in group 1. The subsequent coefficients imply that the spillover effects for larger firms are larger than those for firms in group 1, and are increasing in the firm's size category. T-tests show that the difference between the estimated coefficients of subsequent groups is significant at 5% in each case (denoted by $\circ$ in the table). Larger firms are more likely to respond to import knowledge in their building.

Column 2 reports heterogeneous effects by firm productivity quartile, defined using our TFP estimates introduced in Section 2. Here we find no spillovers for the least productive firms in

[21] Omitting these controls has small effects on the reported results.

[22] In all three columns we assign firms for which we lack information about the characteristic to group 1. The estimated patterns are robust to putting these firms into a separate group.

Table 6: Heterogeneity of peer effect by firm characteristics

Dependent variable: starting to import	Firm groups by		
	size	productivity	ownership
	(1)	(2)	(3)
Same-building importer peer * Group 1	0.07***	0.03	0.11***
	(0.02)	(0.02)	(0.02)
Same-building importer peer * Group 2	0.62***°	0.20***°	0.81***°
	(0.12)	(0.05)	(0.11)
Same-building importer peer * Group 3	1.45***°	0.38***	
	(0.29)	(0.07)	
Same-building importer peer * Group 4	3.32***°	0.61***°	
	(0.87)	(0.09)	
Other types of importing peers * Group indicators	Yes	Yes	Yes
Firm-year FE	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes
Observations	3,778,517	3,778,517	3,778,517

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Dependent variable is an indicator for the firm starting to import from the country in the given year. Right-hand side variables are indicators for peers with prior country-specific import experience interacted with group indicators. Groups are defined in columns, with group 1 the lowest category or domestic firms in column 3. Other types of peers refer to all other (non-same-building) peer categories in Table 3. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. ° denotes that the coefficient is significantly different from that of previous group at 5%.

group 1, but significant and increasingly strong spillovers in the higher productivity quartiles. The coefficients of subsequent groups are significantly different in two of the three cases. Finally, in column 3 we look at ownership: group 1 represents domestically-owned firms and firms without information on ownership, while group 2 represents foreign-owned firms. Spillovers are significant in both groups and significantly larger for foreign firms. Taken together, these results suggest that absorptive capacity (Lychagin 2016), which is more likely to be present in larger, more productive, and foreign firms, is important for the adoption of import knowledge.

Table 7: Heterogeneity of peer effect by peer characteristics

Dependent variable: starting to import	Peer groups by		
	size	productivity	ownership
	(1)	(2)	(3)
Peers with import experience in:			
same building and in group 1	0.17***	0.14***	0.14***
	(0.03)	(0.04)	(0.03)
same building and in group 2	0.26***	0.13***	0.40***°
	(0.05)	(0.05)	(0.05)
same building and in group 3	0.35***	0.19***	
	(0.07)	(0.04)	
same building and in group 4	0.15	0.34***°	
	(0.10)	(0.05)	
Other types of importing peers in different groups	Yes	Yes	Yes
Firm-year FE	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes
Observations	3,778,517	3,778,517	3,778,517

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Dependent variable is an indicator for the firm starting to import from the country in the given year. Right-hand side variables are indicators for peers with prior country-specific import experience by peer group. Groups are defined in columns, with group 1 the lowest category or domestic firms in column 3. Other types of peers refer to all other (non-same-building) peer categories in Table 3. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. ° denotes that the coefficient is significantly different from that of previous group at 5%.

Peer heterogeneity. To estimate heterogeneous effects by peer type, we use

$$Y_{ict} = \sum_{h=1}^{H} \beta_h \cdot X^{sb}_{ic,t-1}(h) + \text{controls}_{ict} + \alpha_{it} + \mu_{ct} + \epsilon_{ict}. \tag{4}$$

Here too we create categories for a characteristic, such as size, and $X^{sb}_{ic,t-1}(h)$ is an indicator for having a same-building peer in category h which has import experience. Thus β_h measures the effect of having an experienced peer in category h. Similar to equation (3) the controls include the analogous variables for the other networks.

Table 7 reports the results. Column 1 shows spillovers by peer size, using the same cutoffs of 5, 20 and 100 employees already used above.[23] Spillovers are significant even from peers in the smallest group. Although the differences are not significant at 5%, the point estimates show that spillovers are larger when peers are larger, except for peers in the highest quartile where the coefficient is imprecisely estimated. Column 2 shows the analogous specification using peers' productivity quartiles. Here too, spillovers are always positive, and point estimates are larger for higher productivity peers. The difference between the third and fourth quartile is significant. Finally, column 3 shows significant spillovers from domestic peers (group 1) and significantly larger spillovers from foreign peers (group 2). Although the coefficients in this table are slightly less precisely estimated, their general pattern strongly suggests that the import knowledge of larger, more productive and foreign firms—perhaps because they are more successful importers or more trusted peers—is more likely to diffuse. To further confirm this logic, in Table A7 in the Appendix we show that spillovers are stronger from "more successful" importer peers, where import success is measured with the persistence of the peer's import experience.

Number of peers. We next explore whether having more peers with country-specific import experience increases the probability of importing. Simple models of diffusion would predict such an effect, as with more informed peers there are more opportunities for learning. We consider a specification in which the effect is linear and use the number of peers with country-specific experience as a right-hand side variable. Column 1 of Table 8 shows that increasing the number of experienced peers in the same building by one increases the average probability of import entry by 0.2 percentage points. Column 2 presents similar results from a more flexible specification in which we separately estimate the effect of having exactly k experienced peers in a specific peer group. These coefficients are comparable in magnitude to the 0.2 effect of the linear specification, and given the standard errors we cannot reject that in this range the number of experienced peers linearly increases the probability of importing.

Taken together, the above results reveal plausible heterogeneity in knowledge spillovers: diffusion is stronger when firms are better, when peers are better, when the quality of knowledge is higher, and when there are more learning opportunities.

[23] As before, we assign peers for which we lack information on the specific characteristic to peer group 1.

Table 8: Effect of peer import experience by number of peers

Dependent variable: starting to import	(1)	(2)
Peers with import experience in same building:		
Number of peers	0.20*** (0.03)	
1 peer		0.17*** (0.03)
2 peers		0.36***° (0.06)
3 peers		0.82***° (0.14)
4 or more peers		1.02*** (0.21)
Number of peers of other types	Yes	No
Indicators for the number of peers of other types	No	Yes
Firm-year FE	Yes	Yes
Country-year FE	Yes	Yes
Observations	3,778,517	3,778,517

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Dependent variable is an indicator for the firm starting to import from the country in the given year. Right-hand side variables are the number of peers with prior country-specific import experience in column (1) and indicators for a specific number of such peers in column (2). Other types of peers refer to all other (non-same-building) peer categories in Table 3. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. ° denotes that the coefficient is significantly different from that of previous group at 5%.

4.2 Interaction between firm and peer characteristics

We turn to explore how firm and peer characteristics interact in shaping diffusion. Interaction effects are potentially important because their presence indicates that sorting firms can further increase the adoption of good business practices.

Productivity complementarities. We explore complementarities between firm and peer productivity using the following specification

$$Y_{ict} = \beta_{ll} \cdot X^{sb}_{ic,t-1} + \beta_{hl} \cdot X^{sb}_{ic,t-1} \times I^h_{it} + \beta_{lh} \cdot X^{sb}_{ic,t-1}(h) + \beta_{hh} \cdot X^{sb}_{ic,t-1}(h) \times I^h_{it} + \text{controls}_{ict} + \alpha_{it} + \mu_{ct} + \epsilon_{ict}. \tag{5}$$

For simplicity we just use binary indicators to proxy for productivity, let h stand for high-productivity and l for low-productivity firms, and let I^h_{it} be an indicator for firm i in year t being in the high-productivity category.[24] Since $X^{sb}_{ic,t-1}$ is an indicator for import experience by (any) peer while $X^{sb}_{ic,t-1}(h)$ is an indicator for import experience by a high-productivity peer, β_{ll} measures the spillover to a low-productivity firm from a low-productivity peer, β_{hl} and β_{lh} capture the additional gains in the spillover for high-productivity firms and peers, respectively. And β_{hh} measures the complementarity effect of interest.

Table 9 shows the results from estimating this regression. Column 1 reports a specification in which high productivity is defined as the top quartile in the productivity distribution. The fact that the coefficients of the non-interacted indicators of high-productivity firm (β_{hl}) and high-productivity peer (β_{lh}) are positive and significant is familiar from the previous subsection. The key novelty in the specification is that the coefficient of the interaction between high-productivity firm and high-productivity peer is a significant 0.5 percentage points. In column 2 we change the definition of the indicator for high-productivity firm to be above the median of the productivity distribution. The patterns obtained here are similar, and in particular the coefficient of the interaction continues to be significant and positive. From these results we conclude that there are statistically and economically significant complementarities between firm and peer productivity for the adoption of good business practices.

One implication of these results concerns the benefits of sorting. Because of positive complementarities, sorting firms by productivity can generate aggregate gains in the overall adoption of good business practices. This force is distinct from the basic idea that having more informed peers increases adoption: it suggests that even holding fixed the average number of informed peers—that is, the neighborhood structure—changing the pattern of sorting can further increase adoption.

We next present a specification that captures the distinct effects of (i) the number of informed

[24] We assign firms with missing productivity data to the low-productivity group.

Table 9: Complementarities between peer and receiver firm productivity

Dependent variable: starting to import	High-productivity defined as	
	top quartile	above median
	(1)	(2)
Peers with import experience in same building if:		
any peer	0.10***	0.04
	(0.03)	(0.03)
any peer * High-productivity firm	0.29***	0.17**
	(0.10)	(0.08)
high-productivity peer	0.13***	0.05
	(0.05)	(0.04)
high-productivity peer * High-productivity firm	0.50***	0.40***
	(0.18)	(0.11)
Other types of importing peers by firm and peer productivity	Yes	Yes
Firm-year FE	Yes	Yes
Country-year FE	Yes	Yes
Observations	3,778,517	3,778,517

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Dependent variable is an indicator for the firm starting to import from the country in the given year. Right-hand side variables are indicators for peers with prior country-specific import experience separately for high-productivity peers and also interacted with high-productivity firm indicator. We define high-productivity as a 3-year average TFP above the 75th percentile of the 2-digit industry in column (1) and above the 50th percentile in column (2). Other types of peers refer to all other (non-same-building) peer categories in Table 3. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

peers and (ii) the quality of the match. This specification will form the basis of our counterfactual analysis in Section 5 in which we quantify the joint implications of these two forces. Specifically, we estimate

$$
\begin{aligned}
Y_{ict} = {} & \beta_{ll} \cdot N^{sb}_{ic,t-1}(l) \times I^{l}_{it} + \beta_{hl} \cdot N^{sb}_{ic,t-1}(l) \times I^{h}_{it} + \\
& + \beta_{lh} \cdot N^{sb}_{ic,t-1}(h) \times I^{l}_{it} + \beta_{hh} \cdot N^{sb}_{ic,t-1}(h) \times I^{h}_{it} + \text{controls}_{ict} + \alpha_{it} + \mu_{ct} + \epsilon_{ict} \quad (6)
\end{aligned}
$$

where $N(l)$ is the number of low-productivity and $N(h)$ is the number of high-productivity peers, I^{l}_{it} is an indicator for firm i in year t being in the low-productivity category, and controls include

Table 10: Complementarities between peer and receiver firm productivity with peer effect increasing in the number of peers

Dependent variable: starting to import	(1)
Number of peers with import experience in same building if:	
low-productivity peer * Low-productivity firm	0.12** (0.05)
high-productivity peer * Low-productivity firm	0.16*** (0.04)
low-productivity peer * High-productivity firm	0.31***° (0.07)
high-productivity peer * High-productivity firm	0.73***° (0.13)
Number of peers of other types by firm and peer productivity	Yes
Firm-year FE	Yes
Country-year FE	Yes
Observations	3,778,517

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Dependent variable is an indicator for the firm starting to import from the country in the given year. Right-hand side variables are the number of peers with prior country-specific import experience by productivity and interacted with receiver firm productivity indicators. We define high-productivity as having a 3-year average TFP above the 75th percentile of the 2-digit industry. Other types of peers refer to all other (non-same-building) peer categories in Table 3. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. ° denotes that the coefficient is significantly different from that of previous group at 5%.

the analogous interactions for the other four peer groups.

Table 10 shows the results, using the top productivity quartile for the definition of high-productivity firms. The positive and significant coefficients show that spillovers are positive for any firm type and peer type, so that having more knowledgeable peers increases the probability of importing. And the fact that β_{hh} is much larger than the other coefficients shows the complementarity effect: diffusion is stronger when both the firm and the peers are more productive.

Same-industry and same-product effects. To further explore the nature of complementarities, we investigate whether spillovers are larger between same-industry firms, and within a given imported product category. For same-industry effects our strategy is to include separate indicators for experienced peers operating in the same 2-digit industry as the observed firm and operating in different industries. We do this for all networks, but only report the results here for the same-building network. Column 1 of Table 11 shows that same-building peers have a larger effect if they operate in the same industry as the firm. Relative to the significant different-industry spillover of 0.17 percentage points, the same-industry spillover is larger by 0.42 percentage points. Column 2 shows a similar pattern for the restricted sample of manufacturing firms, but perhaps due to the reduction in power the difference between the effect of the two peer types is not significant any more. The positive and significant cross-industry spillovers mitigate identification concerns related to clustering by industry. And the larger same-industry spillovers highlight the societal benefit of sorting firms based on industry for increasing the overall adoption rate of good business practices.

Finally, to measure import diffusion within a product category, we modify our specification in two ways. First, we estimate separate regressions for each product category, using a sample of observations in which the firm has not yet imported the given product category from a specific country, and including as controls indicators for whether the firm has imported other product categories from that country before. Second, our right-hand side variables are indicators for "same-product importer peers"—that is, peers which have imported in the past the given product category from the specific country—and "different-product importer peers"—that is, peers which have only imported in the past different product categories from the specific country. The last four columns of Table 11 show the results for each of four product categories defined based on the BEC categories. The effect of different-product importer peers is significant in all four categories; and same-product spillovers are always higher, significantly so in three of the four cases. We conclude that spillovers are larger within a product category, which is intuitive if part of importing knowledge is product-specific and further strengthens the argument about sorting firms based on industry to maximize spillovers.[25]

[25] In Table O8 of the Online Appendix we also explore a related specification in which we show that conditional on a firm starting to import from a country, it is more likely to import the product category in which its peer has had import experience.

Table 11: Effect of peer experience within industry and product

Dependent variable: starting to import	same industry		same product			
	All firms	Manuf. firms	Consumer goods	Industrial supplies	Capital goods	Parts
	(1)	(2)	(3)	(4)	(5)	(6)
Peers with import experience in:						
same building with different industry/product	0.17***	0.36***	0.07***	0.05**	0.06***	0.05***
	(0.02)	(0.12)	(0.02)	(0.02)	(0.01)	(0.01)
same building with same industry/product	0.59***°	1.00**	0.17***°	0.17***°	0.11***	0.18***°
	(0.09)	(0.44)	(0.03)	(0.03)	(0.03)	(0.03)
Other types of importing peers by same/different industry/product	Yes	Yes	Yes	Yes	Yes	Yes
Not yet importer from country	No	No	Yes	Yes	Yes	Yes
Firm-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,778,517	376,739	3,821,755	3,805,958	3,828,759	3,829,629
Baseline hazard (in %):	0.19	0.41	0.07	0.11	0.05	0.05

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Column (2) contains only manufacturing firms. Dependent variable is an indicator for the firm starting to import from the country in the given year. In columns (3)-(6) only imports in the given product category are considered, both for creating the sample and defining the dependent variable. Right-hand side variables are indicators for peers with prior country-specific import experience. Separate indicators are included for peers in the 2-digit industry of the firm or in a different industry in columns (1)-(2), and peers importing the same or different product categories in columns (3)-(6). Consumer goods are BEC 1 & 6, industrial supplies are BEC 2 & 3, capital goods are BEC 41, 51 & 52, and parts and accessories are BEC 42 & 53. Other types of peers refer to all other (non-same-building) peer categories in Table 3. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. ° denotes that the coefficient is significantly different from that of previous group at 5%. Baseline hazard refers to the share of importers in the estimation sample.

5 Counterfactual policy analysis

In the presence of spillovers, policies that encourage firm trade can have additional indirect effects through a social multiplier (Glaeser, Sacerdote and Scheinkman 2003). And when spillover effects are context-dependent, so is the size of the multiplier, opening the possibility that targeted trade policies generate larger social gains. In this section we use our estimates of the import spillover effect in a counterfactual analysis to explore how the size and composition of a firm's peer group shape the social multiplier.

Our goal is to compute the model-implied effect on the number of importers of a non-importer firm's exogenously induced entry into importing. To do this we assume that import spillovers follow a simple diffusion model whose parameters are determined by our estimates. For simplicity, in the model we only allow import spillovers between peers in the same building. We assume that the probability that a non-importer gets "infected" is linear in the number of importing peers.[26] We allow the diffusion probability to depend on both the sender and the receiver firm's productivity type, measured with an indicator which equals one if the firm is in the highest productivity quartile. We also allow firms to become importers independently of spillovers, with a baseline probability which is constant over time and across source countries, but can depend on the firm's productivity type. We assume that all spillover and baseline adoption realizations are independent from each other and over time. Given these assumptions, the model generates a Markov process, and we can track its dynamics, for each building, with four state variables: the number of high/low productivity importer/non-importer firms in the building.[27]

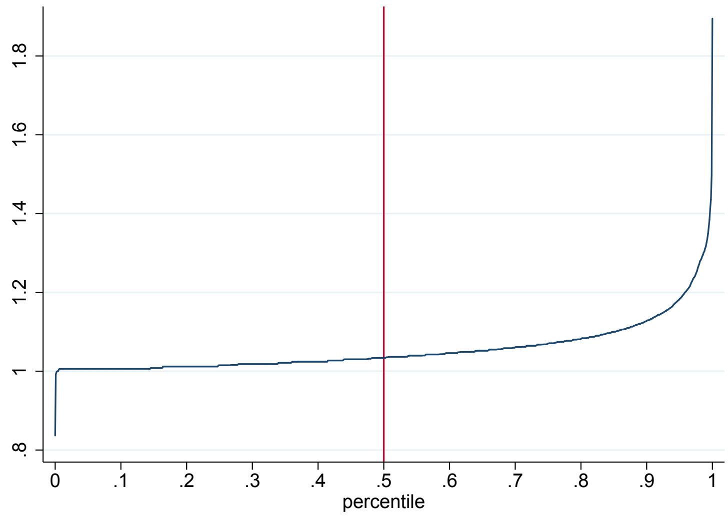
To parametrize this model we use specification (6) in Table 10, which estimates different spillover parameters by firm and peer productivity category, and also reports the change in spillovers by the number of experienced peers. We calculate baseline probabilities by firm productivity category using the subgroup of firms which have no experienced peers in the same building. Starting from an initial year s which we set to 2003, we then study dynamics in the diffusion model in each building over a five-year horizon. In doing this, we assume that firms do not move in or out of the building and do not enter or exit production. We also investigate the benchmark case of the model with no spillovers, in which the diffusion parameters are set to zero.

The social multiplier. A key object of interest is the 5-year social multiplier of importing that results from exogenously inducing a firm i to import from country c, defined for each firm i which has not started to import from country c as

$$\eta_s^c(i) \equiv \frac{E[M^c_{a(i),s+5} \mid T^c_s(i) = 1,\ \text{spillovers}] - E[M^c_{a(i),s+5} \mid T^c_s(i) = 0,\ \text{spillovers}]}{E[M^c_{a(i),s+5} \mid T^c_s(i) = 1,\ \text{no spillovers}] - E[M^c_{a(i),s+5} \mid T^c_s(i) = 0,\ \text{no spillovers}]}. \tag{7}$$

[26] This assumption is consistent with the results of the flexible specification reported in column (2) of Table 8.
[27] We present the transition matrix and the forecasting equations of this Markov chain in Appendix B.

Figure 2: Distribution of the 5-year social multiplier for firms with non-importer peers in the building



Notes: Sample includes firm-country pairs in which the firm and at least one other firm in the building have not yet imported from the country in 2003. The 5-year social multiplier is the additional number of firms in the building starting to import from a specific country within 5 years after one firm in the building is induced to start importing form the country, normalized by the same difference in the absence of spillovers. For the calculations we assume that import spillovers and the baseline probability of starting to import are constant over time and across countries, but heterogeneous across firm and peer productivity groups; spillovers exist only within the same building, and increase linearly in the number of peers; and there are no firm entries, exits or location changes.

Here $M^c_{a(i),s+5}$ is the number of importers from country c on address a of firm i in year $s+5$ and $T^c_s(i)$ refers to the "treatment status" of firm i in year s, taking the value 1 if this firm is induced to start importing from country c. The numerator shows the expected change in the number of importers after 5 years of firm i being treated. This term incorporates import spillovers. The denominator is the corresponding treatment effect in the benchmark model in which import spillovers are set at zero. Thus the multiplier measures how much larger is the treatment effect in the presence, relative to the absence, of import spillovers.[28]

Figure 2 plots, in increasing order, the implied 5-year social multiplier for all non-importer firms that have non-importer peers in our data in $s = 2003$. The figure reveals substantial heterogeneity. Interestingly, for about half a percent of firms the multiplier is smaller than one: treating these

[28] The conditional expectations can be calculated by iterating the Markov chain forward, as we explain in Appendix B.

firms results in a smaller number of total importers in the presence of spillovers than in the absence of spillovers. This is because spillovers have two effects: they increase the impact of treating firm i, but they also increase spillovers from other importers in the building. Because of this second force, the net effect of treating firm i can be reduced when spillovers are introduced, essentially because spillovers from peers of i crowd out spillovers from i. However this subtle crowding-out effect only overcomes the more intuitive positive effect for a small share of observations.

The median multiplier in the figure is 1.03: inducing the median firm to import is 3 percent more effective in the presence than in the absence of import spillovers. The 90th percentile of the multiplier is 1.13. Thus inducing a firm to import which is located at this point of the multiplier distribution is 13 percent more effective once import spillovers are taken into account. While spillovers may not be very important for the typical firm, they seem quite important for a significant share of firms, suggesting that targeting policies to such firms can generate substantial benefits.

Targeted trade policies. We next use our counterfactual to evaluate a hypothetical import-encouraging trade policy, which demonstrates how targeting can improve policy effectiveness. For policy evaluation the object of interest is not the multiplier, but rather the numerator of equation (7), which measures the five-year treatment effect of inducing firm i to import from country c. In our analysis we compare two policies: one in which we target firms for which this treatment effect is large, and another with no targeting. For simplicity we consider an import-encouragement treatment which is completely effective in teaching the firm how to import from a particular country. Thus we assume that treating a firm results in it starting to import from the country under consideration with certainty.

Our targeted policy is to treat the 1,000 firms for whom the estimated treatment effect is largest, while our non-targeted policy is to treat 1,000 randomly chosen firms. To avoid complications arising from treating multiple firms in the same building, we restrict both policies to treat, for any given source country, at most one firm per building. And to induce some amount of diffusion we only treat firms which have not yet imported from the country and which have at least one other non-importer peer in the building. Evaluating the targeted policy is straightforward, as it requires computing the numerator of (7) for the selected firms. For the non-targeted policy the impact also depends on the specific set of firms treated. To measure its average effect, we draw the 1,000 random firms 1,000 times, compute the treatment effect for each draw, and average over draws.

The differences between the impacts of the two policies are remarkable. The targeted policy yields after five years 285 additional importers for a total of 1,285 importers. In contrast, the non-targeted policy yields, on average, 16 additional importers. In this example the targeted policy is 26% more effective than the non-targeted policy ($1,285/1,016 - 1 = 0.26$). Since the targeting is based entirely on observable firm characteristics such as the productivity of the treated firm and its peers in the building, in principle it can be implemented using public data. Overall, the result suggests that there can be large potential gains from targeting interventions to firms which are likely to be good seeds for diffusion.

Internal consistency. We now connect the simulation results of the diffusion model and the estimates of the mover design in Section 3.3. Both of these designs evaluate the dynamic impact of having an additional importer peer. Because they exploit different sources of variation and use a different combination of reduced-form and structural approaches, their comparison provides a useful test of internal consistency. As we have just seen, the counterfactual implies that turning 1,000 random firms in different buildings with non-importer firms into importers would result in an expected 16 additional importers after 5 years. The point estimate of the mover design implies (Table A5) that five years after an importer moves into the building, the probability of an incumbent starting to import increases by 0.73 percentage points. Because the average number of incumbent firms in a building is 4.6, the latter estimate implies that turning 1,000 firms in different "non-importer" buildings into importers would result in $0.0073 \cdot (4.6 - 1) \cdot 1000 = 26.28$ new importers. This has the same order of magnitude as the counterfactual, and given our standard errors we cannot reject that the two are equal.

We can also check intervening years. Table O9 in the Online Appendix reports the expected number of firms starting to import 1-4 years after the above treatment in both designs. Here too, the numbers have the same order of magnitude and given the confidence intervals we cannot reject that they are equal. These patterns are especially remarkable because the mover and the counterfactual design use somewhat different samples: for example, in the mover design the 5-year effect is identified from moves in the subperiod 1994-1998. We conclude that exploiting different designs and sources of variation lead to similar estimates of the dynamics of knowledge spillovers, providing internal consistency to our results.

6 Conclusion

In this paper we documented evidence for import spillovers. Exploiting source-country variation, precise spatial neighborhoods and plausibly exogenous firm moves in two complementary research designs, we obtained credible estimates of diffusion in spatial and managerial networks. We also documented that spillovers are stronger when firms or peers are better, and exhibit complementarities in firm and peer productivity. Taken together, these two results show that both high network density, and positive sorting in a given network, can increase diffusion. We then conducted a counterfactual analysis showing that due to the combination of these two forces the social multiplier of importing is heterogeneous, so that targeted import subsidy policies can have substantially larger effects. In combination, our results highlighted one concrete benefit of firm clusters: that of facilitating the diffusion of good business practices. More broadly, our analysis contributes to a growing literature highlighting the importance of business networks in shaping economic outcomes.

References

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf, "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Aitken, Brian, Gordon H Hanson, and Ann E Harrison, "Spillovers, foreign investment, and export behavior," *Journal of International Economics*, 1997, *43* (1), 103–132.

Amiti, Mary and Jozef Konings, "Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia," *The American Economic Review*, 2007, *97* (5), 1611–1638.

Arzaghi, Mohammad and J Vernon Henderson, "Networking off Madison Avenue," *The Review of Economic Studies*, 2008, *75* (4), 1011–1038.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson, "The diffusion of microfinance," *Science*, 2013, *341* (6144), 1236498.

Barrios, Salvador, Holger Görg, and Eric Strobl, "Explaining firms export behaviour: R&D, spillovers and the destination market," *Oxford Bulletin of Economics and Statistics*, 2003, *65* (4), 475–496.

Bernard, Andrew B and J Bradford Jensen, "Why some firms export," *The Review of Economics and Statistics*, 2004, *86* (2), 561–569.

____ , ____ , and Peter K Schott, "Importers, exporters and multinationals: a portrait of firms in the US that trade goods," in "Producer dynamics: New evidence from micro data," University of Chicago Press, 2009, pp. 513–552.

Cai, Jing and Adam Szeidl, "Interfirm Relationships and Business Performance," Technical Report, National Bureau of Economic Research 2016.

Castillo, Juana and Francisco Requena Silvente, "Export dynamics and information spillovers: evidence from Spanish firms," 2011. Department of Applied Economics II, Universidad de Valencia, Working Papers 1103.

Chaney, Thomas, "The network structure of international trade," *The American Economic Review*, 2014, *104* (11), 3600–3634.

Choquette, Eliane and Philipp Meinen, "Export spillovers: Opening the black box," *The World Economy*, 2015, *38* (12), 1912–1946.

Clerides, Sofronis K, Saul Lach, and James R Tybout, "Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco," *The Quarterly Journal of Economics*, 1998, pp. 903–947.

Dumont, Michel, Bruno Merlevede, Christophe Piette, and Glenn Rayp, "The productivity and export spillovers of the internationalisation behaviour of Belgian firms," 2010. National Bank of Belgium Working Paper, 201.

Fafchamps, Marcel and Simon Quinn, "Networks and Manufacturing Firms in Africa: Results from a Randomized Field Experiment," *The World Bank Economic Review*, 2016.

Fernandes, Ana P and Heiwai Tang, "Learning to export from neighbors," *Journal of International Economics*, 2014, *94* (1), 67–84.

Glaeser, Edward L, Bruce I Sacerdote, and Jose A Scheinkman, "The social multiplier," *Journal of the European Economic Association*, 2003, *1* (2-3), 345–353.

Greenaway, David and Richard Kneller, "Exporting, productivity and agglomeration," *European Economic Review*, 2008, *52* (5), 919–939.

Halpern, László, Miklós Koren, and Adam Szeidl, "Imported inputs and productivity," *The American Economic Review*, 2015, *105* (12), 3660–3703.

Harasztosi, Péter, "Export and import spillovers in Hungary," 2011. Unpublished manuscript, Central European University.

——, "Essays on agglomeration economics and trade." PhD dissertation, Central European University 2013.

Kamal, Fariha and Asha Sundaram, "Buyer–seller relationships in international trade: Do your neighbors matter?," *Journal of International Economics*, 2016, *102*, 128–140.

Koenig, Pamina, "Agglomeration and the export decisions of French firms," *Journal of Urban Economics*, 2009, *66* (3), 186–195.

——, Florian Mayneris, and Sandra Poncet, "Local export spillovers in France," *European Economic Review*, 2010, *54* (4), 622–641.

Lawless, Martina, "Firm export participation: entry, spillovers and tradability," *Applied Economics*, 2009, *41* (5), 665–675.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *The Review of Economic Studies*, 2003, *70* (2), 317–341.

Lovely, Mary E, Stuart S Rosenthal, and Shalini Sharma, "Information, agglomeration, and the headquarters of US exporters," *Regional Science and Urban Economics*, 2005, *35* (2), 167–191.

Lychagin, Sergey, "Spillovers, absorptive capacity and agglomeration," *Journal of Urban Economics*, 2016, *96*, 17–35.

Manski, Charles F, "Identification of endogenous social effects: The reflection problem," *The review of economic studies*, 1993, *60* (3), 531–542.

Masso, Jaan, Kärt Rõigas, and Priit Vahter, "Foreign market experience, learning by hiring and firm export performance," *Review of World Economics*, 2015, *151* (4), 659–686.

Mayneris, Florian and Sandra Poncet, "Chinese firms' entry to export markets: the role of foreign export spillovers," *The World Bank Economic Review*, 2015, *29* (1), 150–179.

Mion, Giordano and Luca David Opromolla, "Managers' mobility, trade performance, and wages," *Journal of International Economics*, 2014, *94* (1), 85–101.

——, ——, and Alessandro Sforza, "The Diffusion of Knowledge via Managers' Mobility," 2016. CEPR Discussion Paper 11706.

Poncet, Sandra and Florian Mayneris, "French Firms Penetrating Asian Markets: Role of Export Spillovers," *Journal of Economic Integration*, 2013, pp. 354–374.

Pupato, Germán, "Non-Market Interactions and Entry into Export Markets," 2010. Available at SSRN 1777327.

Ramos, Roberto and Enrique Moral-Benito, "Agglomeration matters for trade," 2013. Banco de Espana Working Paper.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Raymond Owens III, "Housing externalities," *Journal of Political Economy*, 2010, *118* (3), 485–535.

Sala, Davide and Erdal Yalcin, "Export experience of managers and the internationalisation of firms," *The World Economy*, 2015, *38* (7), 1064–1089.

A Additional Evidence

A.1 Alternative specifications and robustness

Early import history. Since we do not observe the import history of firms before 1992, in our empirical designs such as that in 3.2 we may misclassify some firms which imported earlier than 1992 as non-importers. To address this problem, we re-estimate our baseline specification (1) for the sub-period 1999-2003, which ensures that a firm we classify as a non-importer in a year did not import in the preceding six years. Column (1) of Table A1 shows that this sample restriction gives significant peer effects and qualitatively similar results.

Firm's other country-specific experience. A firm's decision to import from a country may be correlated with its other country-specific experience, such as exporting to that country or having owners from that country. Column (2) of Table A1 re-estimates our baseline specification (1) controlling for these experiences. Both of these experiences predict the decision to import, but our estimate of the spillover coefficient remains essentially unchanged.

Peers' country-specific experience. We next explore whether having peers who have *export* experience with a country (column 3), or who have been *owned* by entities from a country (column 4) also affects a firm's decision to import. If importing from country c requires knowledge specific only to country c, then we expect similar coefficients for these cross-activity spillover effects; but if importing also requires knowledge specific to the activity of importing then these estimates should be smaller. The results seem more consistent with the second hypothesis: although peers' export and ownership experience do predict to some extent the decision to import, the coefficients (not reported) are generally smaller and less significant, and the import spillover coefficients stay

Table A1: Effect of peer experience on same-country imports

Dependent variable: starting to import				
Sample period:	1999-2003	1994-2003		
Type of other experience:			Exporter	Owner
	(1)	(2)	(3)	(4)
Peers with import experience in:				
same building	0.15***	0.21***	0.22***	0.22***
	(0.03)	(0.03)	(0.03)	(0.03)
neighbor building	0.04**	0.04**	0.04*	0.04**
	(0.02)	(0.02)	(0.02)	(0.02)
cross-street building	0.02	0.03	0.03	0.03
	(0.03)	(0.02)	(0.02)	(0.02)
person network	0.37***	0.39***	0.40***	0.40***
	(0.10)	(0.09)	(0.09)	(0.09)
ownership network	0.49***	0.51***	0.51***	0.53***
	(0.05)	(0.05)	(0.05)	(0.05)
Firm exported to the country		2.01***		
		(0.11)		
Firm had owners from the country		0.64***		
		(0.08)		
Peers with other experience	No	No	Yes	Yes
Firm-year FE	Yes	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes	Yes
Observations	2,385,154	3,778,517	3,778,517	3,778,517

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Column (1) includes a shorter sample period: 1999-2003. Dependent variable is an indicator for the firm starting to import from the country in the given year. Right-hand side variables are indicators for specific types of peers with prior country-specific import experience, as well as export experience in column (3) and country-specific owners in column (4). Column (2) includes additional indicators for the firm's own country-specific export experience and for the presence of owners from the same country. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

unchanged.

Connecting person definition. We next explore the robustness of the results in Section 3.2 to different definitions of the person network. For completeness we work with specifications similar to those in columns 3 and 4 of Table A1 that, besides import, also include other types of peer

Table A2: Peer effects with different definitions of person network

Dependent variable: starting to import

	Connecting person definition			
	any connection		from signing right to ownership	
Type of other experience:	Exporter (1)	Owner (2)	Exporter (3)	Owner (4)
Peers with import experience in:				
same building	0.22***	0.22***	0.22***	0.22***
	(0.03)	(0.03)	(0.03)	(0.03)
neighbor building	0.04*	0.04**	0.04*	0.04**
	(0.02)	(0.02)	(0.02)	(0.02)
cross-street building	0.03	0.03	0.03	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
person network	0.26***	0.26***	0.08	0.10
	(0.06)	(0.06)	(0.09)	(0.09)
ownership network	0.50***	0.52***	0.51***	0.53***
	(0.05)	(0.05)	(0.05)	(0.05)
Peers with other experience	Yes	Yes	Yes	Yes
Firm-year FE	Yes	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes	Yes
Observations	3,778,517	3,778,517	3,778,517	3,778,517

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Dependent variable is an indicator for the firm starting to import from the country in the given year. Right-hand side variables are indicators for specific types of peers with prior country-specific import experience, as well as export experience in columns (1),(3) or country-specific owners in columns (2),(4). As connecting people columns (1)-(2) use managers with signing right, owners or supervisory board members, and columns (3)-(4) use owners of the firm who had signing right in the peer before. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

experience. Table A2 reports the results. In columns 1 and 2 we use a broader definition than that in the main text: we relax the requirement that the connecting person needs to have signing rights, and allow her/him to have any kind of measurable connection to both firms. Thus, in this definition a person p is connected to a firm A if (i) p has signing rights in A; or (ii) p is an owner of A; or (iii) p is member of the supervisory board of A. We then use this definition to create our broader measure of firm-to-firm connections by defining B to be a peer of A at t if there is a person connected to A at t who was connected to B at some prior date. Like in the main text, we exclude peers who are in the same ownership networks, which here implies in particular that we eliminate those connections where the connecting person was an owner in both firms.

In columns 3 and 4 we use a narrower definition than the one in the main text. Here, the connecting person must be both an owner of A at t and must have had signing rights in B at a prior date. A potential benefit of this specification is that it reduces the problem of reverse causality emerging if firms purposefully hire managers with specific import experience: it seems less likely that firms purposefully recruit owners with specific import experience.

Table A2 shows that all our results are robust to using the broad definition. However, with the narrow definition the coefficient of person-connected peers remains positive but becomes insignificant. This could be because the reverse causality effect was driving our main estimates, but could also be explained by using a too restrictive measure of person-connections. The other coefficients in these regressions are as expected, showing that the effect of experienced peers in the other peer groups is not sensitive to changes in the definition of person-connected peers.

Different sample definitions. In the baseline specification of 3.2 we estimate the effect of experienced peers on the decision to start importing. To do this we use a sample in which, at every observation, the firm has not yet imported from the particular country until the previous year. In Table A3 we consider different sample definitions, which allows us to answer slightly different questions. For completeness, we estimate regressions analogous to those in columns 3 and 4 of Table A1 which control for both the import and other possible experience of peer firms. In columns (1) and (2) we include all firm-country pairs in all years. These specifications answer the question of whether a firm imports from a country in a year with a higher probability if it has peers with country-specific experience, irrespective of the firm's own import experience. Columns (3) and (4) include each firm only in the single year in which it starts to import from the group of the four countries for the first time, with a separate observation for each of the four countries. This specification asks whether—conditional on starting to import from one of the four countries—the firm is more likely to import from the country with which some of its peers have experience. Columns (5)

Table A3: Peer effects with different sample definitions

Dependent variable: importing

	All firms		First ever importers		Not yet importers	
Type of other experience:	Exporter (1)	Owner (2)	Exporter (3)	Owner (4)	Exporter (5)	Owner (6)
Peers with import experience in:						
same building	0.78***	0.78***	9.12***	9.03***	0.16***	0.16***
	(0.08)	(0.08)	(1.84)	(1.84)	(0.02)	(0.02)
neighbor building	0.08	0.08	1.39	2.20	0.02	0.03
	(0.05)	(0.05)	(2.35)	(2.33)	(0.02)	(0.02)
cross-street building	0.18**	0.17**	0.46	0.10	0.01	0.01
	(0.07)	(0.07)	(2.68)	(2.69)	(0.02)	(0.02)
person network	1.77***	1.77***	10.40***	11.70***	0.37***	0.38***
	(0.26)	(0.26)	(3.82)	(3.73)	(0.09)	(0.09)
ownership network	2.14***	2.21***	21.90***	23.00***	0.43***	0.44***
	(0.15)	(0.15)	(2.69)	(2.67)	(0.05)	(0.05)
Peers with other experience	Yes	Yes	Yes	Yes	Yes	Yes
Firm-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,845,272	3,845,272	23,404	23,404	3,663,512	3,663,512

Notes: Sample includes firm-country pairs in all years in columns (1)-(2); in the year when the firm started to import from the country group for the first time in columns (3)-(4); in those years when the firm has not imported from any of the countries by the previous year in columns (5)-(6). Dependent variable is an indicator for the firm importing from the country in the given year. Right-hand side variables are indicators for specific types of peers with prior country-specific import experience, as well as export experience in columns (1),(3),(5) or country-specific owners in columns (2),(4),(6). Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

and (6) are the closest to our baseline specification, but these specifications exclude entirely those firms which have already imported from one of the four countries. In all these specifications, the patterns are consistent with the spillover hypothesis. In particular, the import experience of same-building neighbors and person-connected peers, as well as the export and ownership experience of same-building peers (not reported in table), robustly predicts firm importing. Thus our results are robust to plausible changes in sample definitions.

A.2 Mover design, export spillovers and other specifications

Mover design. Table A4 presents summary statistics about the data we use in the mover design of Section 3.3. The Table shows that almost half of all incumbent firms—defined as firms which stay in the same building from one year to the next—experience a firm moving into their building, and that about 9% of all incumbent firms have a mover with previous import experience. 88% of the incumbents are on addresses with no import experience from at least one of the four countries. Almost half of these firms have a mover coming into the building, and for about 5% of these firms the mover has country-specific experience that did not exist at the address. 22% of the addresses have no import experience from one of the countries and attract a mover firm, and 2% attract an experienced mover. These numbers show that there are many observations in many distinct addresses which our mover research design can exploit.

Table A4: Descriptive statistics for buildings with new firms moving in

	Number of			
	incumbent firms	addresses of incumbents	incumbent firms	addresses of incumbents
	on all addresses		on addresses without import experience	
Total	211,453	76,433	184,978	66,596
With a mover	105,214	19,976	87,754	16,833
With a mover having previous import experience from				
any of the 4 countries	18,163	2,251	8,951	1,478
the Czech Republic	11,362	1,255	3,415	645
Slovakia	8,907	1,036	3,231	598
Romania	6,696	749	3,124	475
Russia	4,798	563	2,462	352

Notes: We define incumbents as firms staying in the same building as in the previous year. A mover is a firm changing its address within Budapest from one year to another. An address has no import experience with a country if no incumbent firm in the that or in neighboring buildings has imported from the country up to that year. The mover might or might not have import experience.

Table A5 presents the same estimation results we show in Figure 1. We include the table here to show the the precise value of the point estimates and the standard errors.

Export spillovers. Table A6 shows estimates for export spillovers, using the same identification

Table A5: Effect of experienced peer moving into building on same-country imports

Dependent variable: importing	(1)	(2)
Event-year 1	0.38***	0.12
* Experienced mover in building	(0.11)	(0.13)
Event-year 2	0.83***	0.38*
* Experienced mover in building	(0.18)	(0.23)
Event-year 3	0.98***	0.65**
* Experienced mover in building	(0.23)	(0.30)
Event-year 4	1.29***	0.78**
* Experienced mover in building	(0.31)	(0.38)
Event-year N≥5	1.33***	0.73
* Experienced mover in building	(0.45)	(0.55)
Event-year indicators	Yes	No
Firm-year FE	No	Yes
Country-year FE	No	Yes
Observations	1,101,848	1,101,848

Notes: Sample includes firm-country pairs in years after a mover firm entered the building of the firm, conditional on no incumbent firm in the building imported from the country before that. Dependent variable is an indicator for the firm importing from the country. Right-hand side variables are event-year indicators showing the move occurred N years ago, interacted with a country-specific experience indicator for the mover. We winsorize event-years at 5. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

strategy with which we estimate import spillovers in our first research design of Section 3.2. As we discuss in Section 3.4, spillover patterns and magnitudes are comparable to those of import spillovers.

Other specification checks. In Table O10 of the Online Appendix we explore whether spillovers affect not only the decision to import but also the duration and volume of importing. In our data we do not find a clear evidence for either.

Table A6: Peer effect in exporting behavior

Dependent variable: starting to export					Type of other experience	
					Importer	Owner
	(1)	(2)	(3)	(4)	(5)	(6)
Peers with export experience in:						
same building	0.17***			0.17***	0.16***	0.16***
	(0.02)			(0.02)	(0.02)	(0.02)
neighbor building	0.04**			0.04**	0.04*	0.04**
	(0.02)			(0.02)	(0.02)	(0.02)
cross-street building	0.04			0.04	0.05*	0.04
	(0.03)			(0.03)	(0.03)	(0.03)
person network		0.40***		0.38***	0.37***	0.37***
		(0.09)		(0.09)	(0.09)	(0.09)
ownership network			0.49***	0.49***	0.47***	0.48***
			(0.05)	(0.05)	(0.05)	(0.05)
Peers with other experience in:						
same building					0.05**	0.06**
					(0.02)	(0.03)
neighbor building					0.01	0.00
					(0.02)	(0.02)
cross-street building					-0.03	-0.01
					(0.03)	(0.03)
person network					0.09	0.58
					(0.09)	(0.36)
ownership network					0.09**	0.12
					(0.04)	(0.08)
Firm-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,772,739	3,772,739	3,772,739	3,772,739	3,772,739	3,772,739

Notes: Sample includes firm-country pairs in years in which the firm has not exported to the country by the previous year. Dependent variable is an indicator for the firm starting to export to the country in the given year. Right-hand side variables are indicators for specific types of peers with prior country-specific export experience, as well as import experience in column (5) or country-specific owners in column (6). Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

A.3 Additional specification for heterogeneity of peer effect

Here we present a result on heterogeneity by the quality / success of the peer's import experience. We start with a measure that defines a peer as a "successful" importer from country c in year t if it has imported from c in at least two years within the three-year period $[t-1, t+1]$. Column 1 in Table A7 shows that successful importers have an additional diffusion effect relative to peers with different import patterns, consistent with the idea that these importers have more valuable knowledge. In column 2 we measure success with the length of peer import experience, measured as the maximum number of years during which a peer firm imported from country c minus one, so that we can measure the effect of an additional year of experience. And in column 3 we report a similar specification in which length of experience is the longest continuous import experience allowing for single-year gaps. Both of these specifications show that longer import experience by the peer is associated with higher adoption. Thus overall we find that more successful importers, perhaps because they have more import-related knowledge, generate higher spillovers.

Table A7: Heterogeneity of peer effect by peer success in importing

Dependent variable: starting to import Specification:	Recent successful experience (1)	Length of experience measured by number of years (2)	 number of recent years (3)
Same-building importer peers		0.17*** (0.03)	0.14*** (0.03)
Same-building successful importer peers	0.35***° (0.05)		
Same-building non-successful importer peers	0.20*** (0.03)		
Length of peers' import experience in same building		0.06*** (0.01)	0.07*** (0.01)
Other types of experienced peers	No	Yes	Yes
Other types of experienced peers by import success	Yes	No	No
Length of import experience of other peer types	No	Yes	Yes
Firm-year FE	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes
Observations	3,778,517	3,778,517	3,778,517

Notes: Sample includes firm-country pairs in years in which the firm has not imported from the country by the previous year. Dependent variable is an indicator for the firm starting to import from the country in the given year. Right-hand side variables are indicators for peers with prior country-specific import experience by peer success in importing in column (1) and adding maximum length of peer experience beyond a single year in column (2)-(3). A peer is successful if it imports from the country at least twice in the period $[t-2, t]$. We consider only recent continuous experience allowing for single-year gaps in column (3). Other types of peers refer to all other (non-same-building) peer categories in Table 3. Standard errors in parentheses are clustered by building. Coefficients are multiplied by 100 to read as percentage point marginal effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. ° denotes that the coefficient is significantly different from that of previous group at 5%.

B Social multiplier calculations

This subsection shows how to calculate the 5-year social multiplier using a simple model of knowledge flows affecting import entry. Consider a building with N^l low-productivity and N^h high-productivity firms. The states of the Markov process describing the dynamics of importing defined in Section 5 can be represented as (m^l, m^h), where m^l is the number of low-productivity importers, m^h is the number of high-productivity importers, and $m^l \in 0, 1, ..., N^l$ and $m^h \in 0, 1, ..., N^h$. Let $A^{(N^l,N^h)}$ denote the transition matrix for this process. The transition probability from state (m_1^l, m_1^h) to (m_2^l, m_2^h), denoted $a^{(N^l,N^h)}_{(m_1^l,m_1^h),(m_2^l,m_2^h)}$, is defined as follows. When $(m_1^l \leq m_2^l)$ and $(m_1^h \leq m_2^h)$, we have

$$a^{(N^l,N^h)}_{(m_1^l,m_1^h),(m_2^l,m_2^h)} = \binom{N^l - m_1^l}{m_2^l - m_1^l} \cdot (p_l + m_1^l \cdot \beta_{ll} + m_1^h \cdot \beta_{lh})^{m_2^l - m_1^l} \cdot (1 - p_l - m_1^l \cdot \beta_{ll} - m_1^h \cdot \beta_{lh})^{N^l - m_2^l} \cdot \\ \binom{N^h - m_1^h}{m_2^h - m_1^h} \cdot (p_h + m_1^l \cdot \beta_{hl} + m_1^h \cdot \beta_{hh})^{m_2^h - m_1^h} \cdot (1 - p_h - m_1^l \cdot \beta_{hl} - m_1^h \cdot \beta_{hh})^{N^h - m_2^h}. \quad (8)$$

In all other cases we have $a^{(N^l,N^h)}_{(m_1^l,m_1^h),(m_2^l,m_2^h)} = 0$. In the expression p_l denotes the baseline probability of starting to import for a low-, and p_h for a high-productivity firm. $\beta_{gg'}$ is the estimated effect of an additional peer in productivity group g' on the import entry probability of a firm in productivity group g, with $g, g' \in \{\text{l,h}\}$.

The 5-year transition matrix is given by $(A^{(N^l,N^h)})^5$, with elements $a^{(N^l,N^h)5}_{(m_1^l,m_1^h),(m_2^l,m_2^h)}$. Then the expected number of importers conditional on an initial state (m^l, m^h) is

$$\sum_{k^l=m^l}^{N^l} \sum_{k^h=m^h}^{N^h} (k^l + k^h) \cdot a^{(N^l,N^h)5}_{(m^l,m^h),(k^l,k^h)}. \quad (9)$$

It follows that the numerator of our social multiplier (7) for firm i on address a starting to import from country c can be computed as

$$\sum_{k^l=M^{c,l}_{a(i),s}+I_{g(i)=l}}^{N^l_{a(i),s}} \sum_{k^h=M^{c,h}_{a(i),s}+I_{g(i)=h}}^{N^h_{a(i),s}} (k^l+k^h)\cdot a^{(N^l_{a(i),s},N^h_{a(i),s})5}_{(M^{c,l}_{a(i),s}+I_{g(i)=l},M^{c,h}_{a(i),s}+I_{g(i)=h}),(k^l,k^h)} - \\ - \sum_{k^l=M^{c,l}_{a(i),s}}^{N^l_{a(i),s}} \sum_{k^h=M^{c,h}_{a(i),s}}^{N^h_{a(i),s}} (k^l+k^h)\cdot a^{(N^l_{a(i),s},N^h_{a(i),s})5}_{(M^{c,l}_{a(i),s},M^{c,h}_{a(i),s}),(k^l,k^h)}, \quad (10)$$

where $M^{c,g}_{a(i),s}$ is the number of peers in productivity group g in the building of firm i in year s importing from country c; $I_{g(i)=g}$ is an indicator for firm i being in productivity group g; and $N^g_{a(i),s}$ is the number of firms in productivity group g in the building of firm i in year s, with $g \in \{\text{l,h}\}$. We use the same formula for the denominator but set the $\beta_{gg'}$ parameters to zero.