Migration and the location of MNE activities: evidence from Italian provinces

Received: 8 April 2022 | Revised: 13 June 2023 | Accepted: 30 October 2023

DOI: 10.1111/jors.12674

RESEARCH ARTICLE

JOURNAL OF REGIONAL SCIENCE WILEY

Migration and the location of MNE activities: Evidence from Italian provinces

Luigi Benfratello[1] | **Davide Castellani**[2] | **Anna D'Ambrosio**[1]

[1]Politecnico di Torino, Corso Duca degli Abruzzi, Torino, Italy

[2]Henley Business School, University of Reading, Reading, UK

Correspondence
Anna D'Ambrosio, Politecnico di Torino, Corso Duca degli Abruzzi, 24, Torino, Italy.
Email: anna.dambrosio@polito.it

Funding information
Compagnia di San Paolo

Abstract

We investigate the migration-inward foreign direct investment (FDI) nexus in narrow geographies. A novel two-stage empirical strategy allows us to investigate the role of migration as a determinant of multinational enterprises (MNEs) location choices and unpick heterogeneity in foreign investors' preferences towards the presence of migrants in the host location. This allows us to shed light on the relative importance of the underlying mechanisms linking migration and inward FDI. Relying on 1113 greenfield investments by 895 MNEs in Italian NUTS3 regions over 2003–2015, we find that immigrants from the country of origin of the investor exert a positive but highly heterogeneous effect on MNE location choices. Investors are more sensitive to the presence of migrants from their country of origin when they lack experience in the destination country (first-time investors) and when their investments concern market-access or business-services (downstream) activities. This is consistent with the view that migrants act as information brokers that bridge the fixed costs of international business activities.

KEYWORDS

conditional logit, demand effect, foreign direct investment, information effect, location choice, migration, mixed logit

1 | INTRODUCTION

Flows of people and flows of capital contribute to a more and more interconnected global economy, but they feature quite differently in the public debate. Advanced economies' policymakers tend to consider foreign capital as a key asset and compete to attract multinational enterprises (MNEs). Instead, they often perceive migration, especially inward migration, as a public concern.

By contrast, the economic literature has robustly documented that foreign direct investment (FDI) and migration are complements (Buch et al., 2006; Burchardi et al., 2018; Felbermayr et al., 2015; Javorcik et al., 2011; Jayet & Marchal, 2016; Olney & Pozzoli, 2021). Explanations for this complementarity draw on a variety of theoretical and empirical approaches, which ultimately highlight that migrants may affect different components of MNEs' profit functions, that is, labor costs, fixed and variable costs, and demand.

Understanding the mechanism behind migrant effects is important to inform host economies about how they may leverage desirable side effects of migration. In this paper, we study whether migration acts as a factor that makes regions more attractive for foreign capital, implementing a novel empirical strategy to shed light on the sources of heterogeneity in this relation and, ultimately, the underlying mechanisms.

Based on Defever (2006, 2012), we argue that MNEs establishing different types of subsidiaries have heterogeneous valuations of the factors entering the profit function, among which the migrant effects. For instance, when opening a manufacturing branch, MNEs will mainly consider labor cost factors; when opening marketing offices, they will be more sensitive to demand factors. In turn, if migrants affect production costs, they will have stronger effects on MNEs' location when they establish manufacturing, rather than marketing branches.

The weight given to particular location factors may also depend on firm characteristics (Helpman et al., 2004). For instance, inexperienced firms may be more sensitive to the costs of collecting information about the FDI destination, a factor that migrants may influence. Hence, investigating the sources of the heterogeneity in migrant effects provides a way to learn about which mechanism prevails.

We investigate this research question employing a location choice model with heterogeneity in investors' preferences (e.g., Alcácer et al., 2018; Castellani & Lavoratori, 2020), which we augment with migration variables. This allows us to study whether migrants affect the location choice of inward FDI and whether their effects are heterogeneous. We then explore the sources of this heterogeneity to tease out the role that migration plays for firm-specific location choices (Alcácer et al., 2018; Castellani & Lavoratori, 2020; Hornstein & Greene, 2012; Saxonhouse, 1976).

We leverage detailed information from the *fDi Markets* database on greenfield investments and highly disaggregated data about the locations, that is, Italian NUTS3 regions, or "provinces," to investigate the migration-FDI nexus at a quite fine-grained geographical scale. Italy is a relevant setting for this analysis, given its marked subnational heterogeneity in economic performance and FDI attractiveness, and its remarkable feature of being both a relevant country of origin and of destination for migration. We consider location choices among alternative locations within Italy, under the assumption that regions within Italy are closer substitutes with each other than with regions outside Italy, in line with the findings by Basile et al. (2005, 2009).

Previous studies have exploited the heterogeneity in migrant effects to learn about the underlying mechanisms. However, multiple sources of heterogeneity may be correlated. To our knowledge, no contribution directly confronts them to learn about the prevailing mechanism. Our approach allows us to identify the most significant sources of heterogeneity while remaining ex-ante relatively agnostic about the prevailing factors. Moreover, the heterogeneity in MNEs' activities and firm experience that we explicitly consider has so far received limited attention in the literature, leaving crucial determinants of location choice unobserved. Providing evidence on the extent to which migrant effects can be linked to firm- and investment-specific characteristics, our study adds to the small literature that investigates the migration-FDI nexus based on microdata (e.g., Burchardi et al., 2018; Olney & Pozzoli, 2021).

Our econometric strategy draws on a two-stage estimation approach. First, we employ random parameter (mixed) logit models—allowing for repeated location choices by the same investor—to assess whether the effect of migration on the location of inward FDI is on average positive and heterogeneous across investors. Then, we simulate the related individual-specific random parameters associated with migration and examine the drivers of their distribution in a regression framework.

We distinguish between immigrants and emigrants (Parsons, 2012), as well as between bilateral and multilateral migrants. We call "bilateral immigrants" the foreign residents from the country of origin of the investment who live in the province where the firm may locate, and "bilateral emigrants" the residents of the province who moved to the origin country of the investment. "Multilateral immigrants" refers to the overall stock of immigrants from any country residing in a province, while "multilateral emigrants" to the overall stock of residents from the province that expatriated to any country. We deem it important to include both bilateral and total migrants given that they may capture different effects (Olney & Pozzoli, 2021).

Our results confirm a positive and robust effect of bilateral immigrants on FDI but no effect of bilateral emigrants normultilateral migration. The immigrant effect is driven by skilled immigrants and is strongly localized, consistent with the localized nature of knowledge spillovers (Buch et al., 2006; Rauch, 2001) and with the strong geographic concentration of FDI within urban and even suburban areas (Belderbos et al., 2020; Goerzen et al., 2013).

While immigration effects turn out positive for a vast majority of the investors, about 80%–88%, the magnitudes of these effects vary. We find that investors are more sensitive to the presence of migration networks when they lack experience in the destination economy and when their investments concern market-access or business-services activities. We argue that migrant networks enter the MNE profit functions as factors that decrease the cost of searching for information—mainly market information—about the location.

Our results are robust to different sources of endogeneity. We address unobserved heterogeneity at the local level with province-level dummies, and reverse causality and time-varying omitted variable bias implementing a control function based on standard Altonji–Card instruments for bilateral migration. Moreover, our results prove robust to controlling—within the control function approach—for most of the dyadic province–country heterogeneity.

The paper is organized as follows. Section 2 discusses the literature related to our study. Sections 3 and 4 introduce the empirical model, the data and variables, and present summary and descriptive statistics distinguishing between the first and second stages. Section 5 shows our results along with a set of robustness checks. A discussion of their implications and some concluding remarks follow in Section 6. A Supporting Information Appendix provides details on the data sources and variables, as well as supplementary descriptive statistics and results.

2 | RELATED LITERATURE

The standard assumption in the vast literature on the determinants of FDI location choice is that a firm decides to locate its subsidiaries where the achievable profits outweigh those that can be gained in all other available locations (e.g., Basile, 2004; Basile et al., 2008; Defever, 2006; Head et al., 1999; Head & Mayer, 2004; Nielsen et al., 2017; Spies, 2010). In broad terms, reduced-form formulations of such functions include production costs (Amiti & Javorcik, 2008; Tintelnot, 2017), the market potential of the location (Head & Mayer, 2004), fixed costs of market entry (Helpman et al., 2004; Spies, 2010), and information spillovers arising from agglomeration economies (Fujita & Thisse, 1996).

The increasingly compelling evidence of a positive effect of migration on FDI (e.g., Buch et al., 2006; Burchardi et al., 2018; De Simone & Manchin, 2012; Docquier & Lodigiani, 2010; Etzo & Takaoka, 2018; Gao, 2003; Javorcik et al., 2011; Jayet & Marchal, 2016; Kugler & Rapoport, 2007) can be easily integrated into this framework. The variety of theoretical and empirical approaches adopted (see the review in Jayet & Marchal, 2016) can ultimately be traced back to two main explanations: immigrants affect transnational information and enforcement costs, that is, the fixed costs of establishing a plant abroad (Buch et al., 2006; Docquier & Lodigiani, 2010; Gao, 2003; Javorcik et al., 2011; Kugler & Rapoport, 2007), or they reduce production costs.

The pioneering contributions on migrants' information and enforcement effects, originally proposed to explain trade (Gould, 1994; Rauch & Trinidade, 2002; Wagner et al., 2002), have now been fully incorporated into trade models that recognize the central role of contacts and networks (Arkolakis, 2010; Chaney, 2014) and mainly concern the effects of bilateral migration. Migrant knowledge of the home and destination country institutions, practices, tastes, and language facilitates the flow of information regarding procedures and business opportunities—the "information effect." Furthermore, migrants' embeddedness in coethnic networks generates reputational bounds that ensure the enforcement of transnational contracts and can be valuable in countries where the rule of law is weakly enforced—the "enforcement effect" (Dunlevy, 2006; Rauch & Trinidade, 2002).

FDIs entail substantial capital investment, information asymmetries, and cognitive barriers and critically depend on the knowledge of foreign institutions, business opportunities, and labor market pools of specific skills, making information and enforcement effects pivotal (Daude & Fratzscher, 2008; Head et al., 1995; Javorcik et al., 2011; Jayet & Marchal, 2016). FDIs require comparatively greater fixed costs and lower variable costs than trade (Buckley & Casson, 1981; Helpman et al., 2004), and migrant effects have been found to operate precisely at the level of fixed rather than variable costs (Peri & Requena-Silvente, 2010).

Although most studies applied these considerations to migrant inflows, a few studies on trade noted that information and enforcement effects may apply to both directions of migration, that is, immigration and emigration (e.g., D'Ambrosio & Montresor, 2022; Felbermayr et al., 2015; Parsons, 2012).

As for production cost advantages, they may arise from the presence of both bilateral and multilateral immigrants. Greater stocks of multilateral immigrants may affect labor costs primarily because they provide comparatively low-cost labor, which may in turn potentially induce decreases in domestic labor costs as a whole (Jayet & Marchal, 2016; Markusen, 2006; Olney & Pozzoli, 2021). These considerations also apply to bilateral immigrants, to the extent that they originate from countries with lower labor costs, which is largely the case in Italy (see Supporting Information Appendix C). In addition, immigrants from the same country of origin as the investor may provide skills and work habits that are similar or complementary to those of the investing firms, hence particularly valuable for investors (Burchardi et al., 2018).

A third, less investigated potential channel is posited in Nefussi and Schwellnus (2010) and Burchardi et al. (2018): the similarity in preferences. Migrants' consumption bias in favor of goods and services offered by firms from their home countries contributes directly to demand and enters the assessment of market potential for an affiliate. The "preference" or "transplanted-home bias effect" is indeed an established channel of bilateral migrants' effects on trade (Gould, 1994; White, 2007).[1]

The different effects may clearly coexist, and we can only observe their combined result. Burchardi et al. (2018) provide, to the best of our knowledge, the first contribution that seeks to single out the prevailing mechanism behind migrants' effects. They document a robust causal information effect, while they do not find robust evidence of the other effects.

A limitation of standard heterogeneity analysis is that they separately explore the different underlying mechanisms, identifying the effect of migrants on one aspect at a time (e.g., the effect of migrants depending on the type of goods, on the specific activity of the new venture, or on the country of origin of the investor). Yet, different dimensions are likely to be correlated, and standard approaches may fail to highlight which channels end up prevailing empirically. In this paper, we go beyond estimating the average migrant effects on location to provide a

[1]For the sake of completeness, we also consider the effect of multilateral emigration, although we are not aware of any contributions that explicitly argued for a role of this factor in MNEs' location choice. Indeed, larger emigration rates may act as a factor that "drains brains" and entrepreneurial capital from a particular province and, ceteris paribus, reduces a location's attractiveness by changing its human capital and demand composition (Anelli et al., 2023). On the other hand, large emigration rates may open the region to exchanges with other countries and act as an attraction factor.

comprehensive analysis of the drivers of migrants' effects heterogeneity, with a view to learning about the underlying mechanisms.

The first driver of the heterogeneity is the type of activities carried out by MNEs. Defever (2006, 2012) recognizes that different MNE activities have heterogeneous determinants and attributes activity-specific weights to the different terms of the profit function. Hence, if migrants affect a specific component of the profit function, investments that are more sensitive to that component should be more responsive to the presence of migrants. Clearly, FDI are complex operations and we do not claim that FDI functions are exact proxies of single theoretical mechanisms. Moreover, the relative salience of each of them depends on the elasticities of labor supply, information costs, and demand for migration. Nonetheless, we do think that the relative importance of migrants' effects in different functions can provide useful interpretative guidelines in this regard.

Knowledge-seeking ventures tend to pursue the exploitation of localized knowledge assets (Chung & Alcácer, 2002), the access to which requires complex interactions between local and expatriate labor. In this case, migrant effects will arguably depend on relatively specific information effects mediating cultural differences and facilitating the transfer of knowledge and know-how, which may require a comparatively high level of skills (Defever, 2006; Foley & Kerr, 2013). In contrast, we expect investments involved in production or construction to rely less on information and more on the availability of specific inputs and (low-cost) labor (Defever, 2006; Markusen, 2002). Downstream activities, being more sensitive to the proximity to customers and market size, are likely to be more reactive to the market information mediated by migrants and to their own demand for home-country goods (Defever, 2006, 2012).

We also leverage previous contributions to learn about other dimensions of heterogeneity that may explain the mechanism. First, labor intensity may inform us about the relevance of labor costs (Jayet & Marchal, 2016; Ottaviano et al., 2013). Second, the type of goods produced may be instructive about the presence of a demand effect: firms producing final goods and services will rely more on the consumption home bias of their coethnics when planning foreign sales (Burchardi et al., 2018).

Third, some firm characteristics may explain the salience of information effects. Larger firms presumably have more efficient management structures and better information-processing capacity and, therefore, benefit less from the information-brokering effects of immigrants. Firms with more international experience, better organizational structures, and better knowledge of the destination economy will experience lower fixed costs of opening new ventures and will rely less on the information bridged by migrants (Foley & Kerr, 2013). Peri and Requena-Silvente (2010) draw on the theoretical model by Chaney (2008) and argue that migrants exert an information effect by bridging the fixed costs of international trade. If they also bridge the fixed costs of international investments, they should have a different effect on investing firms depending on their experience in the host market. More experienced firms will presumably have better knowledge and a more established network of references supporting the new venture setup. In turn, we expect that they will rely less on the immigrants' facilitating role. According to Girma and Yu (2002) and Tadesse and White (2008), migrants from more culturally distant countries exert stronger trade effects because cultural distance increases the need for a brokering effect. On the other hand, reliance on more codified knowledge and standardized, less culture-specific procedures may decrease it (D'Ambrosio et al., 2019).

3 | EMPIRICAL MODEL AND VARIABLES

In this section, we introduce our empirical model and variables. We start with a presentation of our random parameter model. Random parameter models not only yield estimates of the parameters' average effects but also of their variance. This allows exploring whether location factors affect choice differently depending on the decisionmaker (Alcácer et al., 2018). We then move to the analysis of the drivers of parameter heterogeneity. We first describe the most straightforward way to address this issue, which is to add interaction effects. We then move to the discussion of our proposed approach which involves a second stage, in which individual investor-specific parameters are regressed on firm-level drivers of parameters heterogeneity. Finally, we discuss the strategies we implement to cope with endogeneity concerns.

3.1 | First stage: Location-choice model

3.1.1 | Random parameter (mixed) logit

Our paper draws from the literature on the location choice for FDI and on migrant effects on trade and FDI. Both literature branches have a fairly established set of estimation strategies of reference. The literature on the migration–FDI nexus mainly applies gravity-like models (see, e.g., Buch et al., 2006; Burchardi et al., 2018; De Simone & Manchin, 2012; Felbermayr et al., 2015; Javorcik et al., 2011). However, addressing our research question from a gravity perspective imposes aggregating investments at some geographic scale. In turn, this implies losing information about the decision-making process underlying location choice. Hence, we follow the literature that studies the location choice for each individual investment (e.g., Basile et al., 2008; Castellani & Lavoratori, 2020; Chung & Alcácer, 2002; Defever, 2006; Du et al., 2008; Head et al., 1995; Nielsen et al., 2017; Spies, 2010), which mostly appeals to discrete-choice models (conditional, nested, and mixed logit models; see Marschak, 1974; McFadden, 1974; McFadden & Train, 2000; Train, 2009). These models share an underlying random utility model, that is, a model assuming, in a partial equilibrium setting, that the location chosen by a multinational firm yields the highest utility/profit compared to the other possible locations, subject to uncertainty deriving from unobservables (Train, 2009).

In essence, we model the decision-making process of a firm planning to carry out some foreign production in Italy and has to choose the location of its activities. The dependent variable *"Choice"* is equal to one if a specific alternative is selected, and zero for all other alternatives in the choice set. In our application, the alternatives are the set of Italian provinces where the FDI could locate.[2] The decisionmaker is the investing company f facing investment decision n. Hence, the number of choices under consideration for our analysis is equal to the number of potential locations J times the number of investment projects N. The probability of choosing a specific province depends only on the difference in utility that province i yields to firm f in investing decision n compared with the other alternatives $j \neq i$. The absolute value of utility does not matter. Hence, the choice will be unaffected by attributes of the alternative that do not induce a difference in utility or by attributes of the decisionmaker that do not vary over alternatives; the effects of these factors will not be estimated. This implies that firm-specific characteristics (e.g., its country of origin, the gross domestic product [GDP] of the origin country, its size, knowledge, capital investment, etc.) can be included in the specification only if interacted with alternative-varying variables (see Train, 2009). On the other hand, bilateral variables such as migration from a given country to a given province induce a difference in utility and can be estimated.[3]

The simplest and most widely used discrete-choice model is the conditional logit model, which relies on the independence from irrelevant alternatives assumption and on an assumption of homogeneous effects of the parameters across decisionmakers. However, these assumptions would be at odds with our discussion about the heterogeneity in companies' evaluations of FDI determinants. Therefore, we base our empirical strategy on a set of mixed logit models (also known as random parameter logit models), which relax this assumption and allow for flexible unspecified patterns of correlation among alternatives, so that some alternatives may be closer substitutes than others (Train, 2009). It can be shown that this is equivalent to allowing heterogeneous parameters across

[2]To avoid possible selection issues, we include all Italian provinces, regardless of whether they received MNE investments or not, in our location-choice set. The results are virtually unaffected if we restrict the sample to provinces that received at least one investment.

[3]It is worth mentioning that our analysis does not have a longitudinal dimension, because we only observe each investment once. However, our analysis pools investments across several years. In practice, for an investment happening at time t the investor has a choice among J possible locations and compares the characteristics of these locations prior to the investment. We operationalize this decision-making process by matching investment projects at time t with location determinants at $t - 1$.

decisionmakers. Indeed, as we shall explain, mixed logit models not only yield an estimate of the average effect of a regressor but also of its standard deviation. A significant standard deviation can be interpreted as a sign of unexplained heterogeneity in the effects; an insignificant one does not allow one to reject that the effect is constant across decisionmakers.

As discussed, we are interested in the heterogeneity in migrant effects across investing firms. Specifically, as we discuss in Section 3.2.2, we use the distributional estimates as the starting point for a second-stage analysis of the sources of heterogeneity in the effects (Castellani & Lavoratori, 2020; Hornstein & Greene, 2012; Saxonhouse, 1976), which guides our investigation of the drivers of the immigrant effects. Although discrete-choice models are fairly standard in the FDI literature, their application to the analysis of migration effects on greenfield FDI and the exploration of sources of heterogeneity in the migrant effects with investment and firm-level determinants are, to the best of our knowledge, novel contributions.

Each firm f makes one or more location decisions. At each decision n, the profit deriving from investing in province j is assumed linear in the parameters. Profit is modeled as a function of alternative-specific characteristics. As anticipated, in the mixed logit model some coefficients are allowed to vary by decisionmaker. The expected profit yielded from location j to firm f at decision n is

$$\pi_{fnj} = \beta' x_{fnj} + \delta_f' z_{fnj} + \epsilon_{fnj}. \tag{1}$$

In this equation, x_{fnj} and z_{fnj} are vectors of province–investment–company location determinants. β is a vector of fixed coefficients, assumed to be the same across investors, δ_f is a vector of random (i.e., investor-specific) coefficients, and ϵ_{fnj} is iid extreme value. Hence the location determinants in z_{fnj} may display different effects for different investors, and δ_f can be interpreted as the investor's specific preferences regarding location determinants. Instead, the determinants in x_{fnj} are assumed to have the same effect, captured by β, for all investors.[4] [5]

Ex ante, we do not have strong reasons to expect that particular variables will have the same effect on all investors, so our starting point is that all variables enter our model with random coefficients and we are interested in learning whether empirically they turn out to be significantly heterogeneous or not. As explained in Section 3.1.2, practical considerations impose that some variables enter with fixed coefficients.

To keep model presentation simple, we refer to vectors of variables x_{fnj} and z_{fnj}, varying by destination province, investment, and investing company. In practice, our model contains variables varying at the level of the destination province j, at the province–company level jf, at the province–investment level jn, and at the province–investment–company level fnj (see Section 3.1.2).

Each investor f knows the value of β and of their own δ_f and chooses province i if the utility associated with it exceeds that of all other provinces $j \neq i$. Conditional on β and on a specific δ_f, the choice probability in a particular choice situation reduces to a standard logit choice probability:

$$L_{fni}(\beta, \delta_f) = \frac{e^{\beta' x_{fni} + \delta_f' z_{fni}}}{\sum_j e^{\beta' x_{fnj} + \delta_f' z_{fnj}}}. \tag{2}$$

However, the random coefficients δ_f cannot be directly estimated, and the analyst cannot condition on δ_f. The unconditional probability that a given project locates in a specific province i is the integral of $L_{fni}(\beta, \delta_f)$ over all possible values of its distribution $g(\delta)$ (Train, 2009).

[4]The δ_f may be interpreted to be a δ_{fn}, that is, changing by investor *and* investment. However, controlling for the correlation between investments by the same firm, we cannot estimate investment-specific coefficients, unless we include investment-level variables in interaction with the z_{fnj} (see Section 3.1.3).

[5]As we shall discuss, our two-stage approach is similar to including a full set of interaction effects of migration with investor-specific variables. However, our approach is less subject to collinearity issues, which are a significant hindrance to the convergence of these models, and can accommodate a wider set of potential drivers of the migrant effects (see Section 3.1.3).

$$P_{fni} = P(Choice_{fni} = 1|\beta, x, z) = \int_{\delta} L_{fni}(\beta, \delta_f) g(\delta)\, d\delta. \tag{3}$$

Making assumptions about the distribution $g(\delta)$ across firms, we can obtain estimates about its distributional parameters. We assume the distribution to be normal with density $g(\delta_f; \bar{\delta}; \sigma_\delta)$, where $\bar{\delta}$ represents the vector of the expected values of δ_f and σ_δ their variance, both in the population.[6] This follows common practice in the literature when no particular direction of the effects is expected (e.g., Revelt & Train, 1998; Train, 2009).[7] The probability that firm f facing investment decision n will locate in province i can be approximated through a simulation process for different values of the parameters $\bar{\delta}$ and σ_δ.[8]

The estimated $\hat{\bar{\delta}}$ are measures of the average effect of a given location determinant on the choice, while the magnitude and significance of their estimated standard deviations $\hat{\sigma_\delta}$ are measures of the heterogeneity of the effects of z on the location choice for FDI. As mentioned, the main difference between this specification and the one for the standard logit is that we allow δ to vary by decisionmaker. Model 3 reduces to a conditional logit model if $g(\delta)$ is degenerate at fixed parameters d, such that the cumulative of $g(\delta)$ switches from 0 to 1 at $\delta = d$.

To take into account that firms may face more than one location choice, we treat location determinants as varying over companies but constant over choice situations for the same firm. The logit probability in Equation (2) is modified to consider the probability that the decisionmaker takes a particular sequence of choices, hence the integrand in Equation (3) involves a product of logit formulas, one for each decision situation $i = i_1, \ldots, i_I$. The probability is simulated similarly to the probability with one choice period (Train, 2009).

3.1.2 | Determinants of location choice

We analyze the location choices of 1113 inward greenfield FDI projects in 107 Italian provinces occurring over the 2003–2015 period and reported in the *fDi Markets* database. As discussed, our dependent variable is the binary variable *Choice*, equal to 1 if firm f chose to locate investment project n in province j, and zero otherwise.

Our main variables of interest are *log immigrants* and *log emigrants*. *Log immigrants* is the log of the stock of immigrants who originally come from the same origin country o as the foreign direct investor and reside in province j at time $t - 1$, where t is the year of occurrence of the investment. We refer to these as "bilateral immigrants." *Log emigrants* is the log of the stock of emigrants from province j and residing, at time $t - 1$, in the same country o from where the investment originated. We label these "bilateral emigrants." Drawing on previous literature and on our discussion in Section 2, we also include two variables capturing multilateral immigration and emigration: *log multilateral immigrants*, that is, the log of immigrants from any origin country residing in province n (net of bilateral immigrants), and *log multilateral emigrants*, the log of emigrants having moved from province n to any destination (net of bilateral emigrants).[9]

To mitigate the risk of omitted variable bias, we include a wide set of controls. To proxy for localized market potential, we include province-level GDP and province population, both measured in logs. To proxy for labor cost, we include the log of the average wage (calculated at the regional level due to data unavailability at the province level) and the province-level unemployment rates; as proxies for the innovation capacity of the location, we include the (log of) the count of patent applications filed by applicants based in the province to the European Patent Office, as well as the share of residents holding a tertiary degree as a proxy for the human capital endowment. We include an index of the infrastructure endowment and an index for the institutional quality of province j to account for the fact that better infrastructure and institutions attract FDI. As is customary in location analysis, we include measures of agglomeration aiming to capture the

[6]More precisely, σ_δ is the diagonal of the variance–covariance matrix of δ_f, whose off-diagonal elements are assumed to be zero for tractability.
[7]See Train (2009, p. 138) for a review of the distributional assumptions employed in previous studies.
[8]For details about the simulation process, see Supporting Information Appendix B.1. All mixed logit models were run in Stata using the user-written command mixlogit (Hole, 2007).
[9]A detailed description of the variables we use and of their construction and sources is provided in Supporting Information Appendix A.

effects of, respectively, Jacobian and Marshallian externalities. We measure Jacobian externalities by sectoral diversity, calculated as $1 - H$, where H is the sectoral Herfindahl–Hirschmann concentration index at the NACE 2-digit level. Marshallian externalities are captured by a province-level location quotient calculated on the same industry as the investment project. A location quotient of manufacturing activities is added to control for the economic structure of the province and for the fact that immigrants tend to locate in provinces with a greater concentration of manufacturing activities. As a measure of urbanization economies, we include the density of firms per square kilometer. Furthermore, considering the central role of Rome and Milan in the Italian administration and economy, we also include a dummy equal to one if the investment is located in either of these two cities.

To account for the role of gravity in explaining FDI location, we include the log of the distance between the centroid of province j and of the capital city of the country of the investment origin country o, and a dummy taking a value equal to 1 for investors whose home country is located on the border with an Italian region. To capture other time-invariant unobservables at the dyadic level, we include the pre-2002 stock of FDI from the same country to the same province. We also control for other bilateral time-varying variables that may correlate with bilateral migration: the log of bilateral imports and exports, a binary variable equal to 1 if the same parent company has already invested in the province, and zero otherwise (parent colocation).

All regressors are lagged 1 year when included in the model to account for the timing of the decision-making process and to mitigate simultaneity concerns. As mentioned above, our starting point is that all variables may enter the model with random coefficients, which would allow us to appreciate which parameters have a significant mean and which have a significant variance (Alcácer et al., 2018). However, some practical considerations induce us to include some variables with fixed coefficients. First, colocation enters with a fixed coefficient due to the limited number of firms that locate where they have previously invested (less than 1% of our sample). Moreover, the maximum number of random coefficients that the Stata routine `mixlogit` accommodates is 20. As our mixed logit model includes up to 25 independent variables, in the richest specification we had to set four additional variables as fixed. Based on preliminary estimations, infrastructure endowment, province unemployment rate, share of tertiary educated, and sectoral diversity did not show significant standard deviations, so we opted to set their parameters as fixed.

3.1.3 | Explaining the heterogeneity in migrants' effects based on investment characteristics

As previously mentioned, a significant standard deviation of the estimated migrants' effects may be taken as a sign of unexplained heterogeneous effects. A simple way to explore the role of investment-level variables in explaining such heterogeneity is to interact them with our determinants of interest. Chung and Alcácer (2002) implement this approach by exploiting heterogeneity across industries.

In practice, this implies augmenting our specification with a set of interaction effects of our variables of interest with investment-level factors. These interactions express how different functions move the effects of the variable of interest away from the mean coefficient estimated for the reference category.

This approach is well suited to study how a limited set of investment-level variables drive the effect heterogeneity. It is less suited to accommodate a large number of heterogeneity drivers. Indeed, extending this approach to the full set of potential drivers of migrants' effect heterogeneity outlined in Section 2 may raise substantial collinearity issues, which are a significant hindrance to the convergence of these models and to the identification of the effects of interest. When the number of potential drivers is large and the researcher is interested in their conditional effects, as in our case, an appealing option is to compute firm-specific parameters and regress them on the vector of potential drivers. In turn, a limitation of the two-stage approach is that, in the presence of repeated choices by the same firm, we can derive firm-specific, but not investment-specific parameters (see Section 3.1.1 and Train, 2009).

As both the interactions and two-stage approaches have pros and cons, in what follows, we combine the two methods. We study the impact of functional heterogeneity via interaction effects, and the impact of a broader set of firm-level drivers with the two-stage regression-based approach we detail in Section 3.2.1.

Specifically, we interact our migration variables with six dummies capturing the main MNE activities in our data set. To reflect our hypotheses about the differential role of migration for knowledge-seeking, production-related, and downstream investments, we leverage the information available in *fDi Markets* about the business activities of the investments and consider six categories: R&D, manufacturing, construction, logistics, sales and marketing, and business services. We proxy knowledge-seeking investments with R&D ventures, which we expect to reflect migrants' skill-specific, knowledge-intensive labor, and information effects. By contrast, we expect the low-cost labor cost channel to be especially important for production-related, that is, manufacturing, construction, and logistics ventures. Finally, we consider investments in sales and marketing, and business services as downstream activities, expecting a stronger role of demand and information-related factors.

Following Chung and Alcácer (2002), we include these interaction effects with fixed parameters. This implies that we will not explore within-function heterogeneity in the immigrant effects, which would be hindered by the limited number of investments available in each function.

3.2 | Second stage: Computing and explaining firm-level individual coefficients

3.2.1 | Econometric approach

To study how firm-level factors may explain the heterogeneity in our variables of interest, we resort to the simulation procedure developed by Revelt and Train (2000) and Train (2009), and compute the firm-specific parameters δ_f based on the parameters of their distributions estimated in the first stage, as well as their variances σ_{δ_f}. We focus on the firm-specific effects of bilateral immigrants and emigrants on MNE activities' location choice.[10] The vector of the individual coefficients $\hat{\delta}_f$ can be employed as dependent variables in a second-stage regression that explores the investor-level factors that may increase or decrease the relevance of the parameters of interest for the location of MNE activities (Castellani & Lavoratori, 2020; Hornstein & Greene, 2012; Saxonhouse, 1976). The estimated firm-specific coefficients are time-invariant, so the regression is a simple cross-sectional regression. The firm characteristics that we employ in the second stage are computed based on the entire time span of our availability in *fDi Markets*. The second-stage equation is the following:

$$\hat{\delta}_f = \theta' X_f + u_f, \tag{4}$$

where X_f is a vector of firm, industry, and home-country level characteristics, including a constant, which we detail in Section 3.2.2. Positive (negative) coefficients can be interpreted as factors that are associated with a greater (smaller) than average effect of migration. Because the random coefficients are firm-specific, variables at the investing company level that could not be included in the first stage can enter this second stage as regressors.[11]

3.2.2 | Firm-level heterogeneity drivers

The 1113 inward FDI projects in our data are the result of choices made by 895 investing companies. As discussed in Section 2, firm characteristics may reflect the relative importance of different mechanisms. We compute their characteristics based on all their worldwide investments in the *fDi Markets* database over the entire 2003–2015 period.

[10]For details about the estimation of individual effects, see Supporting Information Appendix B.2.

[11]The variance of the estimated firm-specific parameters is firm-specific, which makes the error structure of the second-stage model heteroskedastic and OLS an inefficient estimator for the second stage (Hornstein & Greene, 2012; Saxonhouse, 1976). Hence, we apply a feasible generalized least-squares estimator and weight each observation on all variables used in the second stage by the inverse of the estimated standard error of the dependent variable $\hat{\sigma}_{\delta_f}$. Arne Risa Hole provided valuable guidance in estimating $\hat{\sigma}_{\delta_f}$.

To capture the relevance of labor costs, we compute the average number of jobs per million dollars invested as a proxy for labor intensity. As for demand, we expect stronger effects of migration for final goods and services if this channel prevails. Hence, we distinguish between the types of goods produced by leveraging the information about the industry sectors, subsectors, and clusters of the firm provided in *fDi Markets*, grouping them into four categories: final goods, services, intermediate goods, other (i.e., they could be both final and intermediate). Third, we expect the information effects to be more salient for firms with less international experience, organizational capacity, and knowledge of the destination economy. To proxy for organizational capacity we compute the total amount invested worldwide by each of the firms in our sample, the average size of the typical investment, as well as the number of activities and of different countries worldwide where the firm has at least one FDI. To proxy for firm investment experience, we compute the years of experience at its first investment in Italy, that is, the number of years since the company's first appearance in *fDi Markets*.[12] To proxy for firm experience in Italy, we compute Italy's share over total investments, and the number of investments in Italy. We also dichotomize the latter variable into a dummy equal to one if the firm has more than one investment in Italy. Finally, we include a set of geographical area dummies for the investor's country of origin, expecting stronger effects for more culturally distant areas, like Southeast Asia.

3.3 | Endogeneity

To attribute a causal interpretation to the effect of migrants on location choices requires a thorough consideration of endogeneity issues in the first-stage model. Unobserved characteristics of the location, including special ties with the origin country of the investment, may affect migration and location choice at the same time, leading to a biased estimated effect of migration. On the other hand, reverse causality may concern both immigration and emigration: MNE activities open job opportunities that may be particularly appealing for workers in their origin country and drive immigration to the destination; they may create new ties between the origin and the destination that lead to emigration, or they may cause internal migration flows between neighboring, or even distant, regions.

To capture time-invariant unobserved effects, both at the local and at the province-country dyad level, which may correlate with both migration and investment location decisions, we include province and province-country dummies in separate specifications.

3.3.1 | A control function approach

To address reverse causality and time-varying omitted variable bias, we follow standard practice in the literature and employ an Altonji–Card shift-share instrument (Altonji & Card, 1991; Card, 2001), which has been widely applied to study bilateral migration (e.g., Bratti et al., 2014, 2020; Ottaviano & Peri, 2006; Peri & Requena-Silvente, 2010). This instrument exploits the path dependence in immigrant settlements to identify an exogenous source of variation in immigration. In our application, we instrument current stocks based on the country-specific shares of immigrants by province in 1995 (shares) and the yearly Italy-wide growth of country stocks (the shifters) (similar to Bratti et al., 2014). A similar instrument is employed for emigration (see Section A.2 in the Supporting Information Appendix for more details and a discussion of instrument validity). To instrument multilateral immigration and emigration, we aggregate by province the imputed bilateral migration variables (subtracting bilateral migration).

[12]As the first year in the *fDi Markets* data set is 2003, we assume 2003 as the earliest year of investment.

We employ our immigration and emigration instruments within a control function approach. Heckman and Robb (1985) provide a seminal paper on control functions, while Petrin and Train (2010) and Train (2009, Chap. 13) generalize this approach to choice models with random coefficients.

The potentially endogenous migration variable y_{fnj} can be expressed as a function of exogenous variables and unobserved factors:

$$y_{fnj} = W(w_{fnj}, \gamma) + \mu_{fnj}, \tag{5}$$

where w_{fnj} includes the instruments for y_{fnj} described above and the full set of exogenous regressors described in Section 3.1.2. γ is the vector of their estimated coefficients. If μ_{fnj} in Equation (5) and ϵ_{fnj} in Equation (1) are correlated, y_{fnj} will be endogenous. The issue can be addressed by decomposing ϵ_{fnj} into its mean conditional on μ_{fnj} and deviations around this mean: $\epsilon_{fnj} = E(\epsilon_{fnj}|\mu_{fnj}) + \tilde{\epsilon}_{fnj}$. By construction, the deviations $\tilde{\epsilon}_{fnj}$ from the mean will be uncorrelated with μ_{fnj}, hence with y_{fnj}. The part of ϵ_{fnj} that is correlated with y_{fnj} can be entered explicitly as an additional explanatory variable, such that the remaining part is not correlated. The conditional expectation is a function of μ_{fnj} and is known as the control function. A choice model that includes the control function as an additional regressor along with the standard deviation of the conditional errors can be estimated by mixed logit.[13]

Intuitively, regressing immigration on the instrument and on covariates, we separate the component of province-level immigration that is due to immigrants' exogenous effects from the one that may correlate with unobserved province characteristics. The residuals from the immigration regression capture the latter correlated error component, which can be included as a control function in the main regression. In this way, we control for the endogenous component in immigration, so that the estimated immigrants' coefficient captures their exogenous effect.[14] Similar considerations apply to emigration. Testing the significance of the control function coefficient is a way to test for endogeneity (Wooldridge, 2015).

4 | DESCRIPTIVE STATISTICS

4.1 | First-stage descriptive statistics

Table 1 reports the summary statistics and Table 2 shows the correlation matrix for our variables. Parent colocation, firm density, bilateral FDI stock and GDP, and the dummy for Rome and Milan display the highest correlations with the dependent variable *Choice*.

Furthermore, *Choice* is positively correlated with the log of multilateral immigrants, suggesting that the determinants of the location choice of FDI may be similar to those for the location choice of immigrants. Correlation patterns support the expectation that bilateral immigration, too, reacts to the structure of opportunities prevailing locally, as immigration is highly correlated with province GDP, population, patent count, institutional quality, and wages, while it is negatively correlated with the province-specific unemployment rates. Bilateral emigration is positively correlated with immigration, which along with the positive signs of the correlations with province GDP, patent count, imports, and exports suggest that the decision to expatriate is related to the openness of the local context, while it also constitutes a reaction to

[13]The control function approach yields a mixed logit model if we assume that the error term ϵ_{fnj} is composed of a normal component ϵ^1_{fnj} with mean μ_{fnj} and constant variance $\sigma\eta_{fnj}$ (where η_{fnj}, is standard normal), and an iid extreme value component ϵ^2_{fnj}, so that utility can be written as $U_{fnj} = V(x_{fnj}, z_{fnj}, y_{fnj}, \beta, \delta_f) + \lambda\mu_{fnj} + \sigma\eta_{fnj} + \epsilon^2_{fnj}$ (Petrin & Train, 2010; Train, 2009).

[14]The error component is included with a random parameter to capture the impact of both its mean and variance on investors' utility (see note 13), which is also a natural approach considering our focus on the heterogeneity in migrant's effects.

TABLE 1 Summary statistics.

Variable	Mean	SD	Minimum	Maximum
Choice	0.010	0.099	0.000	1.000
Log immigrants	4.219	1.586	0.000	11.725
Log emigrants	5.993	2.364	0.000	11.017
Parent colocation	0.009	0.095	0.000	1.000
Unemployment rate	8.065	4.732	1.855	26.882
Infrastructure endowment	97.145	65.134	9.505	522.210
Residents with tertiary education	0.069	0.972	−2.435	3.468
Log prov. population	12.942	0.719	10.956	15.279
Institutional quality	0.576	0.230	0.000	1.000
Common border	0.359	0.480	0.000	1.000
Log prov. GDP	9.084	0.809	6.664	11.868
Sectoral diversity	0.048	0.947	−4.522	1.832
Log patent count	2.835	1.442	0.000	6.402
Log average wage (region)	9.738	0.130	9.410	10.003
Log distance	7.754	1.082	4.545	9.840
Agglomeration (sector)	1.021	0.824	0.019	45.786
Log imports	16.960	2.323	1.946	23.680
Log exports	17.568	2.235	2.197	22.326
Pre-2002 FDI stock	2.006	8.372	0.000	125.000
Log multilateral immigrants	9.783	1.097	6.078	13.163
Log multilateral emigrants	9.967	1.223	0.000	12.630
Firm density	4.829	10.655	0.201	123.973
Manufacturing concentration	1.040	0.365	0.350	2.006
Rome/Milan	0.020	0.140	0.000	1.000

Note: Summary statistics on the estimation sample. Observations: 111,692. The estimation sample is composed of a set of 107 possible locations for all 1113 investments in the sample. For details on the database construction and data cleaning process, see Supporting Information Appendix A.

Abbreviations: FDI, foreign direct investment; GDP, gross domestic product.

unemployment (with which the correlation is positive and quite high). Coherent with these arguments, the dummy for Rome and Milan is positively correlated with both immigration and emigration, but more strongly with immigration. Finally, geography matters for both immigrants and emigrants, as shown by the negative correlation of both variables with distance, but the correlation coefficient is larger for emigrants, suggesting that immigrants travel longer distances than emigrants.

In Figure 1 we report the distribution of Italian inward FDI by province. The subnational heterogeneity in the distribution of FDI is striking, with the vast majority of FDI being directed to the province of Milan. The provinces of Rome, Turin, Bologna, Genova, Florence, Verona, and Naples are also comparatively important attraction poles. Some geographical clustering of the investments in northern provinces can be identified, while it seems almost

TABLE 2 Correlation matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. Choice	1.00																							
2. Log immigrants	0.14	1.00																						
3. Log emigrants	0.06	0.38	1.00																					
4. Parent colocation	0.52	0.11	0.05	1.00																				
5. Unemployment rate	−0.04	−0.18	0.17	−0.03	1.00																			
6. Infrastructure endowment	0.08	0.23	0.06	0.07	−0.18	1.00																		
7. Residents with tertiary education	0.18	0.32	0.12	0.15	−0.23	0.43	1.00																	
8. Log prov. population	0.20	0.47	0.29	0.16	0.04	0.24	0.35	1.00																
9. Institutional quality	0.09	0.31	−0.12	0.07	−0.78	0.38	0.34	0.12	1.00															
10. Common border	0.06	0.11	−0.05	0.04	−0.50	0.18	−0.15	0.04	0.43	1.00														
11. Log prov. GDP	0.22	0.52	0.24	0.18	−0.22	0.31	0.42	0.95	0.38	0.21	1.00													
12. Sectoral diversity	−0.02	−0.09	0.06	−0.01	0.11	0.01	0.15	−0.07	−0.17	−0.03	−0.12	1.00												
13. Log patent count	0.16	0.45	0.07	0.13	−0.59	0.32	0.39	0.63	0.71	0.44	0.80	−0.17	1.00											
14. Log average wage (region)	0.07	0.25	−0.05	0.05	−0.55	0.21	0.10	0.04	0.69	0.56	0.29	−0.15	0.54	1.00										
15. Log distance	−0.01	−0.14	−0.21	−0.01	0.13	−0.02	−0.01	0.01	−0.12	−0.11	−0.03	0.04	−0.09	−0.11	1.00									

TABLE 2 (Continued)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
16. Agglomeration (sector)	0.05	0.00	−0.01	0.04	−0.09	0.06	0.05	−0.01	0.07	0.10	0.03	0.06	0.06	0.06	−0.03	1.00								
17. Log imports	0.13	0.60	0.43	0.10	−0.34	0.26	0.26	0.46	0.45	0.26	0.56	−0.05	0.58	0.34	−0.26	0.07	1.00							
18. Log exports	0.10	0.54	0.46	0.08	−0.44	0.20	0.24	0.44	0.56	0.30	0.57	−0.06	0.65	0.42	−0.17	0.05	0.78	1.00						
19. Pre-2002 FDI stock	0.33	0.27	0.19	0.21	−0.12	0.10	0.27	0.35	0.20	0.17	0.40	−0.05	0.33	0.17	0.04	0.05	0.30	0.29	1.00					
20. Log multilateral immigrants	0.17	0.51	0.14	0.14	−0.43	0.30	0.41	0.72	0.59	0.31	0.86	−0.22	0.83	0.61	−0.07	0.03	0.55	0.60	0.34	1.00				
21. Log multilateral emigrants	0.08	0.12	0.36	0.07	0.32	0.09	0.20	0.56	−0.29	−0.06	0.43	0.24	0.13	−0.20	0.09	−0.00	0.09	0.08	0.12	0.22	1.00			
22. Firm density	0.37	0.32	0.13	0.26	−0.10	0.27	0.34	0.51	0.24	0.16	0.55	−0.20	0.41	0.22	−0.03	0.03	0.33	0.28	0.63	0.47	0.17	1.00		
23. Manufacturing concentration	−0.05	0.03	−0.08	−0.04	−0.48	−0.17	−0.14	−0.03	0.45	0.32	0.10	−0.03	0.43	0.37	−0.07	0.03	0.23	0.38	0.01	0.28	−0.10	−0.00	1.00	
24. Rome/Milan	0.37	0.29	0.13	0.27	−0.05	0.17	0.45	0.44	0.15	0.04	0.47	−0.04	0.29	0.13	−0.01	0.03	0.25	0.17	0.56	0.37	0.22	0.77	−0.18	1.00

Note: Correlation matrix on the estimation sample. Observations: 111,692. The estimation sample is composed of a set of 107 possible locations for all 1113 investments in the sample. For details on the database construction and data cleaning process, see Supporting Information Appendix A.

Abbreviations: FDI, foreign direct investment; GDP, gross domestic product.

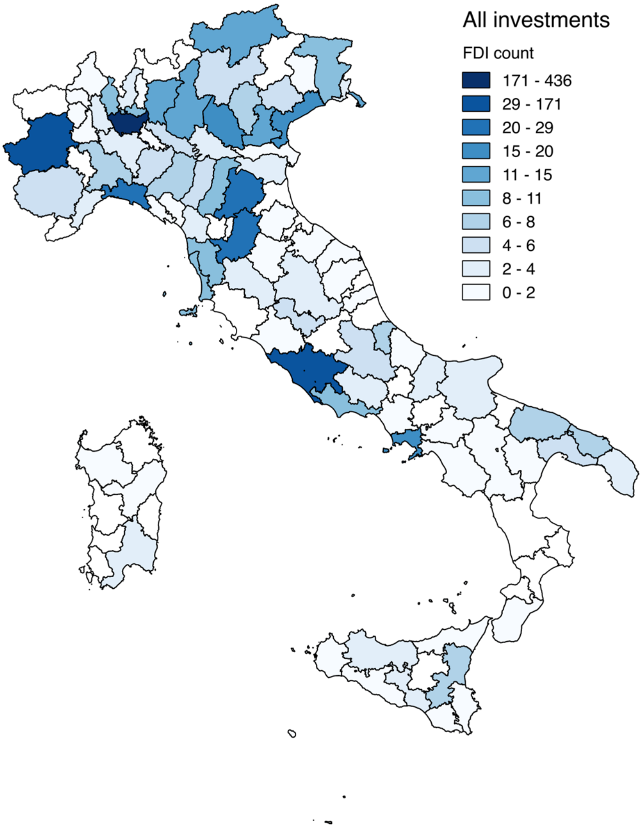

FIGURE 1 FDI distribution by province. Total number of inward FDI projects by province over the 2003–2015 period. *Source*: Own elaborations on *fDi Markets* data. FDI, foreign direct investment. [Color figure can be viewed at wileyonlinelibrary.com]

absent in southern provinces. As for specific functions, the geographical distribution for R&D, market access, and business services investments is similar to the aggregate distribution, while it is more dispersed for manufacturing, logistics, and construction (see figures and further descriptive statistics in Supporting Information Appendix C).

4.2 | Second-stage descriptive statistics

Table 3 reports some summary statistics referring to the characteristics of the 895 investing companies in our sample, retrieved from *fDi Markets*.

Over 2003–2015, the typical company in our data set had invested a total of 913.81 million US$ worldwide in different countries. The size of the average investment of these companies amounts to 44.4 million US$, the typical company in our data set engaging in about 21 FDIs over the considered period. This distribution is highly right-skewed, with a minority of companies making large investments. The number of countries in which these firms invest, which we take as a proxy of their coordination capacity, ranges from 1 to 76, with an average of 6.9. On average, they create 4.52 jobs per million US$ investment, again with substantial heterogeneity across firms.

17% of the firms in our sample invested in Italy more than once, with a maximum of 13 investments. On average, FDI to Italy makes up about 39% of their total worldwide investment portfolio and, at their first

TABLE 3 Summary statistics—Investing companies.

Variable	Mean	SD	Minimum	Maximum
Average investment size (mln US$)	44.40	116.25	0.20	2561.00
Worldwide capital investment (mln US$)	913.81	3049.80	0.20	33,426.85
Italy share over total investments	0.39	0.39	0.00	1.00
Number of investments in Italy	1.34	1.04	1.00	13.00
Dummy: more than one investment in Italy	0.17	0.38	0.00	1.00
Number of countries worldwide	6.90	8.74	1.00	76.00
Years of experience at first investment in Italy	1.62	2.33	0.00	11.84
Number of activities worldwide	2.16	1.98	1.00	15.00
Jobs/mln US$ invested	4.52	8.06	0.04	166.67
Sectoral category dummies				
Services	0.59		0.00	1.00
Final goods	0.05		0.00	1.00
Intermediate goods	0.24		0.00	1.00
Other	0.12		0.00	1.00
World region dummies				
EU	0.55		0.00	1.00
South and East Asia	0.09		0.00	1.00
Non-EU Europe	0.06		0.00	1.00
North America	0.28		0.00	1.00
Rest of the world	0.03		0.00	1.00

Note: Summary statistics about the characteristics of the investing companies engaged in the 1113 investments in the estimation sample over 2003–2015. Observations: 895.
Source: *fDi Markets* data on the global investments about the investing companies engaged in foreign direct investment into Italy.

investment in Italy, they have 1.62 years of international experience. The majority of firms in our sample (59%) are operating in the services sector. 5% of firms produce final goods, 24% intermediate goods, and 12% of them produce goods that can be considered in both categories. Regarding origin areas, 55% of the investing companies come from the EU, 28% from North America, 9% from Southeast Asia, 6% from non-EU Europe, and 3% from the rest of the world.

The correlation matrix of our investing company variables, reported in Table 4, shows high correlations between the total worldwide capital investment of a firm and the number of countries it invests in worldwide, as well as the years of experience the MNE has at the time of its first investment in Italy and the number of different MNE activities in which it is active globally. Hence, firms with greater managerial capacity may have a larger and more diversified portfolio of international investments. Italy's share of total investments—as well as the sheer number of investments in Italy—are also larger for firms with a more diversified portfolio in terms of countries and activities. Some further descriptive statistics about the countries of origin of the FDIs, the functions of the MNE activities, and the breakdown of our migration variables by country and province are reported in Supporting Information Appendix C (Tables C.1–C.3).

TABLE 4 Correlation matrix—Investing companies.

		1	2	3	4	5	6	7	8	9
1	Average investment size (mln US$)	1.00								
2	Worldwide capital investment (mln US$)	0.34	1.00							
3	Italy share over total investments	−0.05	−0.25	1.00						
4	Number of investments in Italy	0.08	0.29	−0.06	1.00					
5	Dummy: more than one investment in Italy	0.11	0.29	−0.06	0.71	1.00				
6	Number of countries worldwide	0.10	0.64	−0.52	0.29	0.28	1.00			
7	Years of experience at first investment in Italy	0.04	0.16	−0.48	−0.01	−0.01	0.37	1.00		
8	Number of activities worldwide	0.16	0.60	−0.45	0.21	0.27	0.78	0.36	1.00	
9	Jobs/mln US$ invested	−0.11	−0.07	0.05	−0.07	−0.08	−0.05	−0.03	−0.03	1.00

Note: Correlation matrix about the characteristics of the investing companies engaged in the 1113 investments in the estimation sample over 2003–2015. Observations: 895.

Source: *fDi Markets* data on the global investments about the investing companies engaged in foreign direct investment into Italy.

5 | RESULTS

5.1 | The effect of migration on location choice

In Table 5, we study the determinants of location choice and their heterogeneity across investing firms.

In Model 1, we report the estimated coefficients of our logit model including the full set of regressors, distinguishing between means in column (1) and standard deviations in column (2).[15] As regards the means, the estimates indicate a positive and significant effect of bilateral immigration but no significant effect of bilateral emigration, nor of multilateral immigration and emigration. Moreover, colocation, agglomeration, imports, and pre-2002 FDI stocks confirm their expected role as significant attraction factors for FDI.[16]

As for the standard deviations, the significant likelihood ratio test on the joint significance of the standard deviations (reported in the last three lines of the table) allows us to strongly reject the null hypothesis of fixed coefficients and supports the expectation of parameter heterogeneity. As discussed, a significant standard deviation can be viewed as evidence of heterogeneity in the investors' preferences for a given location determinant. Significant standard deviations are estimated for immigration, distance, pre-2002 FDI stocks, firm density, and the Rome/Milan dummy.[17] Our estimates imply that, while heterogeneous, the effect of immigrants is positive for a majority of the companies (about 86%, i.e., $100 \times \Phi(0.367/0.335)$), an issue that we shall further explore in what follows.[18] None of the other migration variables turn out to have significantly heterogeneous effects.

[15]The sign of the estimated standard deviations is irrelevant and should be interpreted to be positive (Hole, 2007).

[16]Colocation acts as a sort of lagged dependent variable, and its inclusion erodes the effect of other relevant location determinants, such as province GDP, distance, the share of tertiary educated, manufacturing concentration, and the dummy for Rome and Milan. A limitation of this variable is that, being based on the same *fDi Markets* data set we employ in our analysis, it is unavailable before 2003. Hence, we make the implicit assumption that none of the firms in our sample invested in Italy before 2003, which may be viewed as a strong assumption—especially at the beginning of the period—so that the variable is likely measured with error. Reassuringly, our main results are robust to removing colocation from the analysis (see Table C.9 in Supporting Information Appendix C) and excluding the first 4 years from the analysis to reduce measurement error (see Supporting Information Appendix C, Table C.10). The importance of this variable from both a theoretical and an empirical point of view, as the results will show, leads us to include it in our analysis despite its limitations.

[17]We obtain very similar results when we treat all the factors whose estimated standard deviation turns out to be insignificant as fixed parameter variables. Results are available upon request.

[18]This approximation of the expectation of $\Phi(\hat{\bar{\delta}}_f/\hat{\sigma_\delta})$ follows Hole (2007).

TABLE 5 Estimation results—Mixed logit.

	Model 1 Baseline		Model 2 Province dummies		Model 3 Control function		Model 4 Control function	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log immigrants	0.367***	0.335***	0.485***	0.335***	0.403***	0.348***	0.411***	0.352***
	(0.099)	(0.086)	(0.120)	(0.101)	(0.110)	(0.084)	(0.113)	(0.088)
Log emigrants	0.085	0.010	0.135	−0.013	0.070	−0.000	−0.063	−0.002
	(0.077)	(0.123)	(0.086)	(0.181)	(0.078)	(0.122)	(0.151)	(0.119)
Parent colocation	5.299***		5.271***		5.298***		5.312***	
	(0.149)		(0.154)		(0.149)		(0.152)	
Unemployment rate	−0.035		−0.005		−0.037		−0.033	
	(0.036)		(0.049)		(0.036)		(0.036)	
Infrastructure endowment	0.001				0.001		0.001	
	(0.001)				(0.001)		(0.001)	
Residents with tertiary education	0.117	−0.028			0.113		0.139	
	(0.089)	(0.125)			(0.089)		(0.092)	
Log prov. population	0.192	0.062			0.223	0.017	0.230	−0.003
	(0.513)	(0.159)			(0.512)	(0.151)	(0.512)	(0.161)
Institutional quality	−0.041	−0.169			−0.025	0.171	−0.195	−0.189
	(0.670)	(0.804)			(0.670)	(0.761)	(0.686)	(0.679)
Common border	0.122	−0.127			0.138	−0.264	0.139	0.115
	(0.174)	(0.408)			(0.175)	(0.347)	(0.174)	(0.472)
Log prov. GDP	0.578	−0.012	1.846	−0.025	0.592	0.014	0.587	−0.018
	(0.483)	(0.194)	(1.219)	(0.181)	(0.484)	(0.154)	(0.484)	(0.167)
Sectoral diversity	0.004	−0.033	0.726*	−0.058	0.014	0.036	0.008	
	(0.078)	(0.135)	(0.385)	(0.175)	(0.079)	(0.139)	(0.080)	
Log patent count	0.063	−0.156	0.050	−0.031	0.045	−0.118	0.032	0.027
	(0.140)	(0.099)	(0.234)	(0.159)	(0.142)	(0.135)	(0.138)	(0.186)
Log average wage (region)	0.264	−0.252	8.487	−0.139	0.290	0.302	0.392	−0.199
	(1.461)	(1.856)	(7.306)	(2.362)	(1.461)	(1.826)	(1.470)	(1.966)
Log distance	−0.133	0.976**	−0.247	0.729	−0.126	0.972**	−0.144	1.016**
	(0.291)	(0.433)	(0.305)	(0.500)	(0.295)	(0.447)	(0.297)	(0.452)
Agglomeration (sector)	0.184***	−0.006	0.226***	0.077	0.186***	0.002	0.187***	0.026
	(0.030)	(0.083)	(0.039)	(0.076)	(0.030)	(0.072)	(0.031)	(0.073)
Log imports	0.172**	0.112	0.082	−0.116	0.168**	0.134	0.167**	0.093
	(0.073)	(0.125)	(0.087)	(0.165)	(0.071)	(0.086)	(0.071)	(0.116)
Log exports	0.078	0.017	0.001	−0.020	0.080	0.038	0.102	−0.033
	(0.073)	(0.068)	(0.086)	(0.100)	(0.074)	(0.084)	(0.077)	(0.082)

(Continues)

TABLE 5 (Continued)

	Model 1 Baseline		Model 2 Province dummies		Model 3 Control function		Model 4 Control function	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pre-2002 FDI stock	0.010***	0.012**	0.010***	−0.001	0.010***	0.010*	0.009***	0.011**
	(0.003)	(0.005)	(0.003)	(0.009)	(0.003)	(0.006)	(0.003)	(0.005)
Log multilateral immigrants	−0.254	0.103	0.556	−0.072	−0.289	0.082	−0.256	−0.214
	(0.228)	(0.231)	(0.541)	(0.209)	(0.229)	(0.250)	(0.244)	(0.167)
Log multilateral emigrants	−0.148	−0.041	0.411	0.046	−0.139	0.023	−0.071	0.002
	(0.113)	(0.114)	(0.599)	(0.217)	(0.113)	(0.112)	(0.133)	(0.121)
Firm density	−0.000	0.011***	−0.008	−0.004	−0.001	−0.013***	0.001	−0.012**
	(0.004)	(0.004)	(0.007)	(0.008)	(0.004)	(0.004)	(0.004)	(0.006)
Manufacturing concentration	−0.402	0.013	−3.593*	−0.127	−0.383	0.049	−0.413	0.088
	(0.265)	(0.289)	(2.125)	(0.358)	(0.267)	(0.317)	(0.269)	(0.320)
Rome/Milan	−0.271	1.059***			−0.292	0.937**	−0.382	−0.960**
	(0.425)	(0.303)			(0.426)	(0.392)	(0.436)	(0.441)
Residual (immigrants)					−0.081	−0.016	−0.076	0.017
					(0.115)	(0.159)	(0.115)	(0.162)
Residual (emigrants)							0.148	0.168
							(0.137)	(0.202)
Province dummies	No		Yes		No		No	
Observations	111,692		112,894		111,692		111,692	
AIC	4133.152		4229.098		4136.519		4138.723	
BIC	4546.963		5365.934		4559.953		4571.781	
LR test of joint significance of the SD	115.775		110.235		114.282		113.713	
Degrees of freedom	20		17		20		20	
Test *p*-value	0.000		0.000		0.000		0.000	

Note: Location-choice models for inward FDI targeting Italian provinces. Mixed logit estimates based on 500 Halton draws each. For each model, the table reports the estimated average effect of the variable (Mean) and standard deviation (SD). Model 1 is the baseline estimate. Model 2 includes province dummies, hence excludes collinear time-invariant and variables with limited time variation. This leads to excluding Log prov. population, whose values are missing for some provinces, hence Model 2 has a higher number of observations and underlying firms. For coherence with the other models, in Model 2 we include the dummies for Rome and Milan as random parameters. The interpretation of their coefficient is, however, different, given the different reference category, hence they are not displayed. Model 3 employs the control function including the first-stage residuals for immigrants. Model 4 includes the first-stage residual for both immigrants and emigrants. Due to missing data on immigrants in 1995, the instrument can only be computed for a subset of countries, hence Models 3 and 4 have a smaller number of observations (see Supporting Information Appendix A). Standard errors take into account the correlation among the investments by the same company but are not bootstrapped (see note 22). Standard errors in parentheses.

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; FDI, foreign direct investment; GDP, gross domestic product; LR test, likelihood-ratio test.

**p* < 0.10; ***p* < 0.05; ****p* < 0.01.

Functional, sectoral, and knowledge-base heterogeneity may also explain the significant standard deviations in FDI stocks, whose effects are estimated to be positive for 80% of firms. Larger FDI stocks may signal preferential bilateral ties between the country of origin of the investment and the destination province; heterogeneous bilateral information costs may drive their significant standard deviation. The heterogeneous effects of firm density may instead be attributed to the heterogeneity in MNE strategies. On average, the estimated effect of this type of urbanization economies, conditional on all other regressors, is zero. However, its significant standard deviation indicates that about 50% of firms consider information and urban externalities as positive location factors, while another 50% would refrain from exploiting them, possibly to enjoy first-mover advantages or to locate large and space-consuming establishments. Similarly, the effect of distance is estimated to be negative, as expected, for about two-thirds of the firms, but it is actually positive for the remaining third. This finding is consistent with the results of Castellani et al. (2013), who find that specific types of FDI—namely, R&D investments—travel longer distances. Finally, as for the dummy for Rome and Milan, the mean coefficient is negative but the standard deviation estimated is significant and relatively large. This indicates that, consistent with their distinctive role as the economic and administrative capitals of Italy, about 40% of investors still consider these cities preferential locations for their investments, even conditionally on all other covariates.[19]

We also address potential sources of endogeneity. In Model 2 (columns 3 and 4), we study whether omitted time-invariant province-level factors drive the results by including a dummy for each of the provinces chosen at least once. To avoid collinearity issues, we exclude province-level time-invariant regressors (residents with tertiary education, institutional quality, common border) as well as population, which is time-varying but displays limited time variation in the period under scrutiny and would hamper convergence. For coherence with the previous models, we include the dummies for Rome and Milan as random parameters.[20] As columns (3) and (4) show, the mean coefficient of bilateral immigrants increases in magnitude and in statistical significance and the standard deviation is remarkably stable. As for the other regressors, the signs of the mean coefficients remain stable, although the magnitudes are in some cases quite different. Quite interestingly, the inclusion of the province dummies appears to reduce the heterogeneity of some coefficients; yet, overall, our baseline results on the effects of migrants do not seem to be affected by province-level unobserved heterogeneity. Bilateral emigration and multilateral immigration and emigration remain insignificant and insignificantly heterogeneous.

The main results are unaffected when we implement the control function approach described in Section 3.3 (Models 3 and 4, columns 5–8). In Model 3, we address the potential endogeneity in the province-level bilateral immigrant stocks. As discussed, regressing *Log immigrants* on its Altonji-Card instrument and on the other covariates, we separate the exogenous pull effects of province-level immigration (captured by *Log immigrants*) from the component that may correlate with unobserved province characteristics (captured by the control function).[21]

Our results indicate that the coefficient of the control function is far from being statistically significant, while the effect of *Log immigrants* remains positive, significant, and significantly heterogeneous.[22] Hence, conditional on all other covariates, we cannot reject the null hypothesis of exogeneity, and conclude that the previously detected

[19]Including the dummy for Rome and Milan with a random parameter in a mixed logit model is equivalent to allowing for a nest in a nested logit model, that is, a grouping of alternatives within which error terms are correlated (Train, 2009). This implies recognizing the high subnational concentration of investments and allowing a foreign firm to face a two-step decision: first, whether to invest in either Rome or Milan or elsewhere in Italy; second, in which specific province to invest. The results of the estimates omitting the Rome/Milan dummy are robust and available upon request.

[20]The interpretation of the coefficient is, however, different, given the different reference categories. Hence, we do not display the results to save space and avoid confusion.

[21]The past settlement instrument turns out to fairly accurately predict observed immigration, with a highly significant *F* statistic ($p < 0.0001$), and an R^2 of 0.78. This is in line with the common finding that the past settlement instrument is quite strong.

[22]The standard errors of these models should be bootstrapped to take into account that the control function is estimated (Train, 2009). In our experience, bootstrapping the standard errors of mixed logit models turned out to be so time-consuming to become practically unfeasible. Bootstrapping the standard errors in Model 3 of Table 5 required 25 days of computation (results are available upon request). Bootstrapping the standard errors in Model 4, which includes only one more random coefficient compared to Model 3, turned out to exceed the interval between the monthly maintenance services on the servers that we employed to gain computational speed. For the same reasons, we could not obtain the bootstrapped estimates of Models 5 and 6 in Table 6. Hence, in what follows we present the uncorrected standard errors, that is, standard errors that only take the correlation among the investments of the same company into account. Instead, we will bootstrap the estimates in Model 7 in Table 6 based on the computationally simpler conditional logit.

effect of immigration is attributable to its exogenous effects. Moreover, the insignificant standard deviation in the correlated error component implies that there is no evidence of heterogeneous effects in the endogenous component. Accordingly, Model 3 estimates are very similar to those in Model 1 and so is the interpretation of its main results.

Similar insights are obtained from Model 4, which is based on two preliminary regressions. One regresses bilateral immigrants on their shift-share instrument, on bilateral emigrants' shift-share instrument, and on other covariates; the other regresses emigrants on their instrument, on the immigrants' instrument, and on other covariates.[23] Neither of the residuals is significant, nor significantly heterogeneous, and results for the other regressors are remarkably robust. The inference for the other migration variables is also unaffected. Yet, a comparison of the different estimates indicates that the endogeneity correction provided by the control function operates in opposite ways. Endogeneity appears to induce an upward bias for emigrants and a downward bias for immigrants, making our baseline estimates conservative. In Supporting Information Appendix C.8, we show that no additional insights are gained from including two additional control functions from multilateral immigration and emigration. Hence, we continue with the more parsimonious specification including only the control functions of bilateral migration.[24]

Overall, our first-stage results point to a positive, significant, and significantly heterogeneous exogenous effect of bilateral immigration. Instead, we are unable to detect any significant effects of bilateral emigration nor of multilateral immigration and emigration. Hence, our results do not support any information effects by bilateral emigrants. Neither do they support the low-cost labor channel, proxied by the effect of multilateral immigration, nor any impact of multilateral emigration on the global attractiveness of provinces.

Supporting Information Appendix Section C.2 reports some robustness checks. These confirm that bilateral immigration effects are robust to excluding Rome and Milan from the sample, they are stronger for skilled immigrants and they appear to be the result of a very localized knowledge exchange. They are also robust to including bilateral province-country dummies in our control function specification.[25]

Given that the bilateral immigration effect may be the combined result of different effects (information, demand, and country-specific labor supply effects), we then move to exploring the sources of the heterogeneity in its effects. Clearly, a corresponding analysis of the effects of bilateral emigration and multilateral migration would be pointless due to the insignificant mean and standard deviation of their coefficients.

5.2 | Sources of heterogeneity in the immigrant effects

5.2.1 | Heterogeneity by business function of FDI

The robust heterogeneity detected in the immigrant effects may be due to different factors. One possibility is that the immigrant effect is heterogeneous by function, and one way to test this hypothesis is to augment the control

[23]Also in the case of emigration, the instrument is strong and accurately predicts observed emigration ($p < 0.0001$; $R^2 = 0.81$). These results and those described in the previous note are available upon request.

[24]In the interest of brevity, in what follows we do not report the coefficients for multilateral immigration and emigration, which remain insignificant throughout estimates. Complete results are available upon request.

[25]It may be that the effects of immigration are heterogeneous by alternatives, instead of by decisionmaker. This would imply that immigration has different effects in different provinces. Empirically, our two-step approach is not the right way to study this possibility, given that firm-specific parameters only allow us to explore decisionmaker-level —not alternative-level— heterogeneity. We explore the relevance of this possibility aggregating the destination provinces by geographic groupings (north, center, and south), and studying whether the estimated effects of our variables of interest are different for investments that end up targeting each of these groups. The results, available upon request, do not highlight major discrepancies, except for the effects of multilateral emigration, which appears larger in Southern provinces than in Northern ones. Our data do not allow us to draw stronger conclusions about this intriguing finding, which we invite future research to consolidate and further explore.

function model and interact with the log of immigrants with a set of dummy variables representing the six most frequent functions in our data set (R&D, manufacturing, market access, business services, construction, and logistics). Results are shown in columns (1) and (2) of Table 6. These interaction terms express how different functions move the immigrant effect away from its estimated mean, which is the effect of immigrants for the residual category "Others."

Relative to the reference category, the interaction term is significantly larger for market access and business services FDI, indicating a larger effect of immigrants for these kinds of investments. Instead, the interaction effect is significantly smaller for manufacturing FDI. This suggests that FDI in downstream activities, which are presumably more information-sensitive, reacts more to the presence of bilateral immigrants. Instead, FDIs in manufacturing, which are arguably more intensive in terms of unskilled labor, turn out to be significantly less sensitive, an issue we shall further explore later. Still, the effect of immigration for the reference category remains positive, relatively large, and significant at the 10% level. Moreover, the standard deviation of the immigrant effect is still highly significant and indicates a positive effect for a wide majority of the investors. This suggests that the sources of heterogeneity in the immigrant effect are not exhausted by the functional heterogeneity.

We therefore explore other possible sources of heterogeneity in Section 5.2.2. As we shall discuss, the analysis of the second-stage regressions yields interesting insights, which will allow us to enrich our choice model with additional interactions that exhaust the sources of heterogeneity in the immigrants' effects (Model 6, Table 6) and to obtain a more parsimonious specification (Model 7, Table 6).

5.2.2 | Heterogeneity by characteristics of the investors

As discussed in Section 3.2.1, an alternative approach to explore the sources of heterogeneity in the immigrant effect is to simulate the firm-specific $\hat{\delta}_f^{\mathrm{Immi}}$ from the mixed logit estimates and employ them as the dependent variable of a linear regression model, as in Castellani and Lavoratori (2020) and Alcácer et al. (2018).

The $\hat{\delta}_f^{\mathrm{Immi}}$ from Models 1 to 4 in Table 5 and Model 5 in Table 6 are remarkably similar. Figure 2 shows the distribution of the simulated investor-specific values of $\hat{\delta}_f^{\mathrm{Immi}}$ drawn from Model 4. The average $\hat{\delta}_f^{\mathrm{Immi}}$ corresponds—with minor deviations attributable to numerical issues in simulation—to the mean coefficient of the log of immigrants estimated in Section 5.2.1. From Figure 2, we can appreciate that $\hat{\delta}_f^{\mathrm{Immi}}$ is quite heterogeneous but the distribution is firmly on the positive side. Hence, MNEs are generally attracted to provinces with more immigrants from the MNEs' home countries, but to a varying extent.

The results of the weighted least-squares regressions of Equation (4) are reported in Table 7.

The covariates reported in Table 7 can be interpreted as firm characteristics that are associated with larger or smaller $\hat{\delta}_f^{\mathrm{Immi}}$ and may provide indications on the prevailing channels.

In columns (1)–(4), we report the second-stage estimates corresponding to Models 1–4 in Table 5. These results provide strong support to the interpretation of the migrant effects as an information effect. Indeed, the coefficient of previous experience in Italy is negative and significant in all models.

These results are robust to substituting the dummy with a continuous variable, as well as with a categorical grouping of firms by the number of investments they have in Italy (1, 2, 3, 4, or more).[26] Figure 3 presents the relationship between the average $\hat{\delta}_f^{\mathrm{Immi}}$ estimated in Model 4 and the number of investments. The average coefficient is well above 0.4 for firms with only one investment, while it falls below 0.25 for firms with 4 or more

[26]These results are available upon request.

TABLE 6 Estimation results—Functional interactions.

	Model 5 Mixed logit		Model 6 Mixed logit		Model 7 Conditional logit
	Mean	SD	Mean	SD	Mean
	(1)	(2)	(3)	(4)	(5)
Log immigrants	0.313*	−0.342***	0.348**	−0.002	0.364**
	(0.187)	(0.075)	(0.176)	(0.208)	(0.157)
Log emigrants	−0.207	−0.004	−0.183	0.003	−0.159
	(0.231)	(0.115)	(0.226)	(0.133)	(0.212)
Residuals (immigrants)	−0.082	0.024	−0.084	0.013	−0.071
	(0.114)	(0.167)	(0.113)	(0.178)	(0.089)
Residuals (emigrants)	0.296**	−0.091	0.275**	0.112	0.270**
	(0.135)	(0.227)	(0.130)	(0.292)	(0.117)
Log immigrants × Research & Development	−0.052		0.042		0.050
	(0.247)		(0.233)		(0.205)
Log immigrants × Manufacturing	−0.487**		−0.383**		−0.361**
	(0.197)		(0.181)		(0.160)
Log immigrants × Market access	0.630***		0.525***		0.524***
	(0.192)		(0.183)		(0.159)
Log immigrants × Business services	0.659***		0.623***		0.640***
	(0.220)		(0.207)		(0.188)
Log immigrants × Logistics	−0.137		−0.033		−0.016
	(0.240)		(0.221)		(0.188)
Log immigrants × Construction	−0.206		−0.029		0.004
	(0.243)		(0.231)		(0.249)
Log emigrants × Research & Development	−0.234		−0.265		−0.240
	(0.266)		(0.262)		(0.243)
Log emigrants × Manufacturing	−0.078		−0.105		−0.102
	(0.215)		(0.211)		(0.203)
Log emigrants × Market access	0.083		0.089		0.077
	(0.220)		(0.216)		(0.205)
Log emigrants × Business services	0.142		0.093		0.063
	(0.244)		(0.239)		(0.233)
Log emigrants × Logistics	0.076		0.037		0.008
	(0.262)		(0.260)		(0.229)
Log emigrants × Construction	0.023		0.018		0.002
	(0.273)		(0.273)		(0.303)

TABLE 6 (Continued)

	Model 5 Mixed logit		Model 6 Mixed logit		Model 7 Conditional logit
	Mean	SD	Mean	SD	Mean
	(1)	(2)	(3)	(4)	(5)
Log immigrants × More than one investment in Italy			−0.745***		−0.819***
			(0.127)		(0.131)
Log emigrants × More than one investment in Italy			0.160		0.186
			(0.137)		(0.149)
Observations	111,692		111,692		111,692
AIC	3986.083		3948.955		3919.925[a]
BIC	4534.623		4516.741		4295.242[a]
LR test of joint significance of the SD	46.467		10.971		
Degrees of freedom	20		20		
Test p value	0.001		0.947		

Note: Location-choice models for inward FDI targeting Italian provinces. Models 5 and 6: Mixed logit estimates based on 500 Halton draws. For each model, the table reports the estimated average effect of the variable (Mean) and standard deviation (SD). In these models, standard errors take into account the correlation among the investments by the same company, but are not bootstrapped (see note 22). Model 5 includes first-stage residuals for both bilateral immigrants and emigrants and a set of interaction effects of migration with the main functional categories of the investments. Model 6 further includes an additional interaction with a dummy for whether the investor has previous investing experience in Italy at the time of the investment. The list of control variables is the same as in Model 4, Table 5. Covariates with fixed parameters: parent colocation, unemployment rate, infrastructure endowment, residents with tertiary education, sectoral diversity. Covariates with random parameters: log prov. population, institutional quality, common border, log prov. GDP, log patent count, log average wage (region), log distance, agglomeration (sector), log imports, log exports, pre-2002 FDI stock, log multilateral immigrants, log multilateral emigrants, firm density, manufacturing concentration, Rome/Milan dummy. Model 7 includes the same set of variables but is estimated by conditional logit with standard errors based on 2000 bootstrap samples, hence assumes fixed parameters.

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; FDI, foreign direct investment; GDP, gross domestic product; LR test, likelihood-ratio test.

[a]AIC and BIC statistics referring to the nonbootstrapped model. Standard errors in parentheses.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

investments. Coherently, the relative importance of Italy in the portfolio of a firm's worldwide investments also decreases the salience of the immigrant effect.

The demand channel finds only partial support. In line with our hypotheses, services tend to benefit more from the immigrant effect; however, foreign investors producing final goods are only slightly more reactive to migrants than those producing intermediate goods. As for the area of origin, we do not find significant differences relative to EU MNEs (used as the reference group).[27]

Quite interestingly, our results indicate that firms that make more labor-intensive investments tend to be significantly less sensitive to the presence of bilateral immigrants. This seems consistent with our finding in Table 6 about manufacturing FDI, which can also be assumed to be comparatively labor-intensive. This result is consistent

[27]We also experimented with Douglas Dow's cultural and religious distance measures (http://dow.net.au/), obtaining similar results.

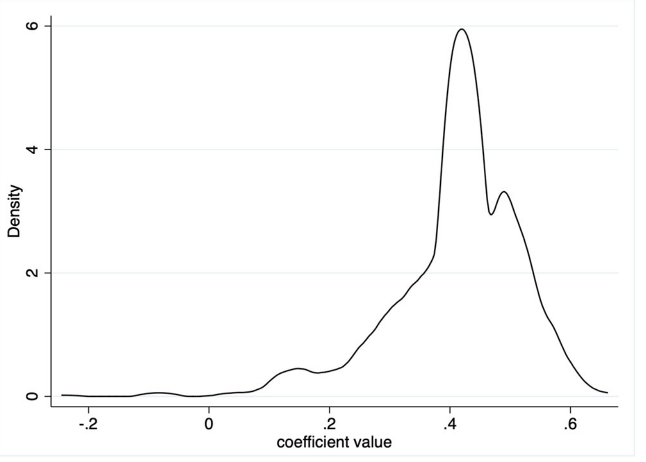

FIGURE 2 Distribution of the simulated $\hat{\delta}_f^{\mathrm{Immi}}$ coefficients. Kernel density distribution of firm-specific coefficients for the bilateral immigrants' effects $\hat{\delta}_f^{\mathrm{Immi}}$ simulated from the distributional parameters estimated in the mixed logit model reported in Table 5 (Model 4). The simulation procedure is described in Section 3.2.1.

with the expectation that other, more information-intensive investments, react more strongly to bilateral migration—the estimated net effect of immigration is still positive even for the investment with the highest labor intensity in our sample—but also raises the question of why labor-intensive investments should rely comparatively less on the labor supply of their co-nationals (see Supporting Information Appendix D for a more detailed discussion).

In a further set of regressions, we replace the type-of-sector dummies in column (4) with a set of 16 dummies indicating the main cluster of economic activity of the investing firm (Supporting Information Appendix Table C.12). Relative to the reference industrial sector, the results indicate that the immigrant effects are largest and most highly significant for the creative industries, financial services, ICT and electronics, and professional services, and, to a lesser extent, for wood, apparel, and related products; instead, investments in the physical sciences rely significantly less on immigrants. This suggests that investments drawing more strongly on intangible and firm-specific knowledge rely more on the immigrant effect to facilitate knowledge transfer with the target location.

In column (5) of Table 7, we report the second-stage results associated with the first-stage specification of Model 5 (Table 6), which includes a set of interaction terms of immigration and emigration with dummy variables identifying the main business function of each investment. These functional dummies correlate with investment size and sectors and appear to absorb the effects of total capital investment worldwide, of the share of Italy in the MNE portfolio, and of the labor intensity of the investment. This confirms the importance of accounting for the activities that the firm performs abroad when investigating the effects of migrants on FDI. On the other hand, conditional on functional heterogeneity, estimates in column (5) identify firm experience in Italy as the single main firm characteristic driving the heterogeneity in the immigrant effect. This is consistent with the arguments in Peri and Requena-Silvente (2010) on migrants' information effect.

5.3 | A more parsimonious specification

With these results in hand, we go back to our first-stage estimations. In fact, results from column (5) in Table 7 suggest that including a further interaction term with firm experience in Italy in the first stage may exhaust the sources of heterogeneity in the immigrant effect. To verify this hypothesis, in Model 6 of Table 6 (columns 3

TABLE 7 Sources of heterogeneity in the immigrant effect $\hat{\delta}_f^{\mathrm{Immi}}$.

Dependent variable: $\hat{\delta}_f^{\mathrm{Immi}}$	Model 1 (1)	Model 2 (2)	Model 3 (3)	Model 4 (4)	Model 5 (5)
Jobs/mln US$ invested	−0.001***	−0.001**	−0.001**	−0.001**	−0.001
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
Type of sector (ref: Services)					
Final goods	−0.032**	−0.033*	−0.033**	−0.036**	0.008
	(0.015)	(0.019)	(0.017)	(0.018)	(0.031)
Intermediate goods	−0.035***	−0.036***	−0.037***	−0.039***	−0.020
	(0.008)	(0.010)	(0.009)	(0.010)	(0.016)
Other goods	−0.029***	−0.029**	−0.029**	−0.031**	0.012
	(0.011)	(0.013)	(0.012)	(0.013)	(0.021)
Dummy: more than one investment in Italy	−0.040***	−0.042***	−0.041***	−0.042***	−0.043**
	(0.010)	(0.012)	(0.011)	(0.012)	(0.020)
Log total capital investment worldwide	−0.014***	−0.015***	−0.015***	−0.015***	−0.006
	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)
Italy share of capital investment worldwide	−0.055***	−0.060***	−0.057***	−0.060***	−0.014
	(0.011)	(0.014)	(0.012)	(0.013)	(0.022)
Area of origin (ref: EU)					
South–East Asia	0.019	0.015	0.022	0.020	0.025
	(0.013)	(0.015)	(0.014)	(0.015)	(0.025)
Non-EU Europe	0.003	0.005	0.002	0.003	0.003
	(0.015)	(0.018)	(0.016)	(0.017)	(0.029)
North America	0.003	0.001	0.002	0.002	0.001
	(0.008)	(0.010)	(0.009)	(0.009)	(0.015)
Rest of the world	0.017	0.015	0.019	0.015	0.006
	(0.020)	(0.024)	(0.022)	(0.023)	(0.039)
Constant	0.458***	0.584***	0.501***	0.513***	0.354***
	(0.014)	(0.016)	(0.015)	(0.015)	(0.026)
Observations	895	898	895	895	895
Test of the joint significance of the regressors	132.687	97.405	122.822	117.517	15.571
Degrees of freedom	11	11	11	11	11
Test p value	0.000	0.000	0.000	0.000	0.158

Note: Variance-weighted least-squares regression of firm-specific immigrant effects on investing company characteristics. Dependent variable: Firm-specific coefficients for the bilateral immigrants' effects $\hat{\delta}_f^{\mathrm{Immi}}$ simulated from the distributional parameters estimated in the mixed logit models reported in Tables 5 (Models 1–4) and 6 (Model 5). The simulation procedure is described in Section 3.2.1. Variances of the individual parameters $\hat{\delta}_f^{\mathrm{Immi}}$ estimated by parametric bootstrapping. Standard errors in parentheses.

*p < 0.10; **p < 0.05; ***p < 0.01.

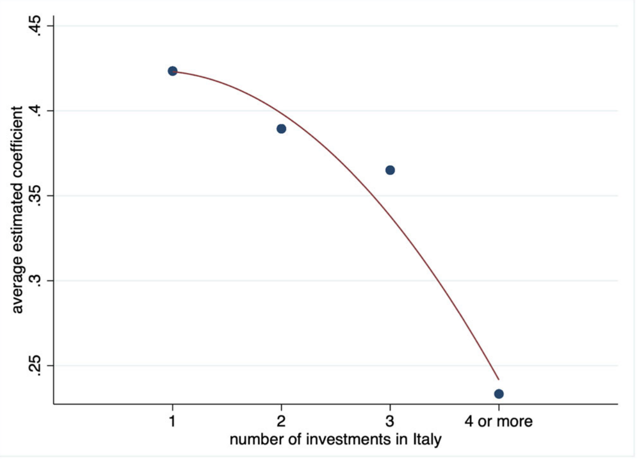

FIGURE 3 Average estimated coefficient $\hat{\delta}_f^{\text{Immi}}$ by the number of investments in Italy. Relationship between firm-specific immigration effects and the number of investments in Italy and quadratic fit. Firm-specific coefficients for the bilateral immigrants' effects $\hat{\delta}_f^{\text{Immi}}$ simulated from the distributional parameters estimated in the mixed logit model reported in Table 5 (Model 4). The simulation procedure is described in Section 3.2.1. [Color figure can be viewed at wileyonlinelibrary.com]

and 4) we include the full set of interaction terms between *log immigrants* and *log emigrants* with the functional dummies and with an additional dummy that identifies whether the firm had already invested in Italy at the time of the new investment. As expected, the interaction between *log immigrants* and the dummy *Already invested in Italy* turns negative and significant. The main effect of immigration remains positive and significant, and the directions of the other interaction effects are confirmed. What is more, the additional interaction effect appears to exhaust the sources of heterogeneity in the immigrants' effects. Indeed, both the standard deviation of *log immigration* and the likelihood ratio test on the joint significance of the standard deviations of the model become insignificant, implying that the null hypothesis that all standard deviations are equal to zero cannot be rejected.

Therefore, once we account for the relevant sources of heterogeneity—identified in the second-stage estimation—our location-choice model with random parameters simplifies to one with fixed parameters. Hence, we can estimate our location-choice model by a more parsimonious and computationally faster estimator, the conditional logit. As the standard errors of the conditional logit model can be easily bootstrapped, we can address remaining concerns about the standard errors in our main regression and take into account that the residuals in the control function are estimated (see note 22).

We report the results of the conditional logit model with bootstrapped standard errors (2000 replications) in Model 7 of Table 6. They are fully consistent with the ones in Model 6 and confirm that the immigrants' effects are strongest for firms investing in Italy for the first time and for Market-Access and Business-Services investments.

By contrast, although slightly positive, the net effect of immigration remains significantly smaller for manufacturing investments, even relative to other comparably labor-intensive functions, such as Construction and Logistics. This is somewhat unexpected: our theoretical arguments suggest that it could even be larger, as the country-specific labor supply channel may add up to the information effects. In Supporting Information Appendix D, we provide an explanation of this result in light of a possible substitution between migrants' protrade and pro-FDI effects. Given the substantially higher costs entailed in FDI compared to trade, an MNE engaging in horizontal FDI may find it more profitable to locate in provinces that are not already served via trade flows, either to avoid competition with other firms from its country of origin or to avoid duplicating efforts in serving a particular region (Blonigen, 2005).[28]

Overall, our analysis confronted different possible sources of heterogeneity in the immigrant effects and identified functional heterogeneity in foreign investments and host country-specific experience of the investors as the main drivers of the heterogeneity in the immigrant effect.

6 | DISCUSSION AND CONCLUSIONS

We confirm a positive, significant, and robust effect of bilateral immigration on FDI, beyond which lies significant heterogeneity. We contribute to the literature by providing an in-depth account of this heterogeneity, confronting different possible drivers, and identifying the most relevant ones. This allows us to shed light on the mechanisms underlying this relationship. Our results indicate that the immigrant effect mainly depends on the experience of the firm investing in Italy and on the main business function of the investment. Instead, our evidence does not support the effects of bilateral emigration or multilateral immigration or emigration.

Bilateral immigrants' effects are especially strong for firms that are investing in Italy for the first time and for investments in downstream activities. This supports the interpretation that immigrants' effects are essentially information effects that bridge the fixed costs of opening new ventures abroad. The role of experience in Italy is consistent with the argument of Peri and Requena-Silvente (2010) that immigrants' information bridges the fixed—rather than variable—costs of international business. The remarkable geographic concentration of the effects and the role of skills that we detected are also compatible with this interpretation. The significantly larger effect of immigration on downstream activities suggests that this information mainly concerns marketing and servicing opportunities.

The effects appear strongest for investments that target intangible assets and for firms operating in the creative industries and ICT and electronics. These investments are strongly tied to firm-specific human capital, organizational resources and capabilities, and corporate culture (Arrighetti et al., 2014; Montresor et al., 2013), and require specific knowledge to develop and adapt the MNE brand in the new market. According to our results, immigrants, especially skilled ones, are well suited to mediating between the MNE organizational routines and the knowledge assets available in the location. In this sense, we complement previous findings on the role of cultural factors in reducing information frictions (e.g., Head et al., 1995).

The effect of immigration significantly decreases after the first investment in Italy, suggesting that, after the fixed costs of entry have been overcome, MNEs develop their own networks and strategies to consolidate and diversify MNEs' presence in Italy in different provinces, and they need to rely less on the information-reducing effect given by the presence of immigrants.

By contrast, our results do not provide any support to the low-cost labor channel: multilateral immigration is insignificant throughout estimates, consistent with the more general finding that labor costs do not play a major role as location determinants for Italian inward FDI and with the Italian labor market rigidity. As for the country-specific labor supply channel, a complex interaction between the protrade and pro-FDI effects of migration for horizontal manufacturing investments emerges, on which we invite future research.

Our paper has implications for policymaking that contrast with the public discourse on immigration as a public concern. To the extent that FDIs contribute to the diffusion of knowledge and innovation, migrants can play a role in local economic development. Migrant information effects may end up supporting the activities of regional investment promotion agencies and can be facilitated by simplifying the communication channels to their homeland.

[28]We are very grateful to an anonymous referee for stimulating reflection in this regard.

On a methodological note, beyond stressing the fundamental fact that parameters vary across decisionmakers, our two-stage mixed logit approach turns out to be general enough to encompass differing extents to which observables can explain parameter heterogeneity. Our application ultimately worked as a model selection procedure, allowing us to reduce a mixed logit model to a more parsimonious conditional logit. This does not need to be the case. Had the second stage confirmed multiple heterogeneity drivers, the related conditional logit with interactions would have been cumbersome to interpret. Furthermore, had the expected set of drivers not exhausted parameter heterogeneity, the conditional logit with interactions would have been inappropriate. By contrast, our two-stage mixed logit approach remains valid and insightful in all cases.

While providing new insights into the mechanism underlying the immigrant effect on FDI, we recognize some limitations of our paper. First, we implicitly exclude non-Italian locations from our choice set. To the extent that foreign destinations are closer substitutes to Italian regions than other Italian regions, one may worry that relevant alternatives are left out. Although we cannot rule out this concern, some considerations make us confident that relevant foreign alternatives do not play a major role in driving our results. Most importantly, by flexibly allowing for unspecified correlation among alternatives, mixed logit models mitigate the relevance of this concern. Moreover, our results are robust to removing the alternatives that are most likely to have close substitutes abroad, that is, the global cities of Milan and Rome. Finally, previous studies have highlighted that the location choices of FDIs in Italy (especially from extra-EU countries) are consistent with the presence of an "Italy effect," hence with our implicit assumption that regions within Italy are closer substitutes with each other than with regions outside Italy (Basile et al., 2005, 2009).

Second, our data source, the *fDi Markets* database, refers to announced rather than implemented projects. We believe that our research question is robust to this limitation, because FDI announcements, too, are the result of a decision-making process that identified a specific destination as yielding the maximum expected utility. Hence, they can be taken to be informative about location determinants. As long as the announced investments that do not ultimately take place (if any) are randomly distributed across investors, locations, and sectors, our results will not be substantially affected. A targeted comparison of location and investment characteristics of investments that ultimately take place with those that do not would be extremely useful to confirm this assumption. It is important to note, however, that *FT Intelligence* constantly updates the *fDi Markets* database, eliminating projects that have not been actually carried out.

Third, due to data limitations, we were unable to distinguish between horizontal and vertical manufacturing FDI, which may help better disentangle the demand effect from the information effect and confirm our conjecture about the possibility that migrants and FDI in manufacturing are substitutes. The role of emigrants' skills and the analysis of migration effects at an even finer scale may also be addressed in future research if the relevant data becomes available.

Fourth, we focused on greenfield FDIs, which are not the only form of international investment, with mergers and acquisitions (M&A) being another common way of investing in foreign countries. Yet, we believe that, by establishing new ventures, greenfield FDIs are more informative about the characteristics and opportunities that drive location choice, while M&A is likely to be more sensitive to the characteristics of the target (merged/acquired) firms.

Finally, we draw conclusions about the effect of migration in facilitating the flow of information about the location but, due to the nature of our data, we cannot say much about whether they also facilitate the flow of information regarding specific firms. This somewhat narrower question would need firm-specific data on the origins of workers/members of the board in acquired/merged MNEs. labor market and demand effects will presumably play a minor role. Yet, cultural affinity between members of the board and managing teams may indeed trigger information effects and facilitate the signing of deals. This research question can be addressed in future research and may usefully complement the findings of the present study.

ACKNOWLEDGMENTS

We wish to thank the Editor, Alessandra Faggian, and two anonymous reviewers for very constructive comments and suggestions. We are indebted to Arne Risa Hole for his indications on the estimation of the standard errors of the individual random parameters, and to Marco Mutinelli for providing part of the data we employ. Kenneth Train and Amil Petrin provided helpful clarifications on the estimation of the control function. We also received helpful comments and suggestions from Laura Abrardi, Alessandro Arrighetti, Marco Bellandi, Luca de Benedictis, Ivan Etzo, Fabio Landini, Andrea Lasagni, Katiuscia Lavoratori, Roberto Leombruni, Francesco Lissoni, Alessia Lo Turco, Giulio Pedrini, Valentina Meliciani, Gianluca Santoni, and the participants in the 2018 AIB, ERSA, EARIE, AISRE, SIE conferences, 2019 SIEPI and ITSG, and 2022 ISGEP workshops. All usual disclaimers apply. Anna D'Ambrosio and Luigi Benfratello gratefully acknowledge financial support by Compagnia di San Paolo under the Starting Grant initiative.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from *fDi Markets*. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from http://www.fdimarkets.com/ with the permission of *fDi Markets*.

ORCID

Luigi Benfratello http://orcid.org/0000-0003-0402-5320
Anna D'Ambrosio https://orcid.org/0000-0002-7946-7695

REFERENCES

Alcácer, J., Chung, W., Hawk, A., & Pacheco-de Almeida, G. (2018). Applying random coefficient models to strategy research: Identifying and exploring firm heterogeneous effects. *Strategy Science*, *3*(3), 533–553.

Altonji, J. G., & Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In J. M. Abowd & R. B. Freeman (Eds.), *Immigration, trade, and the labor market* (pp. 201–234). University of Chicago Press.

Amiti, M., & Javorcik, B. (2008). Trade costs and location of foreign firms in China. *Journal of Development Economics*, *85*(1–2), 129–149.

Anelli, M., Basso, G., Ippedico, G., & Peri, G. (2023). Emigration and entrepreneurial drain. *American Economic Journal: Applied Economics*, *15*, 218–252.

Arkolakis, C. (2010). Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy*, *118*(6), 1151–1199.

Arrighetti, A., Landini, F., & Lasagni, A. (2014). Intangible assets and firm heterogeneity: Evidence from Italy. *Research Policy*, *43*(1), 202–213.

Basile, R. (2004). Acquisition versus greenfield investment: The location of foreign manufacturers in Italy. *Regional Science and Urban Economics*, *34*(1), 3–25.

Basile, R., Benfratello, L., & Castellani, D. (2005). Attracting foreign direct investments in Europe: Are Italian regions doomed? *Rivista di Politica Economica*, *95*, 319.

Basile, R., Castellani, D., & Zanfei, A. (2008). Location choices of multinational firms in Europe: The role of EU cohesion policy. *Journal of International Economics*, *74*(2), 328–340.

Basile, R., Castellani, D., & Zanfei, A. (2009). National boundaries and the location of multinational firms in Europe. *Papers in Regional Science*, *88*(4), 733–748.

Belderbos, R., Du, H. S., & Slangen, A. (2020). When do firms choose global cities as foreign investment locations within countries? The roles of contextual distance, knowledge intensity, and target-country experience. *Journal of World Business*, *55*(1), 101022.

Blonigen, B. A. (2005). A review of the empirical literature on FDI determinants. *Atlantic Economic Journal*, *33*(4), 383–403.

Bratti, M., De Benedictis, L., & Santoni, G. (2014). On the pro-trade effects of immigrants. *Review of World Economics*, *150*, 557–594.

Bratti, M., De Benedictis, L., & Santoni, G. (2020). Immigrant entrepreneurs, diasporas and exports. *Journal of Regional Science*, *60*, 249–272.

Buch, C., Kleinert, J., & Toubal, F. (2006). Where enterprises lead, people follow? Links between migration and FDI in Germany. *European Economic Review*, *50*, 2017–2036.

Buckley, P., & Casson, M. (1981). The optimal timing of a foreign direct investment. *Economic Journal*, *91*, 75–87.

Burchardi, K., Chaney, T., & Hassan, T. (2018). Migrants, ancestors, and foreign investments. *The Review of Economic Studies*, *86*(4), 1448–1486.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, *19*(1), 22–64.

Castellani, D., Jimenez, A., & Zanfei, A. (2013). How remote are R&D labs? Distance factors and international innovative activities. *Journal of International Business Studies*, *44*(7), 649–675.

Castellani, D., & Lavoratori, K. (2020). The lab and the plant: Offshore R&D and co-location with production activities. *Journal of International Business Studies*, *51*(1), 121–137.

Chaney, T. (2008). Distorted gravity: Heterogeneous firms, market structure and the geography of international trade. *American Economic Review*, *98*, 1707–1721.

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

Chung, W., & Alcácer, J. (2002). Knowledge seeking and location choice of foreign direct investment in the United States. *Management Science*, *48*(12), 1534–1554.

D'Ambrosio, A., & Montresor, S. (2022). The pro-export effect of subnational migration networks: New evidence from Spanish provinces. *Review of World Economics*, *158*, 53–107.

D'Ambrosio, A., Montresor, S., Parrilli, M. D., & Quatraro, F. (2019). Migration, communities on the move and international innovation networks: An empirical analysis of Spanish regions. *Regional Studies*, *53*(1), 6–16.

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

De Simone, G., & Manchin, M. (2012). Outward migration and inward FDI: Factor mobility between Eastern and Western Europe. *Review of International Economics*, *20*(3), 600–615.

Defever, F. (2006). Functional fragmentation and the location of multinational firms in the enlarged Europe. *Regional Science and Urban Economics*, *36*(5), 658–677.

Defever, F. (2012). The spatial organization of multinational firms. *The Canadian Journal of Economics/Revue canadienne d'Economique*, *45*(2), 672–697.

Docquier, F., & Lodigiani, E. (2010). Skilled migration and business networks. *Open Economies Review*, *21*, 565–588.

Du, J., Lu, Y., & Tao, Z. (2008). Economic institutions and FDI location choice: Evidence from US multinationals in China. *Journal of Comparative Economics*, *36*, 412–429.

Dunlevy, J. (2006). The impact of corruption and language on the pro-trade effect of immigrants: Evidence from the American States. *Review of Economics and Statistics*, *88*(1), 182–186.

Etzo, I., & Takaoka, S. (2018). The impact of migration on the cross-border M&A: Some evidence for Japan. *The World Economy*, *41*(9), 2464–2490.

Felbermayr, G., Grossmann, V., & Kohler, W. (2015). Migration, international trade, and capital formation. *Handbook of the Economics of International Migration*, *1*, 913–1025.

Foley, C. F., & Kerr, W. R. (2013). Ethnic innovation and US multinational firm activity. *Management Science*, *59*(7), 1529–1544.

Fujita, M., & Thisse, J.-F. (1996). Economics of agglomeration. *Journal of the Japanese and International Economies*, *10*(4), 339–378.

Gao, T. (2003). Ethnic Chinese networks and international investment: Evidence from inward FDI in China. *Journal of Asian Economics*, *14*, 611–629.

Girma, S., & Yu, Z. (2002). The link between immigration and trade: Evidence from the United Kingdom. *Review of World Economics*, *138*, 115–130.

Goerzen, A., Asmussen, C. G., & Nielsen, B. B. (2013). Global cities and multinational enterprise location strategy. *Journal of International Business Studies*, *44*(5), 427–450.

Gould, D. M. (1994). Immigrant links to the home country: Empirical implications for U.S. bilateral trade flows. *Review of Economics and Statistics*, *76*(2), 302–316.

Head, C. K., & Mayer, T. (2004). Market potential and the location of Japanese investment in the European Union. *Review of Economics and Statistics*, *86*(4), 959–972.

Head, C. K., Ries, J. C., & Swenson, D. L. (1995). Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. *Journal of International Economics*, *38*(3–4), 223–247.

Head, C. K., Ries, J. C., & Swenson, D. L. (1999). Attracting foreign manufacturing: Investment promotion and agglomeration. *Regional Science and Urban Economics*, *29*(2), 197–218.

Heckman, J., & Robb, R. (1985). Alternative methods for evaluating the impacts of interventions: An overview. *Journal of Econometrics*, *30*, 239–267.

Helpman, E., Melitz, M. J., & Yeaple, S. R. (2004). Export versus FDI with heterogeneous firms. *American Economic Review*, *94*(1), 300–316.

Hole, A. (2007). Estimating mixed logit models using maximum simulated likelihood. *Stata Journal*, *7*, 388–401.

Hornstein, A. S., & Greene, W. H. (2012). Usage of an estimated coefficient as a dependent variable. *Economics Letters*, *116*(3), 316–318.

Javorcik, B. S., Özden, Ç., Spatareanu, M., & Neagu, C. (2011). Migrant networks and foreign direct investment. *Journal of Development Economics*, *94*(2), 231–241.

Jayet, H., & Marchal, L. (2016). Migration and FDI: Reconciling the standard trade theory with empirical evidence. *Economic Modelling*, *59*, 46–66.

Kugler, M., & Rapoport, H. (2007). International labor and capital flows: Complements or substitutes? *Economics Letters*, *94*(2), 155–162.

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

Markusen, J. (2006). Modelling the offshoring of white-collar services: From comparative advantage to the new theories of trade and FDI. In L. Brainard & S. Collins (Eds.), *Offshoring white-collar work? The issues and the implications*. Brookings Institution.

Marschak, J. (1974). *Binary-choice constraints and random utility indicators* (pp. 218–239). Springer Netherlands.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). Academic Press.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, *15*(5), 447–470.

Montresor, S., Perani, G., & Vezzani, A. (2013). *How do companies "perceive" their intangibles? New statistical evidence from the Innobarometer 2013* (Technical Report). European Commission, Joint Research Centre—Institute for Prospective Technological Studies, Luxembourg, Publications Office of the European Union.

Nefussi, B., & Schwellnus, C. (2010). Does FDI in manufacturing cause FDI in business services? Evidence from French firm-level data. *The Canadian Journal of Economics/Revue canadienne d'Economique*, *43*(1), 180–203.

Nielsen, B. B., Asmussen, C. G., & Weatherall, C. D. (2017). The location choice of foreign direct investments: Empirical evidence and methodological challenges. *Journal of World Business*, *52*(1), 62–82.

Olney, W. W., & Pozzoli, D. (2021). The impact of immigration on firm-level offshoring. *Review of Economics and Statistics*, *103*(1), 177–195.

Ottaviano, G., & Peri, G. (2006). The economic value of cultural diversity: Evidence from US cities. *Journal of Economic Geography 6*, 9–44.

Ottaviano, G. I., Peri, G., & Wright, G. C. (2013). Immigration, offshoring, and American jobs. *American Economic Review*, *103*(5), 1925–1959.

Parsons, C. R. (2012). *Do migrants really foster trade? The trade-migration nexus, a panel approach 1960–2000*. The World Bank.

Peri, G., & Requena-Silvente, F. (2010). The trade creation effect of immigrants: Evidence from the remarkable case of Spain. *The Canadian Journal of Economics*, *43*(4), 1433–1459.

Petrin, A., & Train, K. (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research*, *47*, 3–13.

Rauch, J. (2001). Business and social networks in international trade. *Journal of Economic Literature*, *39*(4), 1177–1203.

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

Revelt, D., & Train, K. (1998). Mixed logit with repeated choices. *Review of Economics and Statistics*, *80*, 647–657.

Revelt, D., & Train, K. (2000). *Customer-specific taste parameters and mixed logit* [Working Paper No. E00-274]. Department of Economics, University of California.

Saxonhouse, G. R. (1976). Estimated parameters as dependent variables. *American Economic Review*, *66*(1), 178–183.

Spies, J. (2010). Network and border effects: Where do foreign multinationals locate in Germany? *Regional Science and Urban Economics*, *40*(1), 20–32.

Tadesse, B., & White, R. (2008). Do immigrants counter the effect of cultural distance on trade? Evidence from US state-level exports. *Journal of Socio-Economics*, *37*(6), 2304–2318.

Tintelnot, F. (2017). Global production with export platforms. *The Quarterly Journal of Economics*, *132*(1), 157–209.

Train, K. E. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge University Press.

Wagner, D., Head, K., & Ries, J. (2002). Immigration and the trade of provinces. *Scottish Journal of Political Economy*, *49*(5), 507–525.

White, R. (2007). Immigrant-trade links, transplanted home bias and network effects. *Applied Economics*, *39*(7), 839–852.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *The Journal of Human Resources*, *50*(2), 420–445.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Benfratello, L., Castellani, D., & D'Ambrosio, A. (2023). Migration and the location of MNE activities: Evidence from Italian provinces. *Journal of Regional Science*, 1–34. https://doi.org/10.1111/jors.12674