The role of immigrants, emigrants and locals in the historical formation of European knowledge agglomerations

Philipp Koch, Viktor Stojkoski & César A. Hidalgo

To cite this article: Philipp Koch, Viktor Stojkoski & César A. Hidalgo (27 Nov 2023): The role of immigrants, emigrants and locals in the historical formation of European knowledge agglomerations, Regional Studies, DOI: 10.1080/00343404.2023.2275571

To link to this article: https://doi.org/10.1080/00343404.2023.2275571

View supplementary material

Published online: 27 Nov 2023.

Submit your article to this journal

Article views: 420

View related articles

View Crossmark data
The role of immigrants, emigrants and locals in the historical formation of European knowledge agglomerations

Philipp Koch\textsuperscript{a,b}, Viktor Stojkoski\textsuperscript{a,c} and César A. Hidalgo\textsuperscript{a,d,e}

\textbf{ABSTRACT}

Did migrants make Paris a mecca for the arts and Vienna a beacon of classical music? Or was their rise a pure consequence of local actors? We use data on more than 22,000 historical individuals born between the years 1000 and 2000 to estimate the contribution of famous immigrants, emigrants and locals to the knowledge specialisations of European regions. We find that the probability that a region develops or keeps specialisation in an activity (based on the birth of famous physicists, painters, etc.) grows with both the presence of immigrants with knowledge about that activity and immigrants with knowledge in related activities. In contrast, we do not find robust evidence that the presence of locals with related knowledge explains entries and/or exits. We address some endogeneity concerns using fixed-effects models considering any location–period–activity-specific factors (e.g., the presence of a new university attracting scientists).

\textbf{KEYWORDS}

migration; knowledge spillovers; relatedness; economic history; economic complexity

\textbf{JEL} N90, O15, R11

\textbf{HISTORY} Received 28 February 2023; in revised form 13 October 2023

\section{1. INTRODUCTION}

Migrants help carry knowledge across space (Cipolla, 1972; Kerr et al., 2017; Lissoni, 2018; Tripl & Maier, 2011; Williams, 2006), shaping the geography of cultural and economic activities (Elekes et al., 2019; Miguelez & Morrison, 2022; Morrison, 2023; Nefik et al., 2018; Puterman & Weil, 2010). But most studies documenting the role of migrants in the diffusion of knowledge use recent data on patents (Bahar et al., 2020; Bernstein et al., 2022; Bosetti et al., 2015; Breschi et al., 2017; Fassio et al., 2019; Hunt & Gauthier-Loiselle, 2010; Miguelez & Morrison, 2022; Miguelez & Noumedem Temgoua, 2020; Miguelez & Moreno, 2013, 2015), research (Bosetti et al., 2015; Tripl, 2013) or product exports (Bahar & Rapoport, 2018), or analyse historical spillovers within activities (Borowiecki, 2012; Borowiecki & Graddy, 2021; Collins, 1974; Diodato et al., 2022; Ganguli, 2015; Hornung, 2014; Mitchell, 2019; Moser et al., 2014; Scoville, 1952a, 1952b; Waldinger, 2010, 2012), leaving questions about the role of migrants in the historical formation of knowledge agglomerations relatively unexplored.

To explore the role of migrants in the historical formation of knowledge agglomerations we use biographical data on more than 22,000 individuals – artists, physicists, explorers, philosophers, etc. – living in Europe between the years 1000 and 2000. We use these data to investigate how immigrants, emigrants and locals explain the probability that famous individuals specialised in an activity – that was not yet present in a region – are born during the next century. That is, we study how the knowledge of migrants and locals contributes to explain, for example, Paris becoming the birthplace of painters and Vienna of composers.

We can explore these questions by creating measures of knowledge spillovers within and between locations and activities. Consider spillovers across locations within the same activity. The knowledge that migrants carry across borders may impact a location’s ability to give birth to famous figures in the activity that the migrants specialise in. That is, immigrant mathematicians may increase the
probability that a city or region begets famous mathematicians. Similarly, emigrating mathematicians may decrease that probability. To capture such spillovers, we identify whether a region experiences a larger than expected inflow or outflow of famous individuals specialised in an activity.

Now consider spillovers across both locations and activities. Migrants and locals specialised in an activity (e.g., mathematics) can impact a region’s ability to give birth to famous figures in a related activity (e.g., physics). To capture such spillovers, we use measures of relatedness (Balland et al., 2022; Boschma, 2017; Hidalgo, 2021; Hidalgo et al., 2007, 2018), which exploit information on the co-location of activities to estimate how ‘cognitively close’ a location is to an activity.

During the past decades, measures of relatedness have been validated as robust predictors of the probability that countries, regions and cities enter or exit an activity, such as product exports (Hidalgo et al., 2007; Pinheiro et al., 2021; Poncet & de Waldemar, 2015), technologies (Balland et al., 2019; Ballard & Boschma, 2021; Boschma et al., 2015; Juhász et al., 2021; Petralia et al., 2017; Rigby, 2015; Uhlbach et al., 2022), industries (Deegan et al., 2021; Esseltzbichler, 2015; Neffke et al., 2011, 2017) and research areas (Boschma et al., 2014; Chinazzi et al., 2019; Guevara et al., 2016). Recent contributions to this literature have focused on unpacking relatedness by considering multiple channels (Bahar et al., 2019; Boschma & Capone, 2015; Cortinovis et al., 2017; Diodato et al., 2018; Farinha et al., 2019; Jara-Figueroa et al., 2018; Jun et al., 2020; Zhu et al., 2017). For instance, does industry- or occupation-specific knowledge contribute to the growth and survival of firms? (Jara-Figueroa et al., 2018). Or do value chains or knowledge agglomerations explain the co-location of firms? (Diodato et al., 2018). To the best of our knowledge no study has yet unpacked relatedness in the context of historical migration. Here we use a dataset spanning 1000 years of history in Europe to explore how the knowledge of immigrants, emigrants and locals explains the probability that a famous cultural figure specialised in an activity is born in a specific region. This contributes to both understanding the role of migrants in the geography of knowledge and unpacking relatedness metrics in the context of migration.

Our findings show that migrants play a crucial role in knowledge agglomerations. Specifically, we find that the probability that a European region enters a new activity grows on average by between 1.7 and 4.6 percentage points if that region received an excess number of immigrants specialised in that activity during the last century. Moreover, we find this correlation is enhanced by immigrants specialised in related activities. Similarly, we find the probability that a European region loses one of its existing specialisations decreases on average by 5.0–10.2 percentage points if that region received an excess number of immigrants specialised in that activity. This correlation is also enhanced by immigrants specialised in related activities. In contrast, we do not find a statistically significant and robust role of the related knowledge of locals (people born in that region) in entries or exits.

To tackle some important endogeneity concerns (migration is often a motivated choice), we employ a highly restrictive fixed effects structure controlling for all possible unobserved factors that are specific to a broad occupational category in a region during a century. These are factors that might affect both, migration patterns and the birth of famous individuals, such as a new university attracting scientists and leading to the birth of more famous scientists in the future, or a prosperous city attracting and begetting more artists. In addition, we control for unobserved factors that are specific to a more granular occupational category in a century which might affect both migration and births, for example, the emergence of a new technology (e.g., photography) begetting a new occupational category (photographers). This captures, for instance, that musicians and singers are likely to have different migration and birth patterns across time than other artists such as painters or actors. Lastly, we tackle some concerns of reverse causality by focusing on excess migration and estimating the expected number of migrants in a location. Although we control for multiple possible observed and unobserved factors to limit endogeneity concerns, we want to stress that we are not able to make strictly causal claims.

Together, these findings advance our understanding of the role of immigrants, emigrants and locals in the historical formation of knowledge agglomerations. They contribute to both the literature on the role of migrants in knowledge diffusion (Bahar et al., 2020; Bahar & Rapoport, 2018; Borowiecki, 2012; Borowiecki & Gradý, 2021; Bosetti et al., 2015; Breschi et al., 2017; Cipolla, 1972; Diodato et al., 2022; Elekes et al., 2019; Fassio et al., 2019; Ganguli, 2015; Hornung, 2014; Hunt & Gauthier-Loiselle, 2010; Kerr et al., 2017; Lissoni, 2018; Migueluez & Morrison, 2022; Migueluez & Noumedem Temgoua, 2020; Migueluez & Moreno, 2013, 2015; Mitchell, 2019; Moser et al., 2014; Neffke et al., 2018; Putterman & Weil, 2010; Trippl, 2013; Trippl & Maier, 2011; Williams, 2006) and that on relatedness (Bahar et al., 2019; Balland et al., 2019, 2022; Ballard & Boschma, 2021; Boschma, 2017; Boschma et al., 2014, 2015; Boschma & Capone, 2015; Chinazzi et al., 2019; Cortinovis et al., 2017; Degan et al., 2021; Diodato et al., 2018; Esseltzbichler, 2015; Farinha et al., 2019; Guevara et al., 2016; Hidalgo, 2021; Hidalgo et al., 2007, 2018; Jara-Figueroa et al., 2018; Juhász et al., 2021; Jun et al., 2020; Neffke et al., 2011, 2017; Petralia et al., 2017; Pinheiro et al., 2021; Poncet & de Waldemar, 2015; Rigby, 2015; Uhlbach et al., 2022; Zhu et al., 2017). Moreover, by developing measures of the related knowledge of migrants, we combine both migration and relatedness in a framework that can be used to study how knowledge spillovers across space and across activities combine in more recent settings. Lastly, this study provides a long-term perspective on the evolution of regional specialisations in Europe, a perspective which is underrepresented in the field of economic geography (Henning, 2019).
2. DATA AND METHODS

2.1. Data
We use the 2020 version of Pantheon (Yu et al., 2016), a publicly available dataset including information on famous individuals with a Wikipedia page in more than 15 different language editions. We focus on the 22,847 famous individuals born or died in Europe between the years 1000 and 2000. We choose Pantheon because it assigns individuals using a controlled taxonomy of 101 occupations, such as painter, writer, composer, physicist, chemist, mathematician, etc. Pantheon provides a good sectoral disaggregation compared with other datasets which either have few sectors (Schich et al., 2014) or use uncontrolled taxonomies with duplicate entries, for example, film director and movie director (Laouenan et al., 2022). This granularity is needed to construct measures of specialisation and relatedness. The full taxonomy and descriptive statistics are provided in Section 1.1 in the supplemental data online.

We use geographical coordinates to assign the place of birth and death of each biography to European administrative regions (NUTS-2 or regions of similar size for countries outside the EU, for example, Russian oblasts; see Section 1.2 in the supplemental data online). Figures 1 (a) and (b) show the places of birth and death of all individuals in our dataset within the applied administrative borders. Due to a lack of data on the full trajectory of individuals, we follow the literature investigating migration patterns of famous individuals (De La Croix & Licandro, 2015; Laouenan et al., 2022; Schich et al., 2014; Serafelli & Tabellini, 2022) and use places of birth and death as rough proxies for migration. Manual inspection of a random sample of 200 biographies revealed places of death to be a valid proxy of an important living place for around 90% of biographies and corresponded to a place of major impact for 75% of biographies (see Section 1.3 in the supplemental data online).

Finally, we assign each individual to a century \( t \) based solely on his or her year of birth. That is, a famous person who is born in the 18th century in Brussels and died in Paris (in the 18th or 19th centuries) is considered a local in Brussels and an immigrant in Paris in the 18th century. We choose this approach since we do not have information on the time of migration. We take this into account in the regression models by lagging the independent variables (see also Section 1.1 in the supplemental data online).

2.2. Descriptive statistics: migration and spatial concentration patterns
We find that most of the migration of famous Europeans over the past 1000 years took place within countries and towards large cities (e.g., from smaller cities in France to Paris). Figures 1(c) and (d) visualise the migration network. Migration is common among famous individuals. In fact, going back to the 11th century, the share of migrants in our dataset never drops below 65%. In the 19th century, almost 80% of famous individuals in our dataset died in a different region than the one in which they were born (Figure 1e).

These migration patterns are not random but follow a process of preferential attachment, clustering individuals in major cities (Borowiecki, 2013; Borowiecki & Dahl, 2021; O’Hagan & Borowiecki, 2010; O’Hagan & Hellmanzik, 2008; Schich et al., 2014; Serafelli & Tabellini, 2022) and leading to a higher spatial concentration for places of death than birth. For instance, 416 famous individuals were born in Paris in the 19th century, but 934 died there (see Section 2.1 in the supplemental data online).

We use information entropy \( H \) to quantify the spatial concentration of births and deaths across regions. Information entropy (base 2) estimates the number of yes/no questions that we would need to answer – on average – to find the place of birth or death of an individual (see Section 2.1 in the supplemental data online). If deaths are more concentrated than births, we will need less questions to guess a place of death than one of birth. We can use entropy \( H \) to estimate the effective number of places of birth or death as \( E = 2^H \), which is the number of regions effectively experiencing the birth or death of a famous individual.

Figure 1(f) shows the effective number of places of birth and death \( E \) for each century. Before the 15th century, the spatial concentration of famous births and deaths was similar. But starting in the 15th century, places of death have become more spatially concentrated and places of birth more widespread. In fact, by the 19th century famous individuals were effectively born in more than 200 (out of 405) regions across Europe, while they effectively died in only 100 regions (Figure 1f).

2.3. Methods

2.3.1. Relatedness of immigrants, emigrants and locals
To explore how the knowledge of immigrants, emigrants and locals shapes the geography of knowledge, we estimate the probability that a region gives birth to a famous individual specialised in an activity as a function of estimates of knowledge spillovers within and between regions and activities.

To capture the knowledge spillovers of migration within the same activity, we calculate the ratio between the observed number of famous immigrants \( N_{ik,t}^{imm} \) or emigrants \( N_{ik,t}^{exp} \) with a certain activity and their expected number (respectively \( N_{ik,t}^{imm} \) and \( N_{ik,t}^{exp} \)), where \( i \) denotes the region, \( k \) the occupation and \( t \) the century.

Taking the ratio between the observed and expected number of migrants allows us to create measures of excess immigration or excess emigration, and thus, to control for the natural attractiveness of a location and the characteristics of an activity. This is important to address reverse causality concerns, since the effects of migrants could be simply a reflection of local factors making a place attractive for migrants with a certain specialisation.
Figure 1. Places of birth, places of death and migration patterns of famous individuals in Europe over the past 1000 years. Maps of (a) places of birth and (b) places of death included in the analysis (NUTS-2 regions for the EU, comparable regions for other countries, for example, oblasts in Russia; see Section 1.2 in the supplemental data online). Migration network of famous individuals within Europe over the past 1000 years, using (c) geography or (d) a force-directed algorithm for visualisation. The latter reveals that famous individuals tend to move within countries towards large regions. (e) Share of migrants in the dataset per century. (f) Effective number of places of birth and death $E$ derived from Shannon entropy (see Section 2.1 in the supplemental data online). Starting in the 15th century, the places of death of famous individuals are more spatially concentrated than their places of birth.
It is worth mentioning that migration decisions can be influenced by multiple local factors. Creatives, for instance, are more likely to move towards places that are already populated by other creatives (Borowiecki & Graddy, 2021) or potential patrons (Haskell, 1996/2000; Oevermann et al., 2007). Geographical and cultural distance (Caragliu et al., 2013; Lever & Van den Berg, 2008), such as a common language or the presence of fellow countrymen, can also play a role (Rephann & Vencatasawmy, 2000). Lastly, migration can also be exogenously forced due to conflict (Borowiecki, 2012) or climate (Abel et al., 2019). By focusing on excess migrants instead of total migrants, in a restrictive fixed-effects model we help mitigate the risks of reverse causality. Mathematically, this involves taking the ratio between the observed and expected number of immigrants or emigrants:

\[
R_{ik,t}^{\text{in}} = \frac{N_{ik,t}^{\text{in}}}{N_{ik,t}^{\text{in},*}},
\]

\[
R_{ik,t}^{\text{out}} = \frac{N_{ik,t}^{\text{out}}}{N_{ik,t}^{\text{out},*}},
\]

where the values are for individuals in region \(i\) and activity \(k\) born in century \(t\).

Here we use two models for the expected number of migrants (\(N_{ik,t}^{\text{in}}\)). The first considers the number of individuals in a location and the number of individuals specialised in an activity. That is a ‘bins and balls’ model for the expected number of immigrants or emigrants, making equation (1) the revealed comparative advantage (Balassa, 1965) or location quotient, a common measure of specialisation:

\[
\hat{N}_{ik,t} = \frac{\sum_i N_{ik,t} \sum_j N_{ik,t}}{\sum_j N_{ik,t}}.
\]

The second model expands on this by taking into account the attractiveness of a location in a specific activity (Neffke et al., 2011). We model \(\hat{N}_{ik,t}\) using a negative binomial regression where we control for the observed number in the previous century (\(N_{ik,t-1}\)), the previous specialisation of the location in the activity based on famous individuals born there (\(S_{t,\text{births}}^{\text{births}}\)),

\[
S_{t,\text{births}}^{\text{births}} = \frac{N_{t,\text{births}}^{\text{births}}}{\sum_i N_{t,\text{births}}^{\text{births}} - \sum_i N_{t,\text{births}}^{\text{births}}}.
\]

where \(N_{t,\text{births}}^{\text{births}}\) denotes the number of famous individuals born in location \(i\) specialised in activity \(k\) in century \(t\), and fixed effects for each location-time (\(\theta_{ik}\)) and activity-time (\(\theta_{ik}\)) to account for unobserved factors. That is, we estimate:

\[
\hat{N}_{ik,t} = f(\alpha_0 + \alpha_1 S_{t,\text{births}}^{\text{births}} + \theta_u + \theta_{ik}),
\]

where \(f\) denotes the negative binomial probability density (see Section 3.4.1 in the supplemental data online for results).

If the observed number of immigrants or emigrants in an activity exceeds the expected number, we say that region received excess immigrants, or produced excess emigrants, on that activity and time period.

Next, we create two specialisation matrices for famous immigrants (\(M_{ik,t}^{\text{in}}, \text{ died ‘here’ but born elsewhere}\)) and emigrants (\(M_{ik,t}^{\text{out}}, \text{ born ‘here’ but died elsewhere}\)):

\[
M_{ik,t}^{\text{in}} = \begin{cases} 1 & \text{if } R_{ik,t}^{\text{in}} \geq 1 \\ 0 & \text{otherwise} \end{cases}
\]

\[
M_{ik,t}^{\text{out}} = \begin{cases} 1 & \text{if } R_{ik,t}^{\text{out}} \geq 1 \\ 0 & \text{otherwise} \end{cases}
\]

Figures 2(a) and (b) show these two matrices using data for individuals born in the 19th century. The matrices are characterised by a nested structure that we recover by sorting locations by diversity (respectively \(\sum_i M_{ik,t}^{\text{in}}\) and \(\sum_i M_{ik,t}^{\text{out}}\)), and activities by ubiquity (respectively \(\sum_j M_{ik,t}^{\text{in}}\) and \(\sum_j M_{ik,t}^{\text{out}}\)). This structure is typical for matrices summarising the geography of activities (Bustos et al., 2012; Hausmann & Hidalgo, 2011) (see Section 2.2 in the supplemental data online), but also for networks describing species interactions in ecology (Almeida-Neto et al., 2008; Bascompte et al., 2003; Bastolla et al., 2009).

To capture spillovers across activities we use measures of relatedness (Balland et al., 2022; Boschma, 2017; Hidalgo, 2021; Hidalgo et al., 2007, 2018). Relatedness exploits information on the co-location of activities to estimate their affinity with a location. We create three separate measures of relatedness for immigrants, emigrants and locals.

These measures build on the specialisation matrices described in equation (5). This time, however, we need to create specialisation matrices for locals, which we define as famous individuals who were born in a region, no matter if they died there or elsewhere. We use this definition because of the large share of migrants among famous individuals (Figure 1d), which would reduce our number of observations drastically if we defined locals as individuals who were born and died in the same place. Controlling for the related knowledge of emigrants, however, relatedness based on all births is a valid proxy for the related knowledge of individuals who were born and died in the same region (see Section 2.3 in the supplemental data online).

That is, as before, we calculate the ratio between observed and expected births of famous individuals:

\[
R_{ik,t}^{\text{births}} = \frac{N_{ik,t}^{\text{births}}}{N_{ik,t}^{\text{births},*}}.
\]

Again, we can apply both the naïve model described in equation (2) or estimate the expected number of births given local factors (see equation 3 in Section 3.4.1 in the supplemental data online) before creating binary
specialisation matrices for locals:

\[
M_{births}^{ik, t} = \begin{cases} 
1 & \text{if } R_{births}^{ik, t} \geq 1 \\
0 & \text{otherwise} 
\end{cases}
\]

This matrix also exhibits a nested structure (Figure 2c).

Next, we define the proximity or similarity between two activities as the minimum of the conditional probability that a location is specialised in both of them (Hidalgo et al., 2007):

\[
\varphi_{imi, kk, t} = \frac{\sum_i M_{imi, ik, t} M_{imi, ik, t}'}{\max(\sum_i M_{imi, ik, t}, \sum_i M_{imi, ik, t}')} 
\]

\[
\varphi_{emi, kk, t} = \frac{\sum_i M_{emi, ik, t} M_{emi, ik, t}'}{\max(\sum_i M_{emi, ik, t}, \sum_i M_{emi, ik, t}')} 
\]

and use these proximities to calculate the relatedness between locations and activities as:

\[
\omega_{imi, kk, t} = \frac{\sum_i M_{imi, ik, t} M_{imi, ik, t}'}{\sum_i \varphi_{imi, kk, t}} 
\]

\[
\omega_{emi, kk, t} = \frac{\sum_i M_{emi, ik, t} M_{emi, ik, t}'}{\sum_i \varphi_{emi, kk, t}} 
\]

\[
\omega_{births, kk, t} = \frac{\sum_i M_{births, ik, t} M_{births, ik, t}'}{\sum_i \varphi_{births, kk, t}} 
\]

These measures quantify how far, for example, immigrants to Paris are from being specialised in archaeology, emigrants from Madrid are from being specialised in singing or locals in Berlin are from being specialised in philosophy.

We note that the relatedness densities calculated with the naïve and binomial model are highly correlated ($R^2 > 0.9$). So, going forward, we present results using the

Figure 2. Nested specialisation matrices and the similarity between activities in the 19th century. Specialisation matrices based on (a) immigrants, (b) emigrants and (c) locals in the 19th century (see equations 5 and 7). Examples of locations and activities (I, writer; II, mathematician; III, physicist; IV, journalist; V, pilot) are highlighted. (d) Proximities between activities based on the co-location of famous immigrants born in the 19th century using the naïve model described in equation (2) to determine the expected number of immigrants. Node size is proportional to the number of famous individuals specialised in the respective activity and born in the 19th century.
naive model and provide additional results using the negative binomial model in Section 3.4.1 in the supplemental data online.

Since the multiple factors contributing to the co-location of activities can be different when looking at immigration, emigration and births, we create separate measures of proximity \( \phi^{\text{imm}}_{ik,t} \), \( \phi^{\text{em}}_{ik,t} \), and \( \phi^{\text{birth}}_{ik,t} \), equation 8). But as a robustness check, we also consider a joint measure of proximity \( \phi^{\text{joint}}_{ik,t} \) using co-location at birth and death (see Section 2.5 in the supplemental data online). Nevertheless, we find the separate measures of proximity provide valuable nuance (see Figure A6 online). Consider explorers and military personnel. Explorers and military personnel share many required capabilities such as navigating, planning, commanding, etc., that may be explained by local factors such as military academies for education, distance to the sea, recency of a war or naval technology. Also, exploration teams often involve soldiers and military personnel, which could then become famous as explorers. Hence, explorers and military personnel are likely to share a geographical origin. Yet, since exploration and military campaigns tend to involve different locations, these two activities are less likely to collocate at death. Now consider composers and noblemen. For these two activities, the proximity based on immigration patterns is higher than the proximity based on births. It makes sense that these activities are to some extent related when looking at places of birth: noblemen are known to be patrons for the arts. Hence, noblemen born in a location will likely create institutions that promote the cultivation of the talent of composers born in this location. But it is likely to be a highly related activity, the proximity based on immigration patterns is higher than the proximity based on births. It makes sense that these activities are to some extent related when looking at places of birth: noblemen are known to be patrons for the arts. Hence, noblemen born in a location will likely create institutions that promote the cultivation of the talent of composers born in this location. But it is also plausible that these activities are even more related when looking at immigration patterns. Given that we observe a disproportional migration flow of noblemen towards a certain location, we can view this location as highly related to composers, since the institutional factors attracting noblemen likely play a role in attracting and cultivating the talent of composers as well.

These examples highlight why we believe that generating separate measures of proximity for immigrants, emigrants and births provides a nuanced perspective that helps unpack relatedness (see Section 2.5 in the supplemental data online for more details).

We illustrate the structure of these proximity networks for immigrants born in the 19th century \( \phi^{\text{imm}}_{ik,19} \) (Figure 2d). A high proximity between two activities indicates similarity or complementarity among them. Like measures of propensity, measures of proximity capture the combined presence of multiple factors that may be contributing to the co-location of two activities. For example, we find a high proximity between biologists and physicians, mathematicians and physicists, and musicians and actors (Figure 2d). While the latter may be considered an example of co-location due to high complementarity (musicians and actors may perform together), associations between mathematicians and physicists, or biologists and physicians, may indicate similarity in knowledge or skills.

### 2.3.2. Entries and exits

We use our measures of relatedness to study the entry and exit of activities in European regions. We do this by estimating logistic models explaining the probability that a region starts to give birth to a disproportionately large number of famous individual specialised in an activity (entries) or stops doing so (exits). That is, a region enters the activity ‘philosophy’ if more philosophers are born there in a certain century than expected, while this has not been the case in the prior century. Similarly, we explain the probability that a region loses an existing specialisation. A region exits the activity ‘physics’ if fewer physicists are born there than expected, while this has not been the case in the prior century. The variables \( E_{ik,1} \) and \( E_{ik,1} \) emerge directly from the specialisation matrix defined in equation (7).

Specifically, we define:

\[
E_{ik,1} = \begin{cases} 1 & \text{if } M^{\text{birth}}_{ik,1} = 0 \text{ and } M^{\text{birth}}_{ik,1} = 1, \\ 0 & \text{otherwise} \end{cases}
\]

\[
E_{ik,1} = \begin{cases} 1 & \text{if } M^{\text{birth}}_{ik,1} = 1 \text{ and } M^{\text{birth}}_{ik,1} = 0, \\ 0 & \text{otherwise} \end{cases}
\]

That is, a region \( i \) enters (exits) an occupation \( k \) in century \( t \) if the observed births of famous individuals with that occupation during the considered century is larger (lower) than expected, while this was not the case in the prior century.

Defining entries and exits looking at places of birth is a rather conservative approach. For a region to enter an activity, it needs to become a place where the required knowledge to cultivate a certain talent can be absorbed through formal or informal institutions and social ties. Indeed, early exposure to local knowledge in an individual’s life is highly relevant in shaping his or her career, both for inventors nowadays (Bell et al., 2019) and artists centuries ago (Galenson, 2009, p. 278). A different approach to describing the geography of knowledge would be, for instance, to focus on all individuals living at a certain place. But this would require having data on all places of living.

We explain entries and exits using measures of the presence of immigrants and emigrants in that activity \( \phi^{\text{imm}}_{ik,t} \) and \( \phi^{\text{em}}_{ik,t} \) and of the related activities that we can attribute to immigrants, emigrants and locals \( \phi^{\text{imm}}_{ik,t} \), \( \phi^{\text{em}}_{ik,t} \), and \( \phi^{\text{birth}}_{ik,t} \). For instance, a significantly positive correlation between \( \phi^{\text{imm}}_{ik,t} \) and entries would point towards migrants bringing into the region the knowledge needed to carry out activity \( k \). That is, a high influx of mathematicians would increase the probability that the region begets its own famous mathematicians. This would be consistent with research showing that migrants help carry the knowledge needed to enter an activity (Bahar et al., 2020; Bahar & Rapoport, 2018; Borowiecki, 2012; Borowiecki & Graddy, 2021; Bosetti et al., 2015; Breschi et al., 2017; Diodato et al., 2022; Fassio et al., 2019; Ganguli, 2015; Hornung, 2014; Hunt & Gauthier-
Similarly, a significant correlation between $a^i_{k,t}$ and entries would support the idea that the related knowledge brought by migrants also impacts the probability that a region develops a new activity. That is, the knowledge of famous immigrants specialised in mathematics diffuses to related fields, such as physics or chemistry, and increases the probability that a region begets its own physicists or chemists.

Lastly, a significant correlation between $a^i_{k,t}$ and entries, after controlling for $a^m_{k,t}$, would indicate that the related knowledge of locals contributes to entering a new activity. That is, a region with many locals already specialised in mathematics has a higher probability of branching into physics or chemistry.

The entry of a region into a new activity could be the result of multiple factors other than migration. For instance, the creation of a new university could attract scientists, and the expansion of a port could create conditions attractive to merchants. We address such endogenous concerns by using highly restrictive conditions attractive to merchants. We address such endogenous concerns by using highly restrictive conditions. This includes an activity’s ubiquity (i.e., the number of locations specialised in it) and how close a region already is to having or losing a specialisation ($b^i_{k,t}$, see equation 6). Lastly, we account for knowledge diffusion across space due to other reasons than migration by creating measures of the spatial proximity to other regions with specialisations in that specific activity or in related activities (see Section 2.4 in the supplemental data online). We provide descriptive statistics and discuss the explanatory variables in more detail in Section 3.1 online.

In sum, we define $Y_{k,t} = \{\text{Entry}_{k,t}, \text{Exit}_{k,t}\}$ and estimate:

$$P(Y_{k,t}) = g(b_1M_{k,t-1} + \beta_2M_{k,t}^\text{imm} + \beta_3a^i_{k,t-1} + \beta_4a^m_{k,t-1} + \beta_5\delta_X + \gamma_{nit} + \delta t + \epsilon_{ik,t},)$$

where $g$ denotes the logistic probability density, $X_{k,t-1}$ denotes a vector of observed control variables, and $\gamma_{nit}$, $\delta t$ the fixed effects.

We calculate average marginal effects based on this logistic regression by computing the marginal effect for each data point and taking the average.

### 3. RESULTS

Table 1 and Figure 3 show the relationship between the activities of immigrants, emigrants and locals and the number of observed entries and exits. For entries (Table 1, columns 1–5), the probability correlates positively with an excess inflow of migrants specialised in an activity during the previous century ($M_{k,t}^\text{imm} = 1$). Specifically, an excess of immigrants increases the probability of entry on average by 4.6 percentage points (Figure 3a). Figure 3(b) plots the probability of entry as a function of $M_{k,t}^\text{imm}$.

We also find that the probability of entry grows with the related knowledge of immigrants. A standard deviation increase of $a^i_{k,t}$ increases the probability of entry on average by 5.8 percentage points (Figure 3a). Figure 3(c) visualises the results by plotting the average probability of entry as a function of the relatedness density of immigrants ($a^i_{k,t}$). In accordance with the literature (Alshamsi et al., 2018; Hidalgo et al., 2018), the average probability of entry grows super-linearly, from 1.1% if no related knowledge of famous immigrants is present in a region ($a^i_{k,t-1} = 0$) to 16.4% if all related activities are present ($a^i_{k,t-1} = 100$). Moreover, we find a positive correlation ($\rho < 0.1$) between $a^i_{k,t}$ and entries, but unlike the estimate of the related knowledge of immigrants, this correlation is not robust (see Section 3.4 in the supplemental data online).

When we look at exits (Table 1, columns 6–10), we find similar relationships but with the opposite sign. An excess inflow of famous individuals specialised in an activity during the previous century ($M_{k,t}^\text{imm} = 1$) reduces the probability of exit significantly by 10.2 percentage points on average (Figure 3d). Also, the related knowledge of immigrants ($a^i_{k,t}$) helps prevent losing specialisation in an activity. Figure 3(e) and (f) visualise these results by plotting the probability of exit as a function of $M_{k,t}^\text{imm}$ and $a^i_{k,t}$, respectively.

These results are robust to estimating the expected number of immigrants, emigrants and locals in equations (1) and (6) using the negative binomial regression model described in equation (3) (see Section 3.4.1 in the supplemental data online). By accounting for local factors, we are able to obtain a more accurate estimate of the
Table 1. Main results of logistic regression models explaining entries and exits of activities.

|                  | Dependent variable: *Entry*_{ikt} |                  | Dependent variable: *Exit*_{ikt} |
|------------------|-----------------------------------|------------------|----------------------------------|
|                  | (1)                               | (2)              | (3)                             | (4)                             | (5)                             | (6)                              | (7)                             | (8)                             | (9)                             | (10)                            |
|                  | 0.334***                          | 0.303***         | 0.336***                        | 0.331***                        | 0.300***                        | −0.603***                       | −0.584***                       | −0.591***                       | −0.587***                       | −0.571***                       |
|                  | (0.080)                           | (0.075)          | (0.086)                         | (0.080)                         | (0.076)                         | (0.127)                         | (0.134)                         | (0.120)                         | (0.126)                         | (0.126)                         |
|                  | 0.115                             | 0.045            | 0.106                           | 0.121                           | 0.018                           | 0.310                           | 0.330                           | 0.233                           | 0.306                           | 0.291                           |
|                  | (0.261)                           | (0.278)          | (0.261)                         | (0.255)                         | (0.270)                         | (0.240)                         | (0.232)                         | (0.216)                         | (0.222)                         | (0.203)                         |
|                  | 0.027***                          |                  | 0.028***                        | (0.007)                         | −0.067***                       | (0.016)                         | −0.064***                       | (0.011)                         | −0.064***                       | (0.011)                         |
|                  | (0.006)                           |                  |                                |                                |                                |                                |                                |                                |                                |                                |
|                  | −0.006                            |                  | −0.024                          | (0.019)                         | −0.048                          | (0.038)                         | −0.025                          | (0.063)                         | −0.025                          | (0.063)                         |
|                  | (0.012)                           |                  |                                |                                |                                |                                |                                |                                |                                |                                |
|                  | 0.011                             |                  | 0.027*                          | (0.008)                         |                                | (0.015)                         |                                |                                |                                | (0.018)                         |
|                  | (0.015)                           |                  |                                |                                |                                |                                |                                |                                |                                | (0.041)                         |
| Further controls | ✓                                 | ✓                | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               |
| Fixed effects    |                                   |                  |                                |                                |                                |                                |                                |                                |                                |                                |
| Broad category–region–century | ✓                                 | ✓                | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               |
| Category–century | ✓                                 | ✓                | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               | ✓                               |
| Observations     | 3944                              | 3944             | 3944                            | 3944                            | 3944                            | 1051                            | 1051                            | 1051                            | 1051                            | 1051                            |
| Pseudo-\(R^2\)  | 0.213                             | 0.214            | 0.213                           | 0.213                           | 0.215                           | 0.224                           | 0.230                           | 0.226                           | 0.226                           | 0.232                           |
| BIC              | 9537.0                            | 9539.4           | 9545.0                          | 9544.5                          | 9553.1                          | 3619.6                          | 3618.0                          | 3623.4                          | 3623.3                          | 3628.8                          |

Note: The fixed effects in these models are highly restrictive, amounting to more than 700 parameters in columns (1) to (5) and more than 350 parameters in columns (6) to (10). All regions included in the regression model exhibit a minimum number of births and migrants such that measures of specialisation and relatedness are defined (see Section 2.2 in the supplementary data online). Standard errors are clustered by region and period. For the full regression tables with all control variables, see Sections 3.2 and 3.3 online. BIC, Bayesian information criterion. *\(p < 0.1\), **\(p < 0.05\), ***\(p < 0.01\).
expected number of immigrants, emigrants and locals and, thus, of a disproportionate migration flow. This mitigates some of the endogeneity concerns.

The highly restrictive fixed effects specification, however, reduces the number of observations in the regression model. To assure the robustness of our results, we estimate the logistic regression models with several less restrictive specifications. This also allows us to include observed variables previously captured by the fixed effects, such as urban population (Bairoch et al., 1988; Buringh, 2021) or a location’s diversity of activities. We find that the knowledge of immigrants remains a significant predictor for both entries and exits (see Section 3.4.2 in the supplementary data online). We acknowledge that, over such long periods, travel times are not constant but decrease with improvements in infrastructure and/or technology. Hence, we allow for century-specific effects of spatial proximity, $\rho_{ik,t-1}^M$ and $\rho_{ik,t-1}^E$, leaving our results unchanged (see Section 3.4.3 online). Also, our sample of famous individuals is not balanced over time. Our findings, however, are robust to excluding the twentieth century from the analysis as well as looking at the 20th century alone (see Section 3.4.4 online). Moreover, our results do not change if we redefine entries and exits as the first or last birth of a famous individual with a specific occupation in a location instead of developing or losing specialisation in an activity (see Section 3.4.5 online). In addition, we explore the explanatory power of interaction terms between various relatedness densities on entries, following the literature on migrants as agents of structural change (Elekes et al., 2019; Miguelez & Morrison, 2022; Neffke et al., 2018). We find a significant, but quantitatively
Role of immigrants, emigrants and locals in the historical formation of European knowledge agglomerations

negligible negative interaction term between $\omega_{imm}^{i\text{imm}}$ and $\omega_{eth}^{i\text{eth}}$, indicating that the related knowledge of immigrants and locals are weak substitutes (see Section 3.4.6 online). Also, it may be that our findings of knowledge spillovers are different for different activities. We explore potential heterogenous effects by estimating our regression model separately for aggregate occupational categories (see Section 3.4.7 online). We find, for instance, a stronger correlation of the presence of immigrants specialised in the same activity ($M_{d,t}^{i\text{imm}}$) on entries in sciences and public institutions, and an increased correlation of related knowledge of immigrants ($\omega_{imm}^{i\text{imm}}$) in humanities and sports. Another source of heterogeneity can be city size when size plays a relevant role in generating knowledge spillovers (see Section 3.4.8 online). We find that most entries take place in large cities, and thus, the effects of migration and relatedness are mainly urban. Whereas for exits, we see that spillovers across activities are more important in larger cities. But the probability of exiting an activity in small cities grows massively with the emigration of individuals specialised in the same activity, pointing towards a pronounced role of talent loss in shaping regional specialisations of small agglomerations. This finding relates to the recent literature on left-behind places (Rodríguez-Pose, 2018; Rodríguez-Pose et al., 2023).

Lastly, although the ratio of observed above expected numbers (equations 1 and 6) fulfils the purpose of controlling for size and reverse causality, these models are opaque, not telling us whether our results are driven by changes in the observed or expected number (or both). Hence, we run our main regression model including all terms of the ratio as a robustness check (see Section 3.4.9 in the supplemental data online). We find that the observed number of immigrants with a specific occupation ($N_{d,t}^{i\text{imm}}$) correlates positively with future entries and negatively with future exits, confirming our main results with composite indices. One additional immigrant to a region with a specific occupation correlates with an average increase in the probability of entry by 1.68 percentage points and a reduction in the probability of exit by 5.04 percentage points (see Section 3.4.9 online).

4. DISCUSSION

Labour mobility and migration are core tenets of the United States and the European Union, because policymakers intuit that migrants carry knowledge across space and activities (Cipolla, 1972; Kerr et al., 2017; Lissoni, 2018; Trippi & Maier, 2011; Williams, 2006). Yet, despite multiple studies documenting the role of migrants in the diffusion of knowledge (Bahar et al., 2020; Bahar & Rapoport, 2018; Borowiecki, 2012; Borowiecki & Graddy, 2021; Bosetti et al., 2015; Breschi et al., 2017; Cipolla, 1972; Collins, 1974; Diodato et al., 2022; Elekes et al., 2019; Fassio et al., 2019; Ganguli, 2015; Hornung, 2014; Hunt & Gauthier-Loiselle, 2010; Kerr et al., 2017; Lissoni, 2018; Miguelez & Morrison, 2022; Miguelez & Noumedem Tengoua, 2020; Miguelez & Moreno, 2013, 2015; Mitchell, 2019; Moser et al., 2014; Scoville, 1952a, 1952b; Trippi, 2013; Waldinger, 2010, 2012), there is little historical quantitative evidence of the role of migrants in the historical evolution of knowledge agglomerations.

Here, we used biographical data on more than 22,000 famous individuals – sculptors, composers, politicians, chemists, etc. – living in Europe between the years 1000 and 2000 to explore how the knowledge of immigrants, emigrants and locals explains the probability that a region enters or exits an activity.

Our findings show that migrants play a crucial role in the historical geography of knowledge. Specifically, we find that the probability that a European region enters a new activity grows with the presence of immigrants with knowledge on that activity. Also, using measures of relatedness (Balland et al., 2022; Boschma, 2017; Hidalgo, 2021; Hidalgo et al., 2007, 2018), we find that this correlation is enhanced by spillovers across related activities. Put differently, the probability that a region begets famous mathematicians grows with an excess immigration of mathematicians and with immigrants from related fields, such as physics or chemistry. Similarly, we find that the probability that a European region loses one of its existing areas of specialisation decreases with the presence of immigrants specialised in that activity and in related activities. However, we do not find that locals with related knowledge play the same statistically significant and robust role in entries or exits.

These findings advance our understanding of the evolution of European agglomerations over the past millennium and of the role of migrants and locals therein. Specifically, we find robust evidence that European agglomerations did not only evolve path-dependently (Nunn, 2021), but also that they benefited from spillovers generated by the migration of famous individuals. This supports the literature on the role of migrants in the diffusion of knowledge (Bahar et al., 2020; Bahar & Rapoport, 2018; Borowiecki, 2012; Borowiecki & Graddy, 2021; Bosetti et al., 2015; Breschi et al., 2017; Cipolla, 1972; Collins, 1974; Diodato et al., 2022; Elekes et al., 2019; Fassio et al., 2019; Ganguli, 2015; Hornung, 2014; Hunt & Gauthier-Loiselle, 2010; Kerr et al., 2017; Lissoni, 2018; Miguelez & Morrison, 2022; Miguelez & Noumedem Tengoua, 2020; Miguelez & Moreno, 2013, 2015; Mitchell, 2019; Moser et al., 2014; Neffke et al., 2018; Trippi, 2013; Trippi & Maier, 2011; Williams, 2006) and contributes to the literature on relatedness (Balland et al., 2022; Boschma, 2017; Hidalgo, 2021; Hidalgo et al., 2007, 2018) explaining changes in specialisation patterns (Balland et al., 2019; Balland & Boschma, 2021; Boschma et al., 2014, 2015; Chinazzi et al., 2019; Deegan et al., 2021; Essletzbichler, 2015; Guevara et al., 2016; Juhász et al., 2021; Neffke et al., 2011, 2017; Petralia et al., 2017; Pinheiro et al., 2021; Poncet & de Waldemar, 2015; Rigby, 2015; Ulhich et al., 2022).

Migrants are known agents of structural change enabling the development of unrelated activities (Elekes et al., 2019; Miguelez & Morrison, 2022; Neffke et al., 2018). Our findings differ slightly from that by emphasising migration as a channel of related diversification and path-dependent development, adding to the literature unpacking
the principle of relatedness (Bahar et al., 2019; Boschma & Capone, 2015; Cortinovis et al., 2017; Diodato et al., 2018; Farinha et al., 2019; Jara-Figueroa et al., 2018; Jun et al., 2020; Zhu et al., 2017). Recently, this intersection between evolutionary economic geography, regional diversification and migration has been identified as a promising field of research (Morrison, 2023). We contribute methodologically to this literature by disentangling relatedness measures for immigrants, emigrants and locals. These novel measures make it possible to explore how knowledge spillovers across space and across activities combine (see Section 2.5 in the supplemental data online). Lastly, this study provides a long-term perspective on the evolution of regional specialisations in Europe, a perspective which has been underrepresented in economic geography (Henning, 2019).

Unfortunately, we do not observe the mechanisms explaining the entry or exit of regions in activities. There are, however, several potential mechanisms responsible for these results, which can be subsumed as horizontal and vertical socialisation (Mokyr, 2017, p. 37). For instance, immigrant physicists could teach at a university, leading to a local flourishing of the field of physics and increasing the probability that a famous physicist emerges in the future. Also, immigrant physicists may bring new ideas and approaches with them, which can stimulate creative thinking and cross-pollination of ideas among local scientists in related fields such as chemistry or mathematics. This could lead to the development of new methods as well as new ways of thinking about problems, which could in turn contribute to an increased probability of giving birth to famous chemists or mathematicians in the future. The mechanisms may be different in other activities such as the arts or humanities. The presence of immigrating musicians may create a critical mass of artists, making it profitable to build cultural infrastructure due to economies of scale (Borowiecki & Graddy, 2021), from which artists in related activities such as singers, composers or dancers benefit as well. Shedding light on these different mechanisms is a promising avenue for future research.

Our study has also other limitations. First, we observe only a small and highly mobile subset of the overall population. That is, 22,000 of the most famous individuals living in Europe over the past 1000 years. A more comprehensive dataset would allow for a more accurate and granular estimation of a location’s related knowledge and the geography of activities. Indeed, we suspect that the limited sample is a likely reason for why we do not observe a statistically significant and robust relevance for locals in shaping the historical geography of knowledge. That being said, the related knowledge of locals plays a significant role in several specifications, for instance if estimating the expected number of famous individuals to define specialisations (see Section 3.4.1 in the supplemental data online) or for large cities (see Section 3.4.8 online). Continuing to investigate the role of locals in the historical geography of knowledge can be an interesting avenue for future research.

Second, we do not observe the full migration trajectory of individuals, but only their place of birth and place of death. Although this approach follows the literature (De La Croix & Licandro, 2015; Laouenan et al., 2022; Schich et al., 2014; Serafinelli & Tabellini, 2022) and provides a good proxy of migration (see Section 1.3 in the supplemental data online), more detailed data on where famous individuals lived and when could provide a better analytical basis to explore the evolution of agglomerations (Lucchini et al., 2019; Menini et al., 2017). Indeed, based on a small number of famous individuals living between 1450 and 1750, it is estimated that they moved on average 3.72 times during their lifetime (Mokyr, 2005). Third, we focus only on Europe. So, it may be that the principles behind the historical geography of knowledge uncovered here are different for other parts of the world. Lastly, migration is influenced by multiple factors such as geography and culture (Caragliu et al., 2013; Lewer & Van den Berg, 2008), agglomeration (Borowiecki & Graddy, 2021), patrons (Haskell, 1996/2000; Oevermann et al., 2007) or conflict (Abel et al., 2019; Borowiecki, 2012), evoking reverse causality and endogeneity concerns in our study. We tackled these concerns by using highly restrictive fixed-effects and estimating the expected number of immigrants, emigrants and locals to define specialisations. Despite these efforts, we want to stress that we are not able to make strictly causal claims, a task that can be challenging using historical observational data.

Yet, despite these limitations, our study provides evidence of migration playing a central role in the evolution of European knowledge agglomerations. Also, while being a historical study, our study concerns a topic that is highly relevant in today’s economic policy. The effects of migration on local economies have been debated intensively, both in academia (Borjas, 1994; Card, 2001; Ottaviano & Peri, 2012; Puttermann & Weil, 2010) and in policy circles (OECD, 2021; World Bank, 2019). Our findings add to this debate by showing that the immigration of high-skilled individuals correlates with entering and exiting specialisations of regions. Yet, our results can neither be interpreted causally nor tell us whether these findings remain for migration that is incentivised by policy instruments, since we observe migration involving multiple forces, from forced displacement due to war, to organic forms of migration.

ACKNOWLEDGEMENTS

We thank two anonymous referees, Andrea Morrison, Eva Coll, Jesús Crespo Cuaresma, Ron Boschma, Andrea Belmartino, the attendees of the 2022 Economic Geography PhD school in Utrecht, the attendees of the WICK#10 PhD Workshop in Economics of Innovation, Complexity and Knowledge in Turin, the attendees of the 13th Geoffrey J. D. Hewings Regional Economics Workshop in Vienna, members of the Complexity Science Hub and EcoAustria in Vienna, and the members of the Center for Collective Learning for valuable feedback. Working paper versions of this article have been published in the Papers in Evolutionary Economic Geography series of Utrecht University (PEEG #22.31) and on arXiv (arXiv:2210.15914).
DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

FUNDING

This project was supported by the Agence Nationale de la Recherche [grant number ANR–19–P3IA–0004], the 101086712–LearnData–HORIZON-WIDER-2022–TALENTS-01 financed by the European Research Executive Agency (REA), and the European Lighthouse of AI for Sustainability [grant number 101120237–HORIZON-CL4-2022–HUMAN-02].

ORCID

Philipp Koch https://orcid.org/0000-0002-3248-9686
Viktor Stojkosi https://orcid.org/0000-0002-7594-4431
César A. Hidalgo https://orcid.org/0000-0002-6977-9492

REFERENCES

Abel, G. J., Brottrager, M., Crespo Cuarena, J., & Muttarak, R. (2019). Climate, conflict and forced migration. Global Environmental Change, 54, 239–249. https://doi.org/10.1016/j.gloenvcha.2018.12.003
Almeida-Neto, M., Guimaraes, P., Guimaraes, P. R., Loyola, R. D., & Ullrich, W. (2008). A consistent metric for nestedness analysis in ecological systems: Reconciling concept and measurement. Oikos, 117(8), 1227–1239. https://doi.org/10.1111/j.0030-1299.2008.16644.x
Alshamsi, A., Pinheiro, F. L., & Hidalgo, C. A. (2018). Optimal diversification strategies in the networks of related products and of related research areas. Nature Communications, 9(1), 1328. https://doi.org/10.1038/s41467-018-03740-9
Bahar, D., Choudhury, P., & Rapoport, H. (2020). Migrant inventors and the technological advantage of nations. Research Policy, 49(9), 103947. https://doi.org/10.1016/j.respol.2020.103947
Bahar, D., & Rapoport, H. (2018). Migration, knowledge diffusion and the comparative advantage of nations. The Economic Journal, 128(612), F273–F305. https://doi.org/10.101111/ecoj.12450
Bahar, D., Rosenow, S., Stein, E., & Wagner, R. (2019). Export take-offs and acceleration: Unpacking cross-sector linkages in the evolution of comparative advantage. World Development, 117, 48–60. https://doi.org/10.1016/j.worlddev.2018.12.016
Bairoch, P., Batou, J., & Chèvre, P. (1988). Population des villes européennes de 800 à 1850. Librairie Droz.
Balassa, B. (1965). Trade liberalisation and revealed comparative advantage. The Manchester School, 33(2), 92–123. https://doi.org/10.1111/j.1467-9557.1965.tb00050.x
Balland, P.-A., & Boschma, R. (2021). Complementary interregional linkages and smart specialisation: An empirical study on European regions. Regional Studies, 55(6), 1059–1070. https://doi.org/10.1080/00343404.2020.1861240
Balland, P.-A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European union: Relatedness, knowledge complexity and regional diversification. Regional Studies, 53(9), 1252–1268. https://doi.org/10.1080/00343404.2018.1457900
Balland, P.-A., Broekel, T., Diudato, D., Giuliani, E., Hausmann, R., O’Clergy, N., & Rigby, D. (2022). The new paradigm of economic complexity. Research Policy, 51(3), 104450. https://doi.org/10.1016/j.respol.2021.104450
Bascompte, J., Jordano, P., Melián, C. J., & Olesen, J. M. (2003). The nested assembly of plant–animal mutualistic networks. Proceedings of the National Academy of Sciences, 100(16), 9383–9387. https://doi.org/10.1073/pnas.1633576100
Bastolla, U., Fortuna, M. A., Pascual-Garcia, A., Ferrera, A., Luque, B., & Bascompte, J. (2009). The architecture of mutualistic networks minimizes competition and increases biodiversity. Nature, 458(7241), 1018–1020. https://doi.org/10.1038/nature07950
Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. (2019). Who becomes an inventor in America? The importance of exposure to innovation. The Quarterly Journal of Economics, 134(2), 647–713. https://doi.org/10.1093/qje/qjy028
Bernstein, S., Diamond, R., Jiranaphawiboon, A., McQuade, T., & Pousada, B. (2022). The contribution of high-skilled migrants to innovation in the United States (NBER Working Paper Series w30797). https://doi.org/10.3386/w30797
Borjas, G. J. (1994). The economics of immigration. Journal of Economic Literature, 32(4), 1667–1717.
Borowiecki, K. J. (2012). Are composers different? Historical evidence on conflict-induced migration (1816–1997). European Review of Economic History, 16(3), 270–291. https://doi.org/10.1093/erconhr/hec001
Borowiecki, K. J. (2013). Geographic clustering and productivity: An instrumental variable approach for classical composers. Journal of Urban Economics, 7(1), 94–110. https://doi.org/10.1016/j.jue.2012.07.004
Borowiecki, K. J., & Dahl, C. M. (2021). What makes an artist? The evolution and clustering of creative activity in the US since 1850. Regional Science and Urban Economics, 86, 103614. https://doi.org/10.1016/j.regsciurbeco.2020.103614
Borowiecki, K. J., & Graddy, K. (2021). Immigrant artists: Enrichment or displacement? Journal of Economic Behavior & Organization, 191, 785–797. https://doi.org/10.1016/j.jebo.2020.10.007
Boschma, R. (2017). Relatedness as driver of regional diversification: A research agenda. Regional Studies, 51(3), 351–364. https://doi.org/10.1080/00343404.2016.1254767
Boschma, R., Ballard, P.-A., & Mogler, D. F. (2015). Relatedness and technological change in cities: The rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. Industrial and Corporate Change, 24(1), 223–250. https://doi.org/10.1093/icc/dtu012
Boschma, R., & Capone, G. (2015). Institutions and diversification: Related versus unrelated diversification in a varieties of capitalism framework. Research Policy, 44(10), 1902–1914. https://doi.org/10.1016/j.respol.2015.06.013
Boschma, R., Heinemiks, G., & Ballard, P.-A. (2014). Scientific knowledge density dynamics and relatedness in biotech cities. Research Policy, 43(1), 107–114. https://doi.org/10.1016/j.respol.2013.07.009
Bosetti, V., Cattaneo, C., & Verdolini, E. (2015). Migration of skilled workers and innovation: A European perspective. Journal of International Economics, 96(2), 311–322. https://doi.org/10.1016/j.jinteco.2015.04.002
Breschi, S., Lissoni, F., & Miguelez, E. (2017). Foreign-origin inventors in the USA: Testing for diaspora and brain gain effects. Journal of Economic Geography, 17(5), 1099–1108. https://doi.org/10.1093/jeg/lbw044
Buringh, E. (2021). The population of European cities from 700 to 2000: Social and economic history. Research Data Journal for the Humanities and Social Sciences, 6(1), 1–18. https://doi.org/10.1163/24523666-06010003
Bustos, S., Gomez, C., Hausmann, R., & Hidalgo, C. A. (2012). The dynamics of nestedness predicts the evolution of industrial ecosystems. PLoS One, 7(11), e49393. https://doi.org/10.1371/journal.pone.0049393
