Skilled Migration and Innovation in European Industries*

Claudio Fassio[a, d], Fabio Montobbio[f, c, d], Alessandra Venturini [b, e]

Abstract. This paper studies the effects of skilled migration on innovation in European industries between 1994 and 2005, using the French and the UK Labour Force Surveys and the German Microcensus. We tackle the endogeneity of migrants with a set of external and internal instruments (GMM-SYS). Our results show that highly-educated migrants have a positive effect on innovation and the effect is about one third the one of the skilled natives. The effect of skilled migrant is stronger in industries with high R&D intensity, high FDIs and openness and in industries with higher ethnical diversity at industry level. The estimated coefficient for migrants within Europe is more than twice the one of the non European migrant.

Keywords: Innovation, Migration, Skills, Patents, Europe.
JEL Codes: O31, O33, F22, J61

[*]Acknowledgents:

We would like to thank, for their extraordinary research assistance, Metin Nebiler, Stella Capuano, Laura Bartolini, Davide Cannito, Marco Le Moglie, Sona Kalantaryan; for discussion and suggestions Herbert Brucker, Christian Dustmann, Alessandro Sembenelli, Paula Stephan, Cristian Bartolucci, Agnese Romiti, and Giovanni Facchini. We thank the financial support of the Migration Policy Center, EUI. This paper is a joint research between the Migration Policy Center, EUI and the University of Turin. The paper has benefited from comments and suggestions received during the following seminar and conference presentations: XXIX AIEL Conference, Pisa, September 2014; 5th ZEW/MaCCI Conference on the Economics of Innovation and Patenting, Mannheim, June 2014; Seminar at the Department of Spatial Economics, VU University, Amsterdam, January 2014; NORFACE-TEMPO Conference, Nottingham, September 2013; 3rd SEEK Conference, Mannheim, April 2013; and Seminar at the Department of Economics, University of Trieste, March 2013. The usual disclaimers apply.

[a]: Centre for Innovation, Research, and Competence in the Learning Economy (CIRCLE), Lund University, Box 118, 221 00, Lund, Sweden; email: claudio.fassio@circle.lu.se.
[b]: Department of Economics and Statistics Cognetti De Martiis, University of Turin, Lungo Dora Siena 100 A - 10153 Torino, Italy; email: alessandra.venturini@unito.it.
[c]: iCRIOS, Bocconi University, Via Roentgen 1, 20136, Milano, Italy.
[d]: BRICK, Collegio Carlo Alberto, Via Real Collegio 30, 10024 Moncalieri (Torino), Italy.
[e]: Migration Policy Center, European University Institute, Via Boccaccio 151, 50133 Firenze, Italy .
[f]: Institute of Economic Policy, Università Cattolica del Sacro Cuore, via Necchi 5, 20123 Milan, Italy; fabio.montobbio@unicatt.it (corresponding author).

1. Introduction.

A growing number of studies investigates the impact of high skilled migration on innovation activities. The evidence suggests, with some exceptions, that the impact is generally positive and can be explained by two sets of reasons. First skilled migration contributes directly to research activities and innovation. Using individual data on workers in Science & Engineering (S&E) in the United States, a set of studies shows that total invention increases through the direct contributions of immigrant inventors (Kerr and Lincoln, 2010; Hunt, Gauthier-Loiselle, 2010; Chellaraj et al., 2008; No and Walsh, 2010; Stephan and Levin, 2001). In the US a combination of immigrant policies and self-selection make immigrants on average more educated than natives (Batalova and Fix, 2017), more likely to work in S&E occupations and, possibly, of higher entrepreneurial and inventive ability (Hunt, Gauthier-Loiselle, 2010). Skilled migrants are more likely to be employed than natives in S&E occupations because S&E knowledge is more codified, it transfers more easily across countries and it does not rely on location specific institutional or cultural knowledge, like political sciences, law or medicine (Chiswick and Taengnoi 2007; Sparber, 2010).

A second set of explanations focuses on the impact of ethnical diversity on innovation (e.g. Ottaviano and Peri 2012; Østergaard et al. 2011; Ozgen et al. 2012; Nathan and Lee, 2013; Parrotta et al. 2014; Nathan, 2015). Some evidence suggests that firms with a highly ethnically diverse workforce tend to be more innovative (e.g. Parrotta et al. 2014). Diversity at the firm level enhances innovation because diverse immigrants may provide complementary skills to natives, enhance critical mass and specialization of tasks within the firm and favour knowledge spillovers. Some studies have also analyzed the impact of ethnic diversity on innovation at regional level with mixed results. At the regional level ethnic diversity enhances positive externalities based on cultural diversity and creativity, as well as complementarities in labour market conditions (Ozgen et al. 2012 and 2017; Niebhur, 2010; Nathan and Lee, 2013).

This literature has enabled a deeper understanding of the role of skilled migration and ethnic diversity on innovation, however there is a clear need to consolidate this research field. We do not know, for example, how the industrial composition of the economy affects how migrants contribute to innovation or the extent to which skilled migrants differ from skilled natives in their impact on innovation (in particular in Europe) . Recent papers studying the impact of migration on patents in Europe take a country/region/province perspective (Ozgen et al. 2012 and 2017; Bosetti et al. 2015; Bratti and Conti 2014). However an empirical strategy based on regions and provinces as a unit of analysis is not able to provide information on whether immigrants (in particular skilled immigrants) are really employed in the patenting sectors. In addition the production of patents and the overall innovative processes vary dramatically across industries.

For this reason, in this paper we adopt an industry perspective and focus our study on the manufacturing sectors, since manufacturing accounts for a large share of overall patenting activities. This provides an improvement and a complementary view relative to the existing literature. In our analysis we identify specific conditions under which skilled migration affects innovation activities, exploiting the heterogeneity across sectors. We are able to test the size of the impact of migrant and native workers on innovation conditioning our results to the different characteristics of the manufacturing sectors (in different countries). In particular we focus on the technological intensity of the industry, its international openness and the role of multinational corporations. We also contribute to study ethnic diversity measuring the distribution of different nationalities at the industry level. Studies at the regional level highlight the role of the externalities originating by a diverse environment, which favors complementarity and induces creativity and problem solving. We test whether the analysis of diversity at the industry level strengthens this idea.

In order to address these issues, this paper estimates the impact of migrant (and native) workers on patent production in 16 manufacturing industries, between 1994 and 2005, in France, Germany and UK: the three largest European countries in terms of population and GDP and top destinations for highly skilled workforce in Europe (and in the world after US and Canada; see OECD, 2015). Our paper measures innovation using patents (weighted with forward citations) applied at the European Patent Office. The characteristics of the labour force are based on the Labour Force Surveys in France and the UK and the Microcensus in Germany. Our database allows to fully control for the different characteristics of the labour force, in particular level of education and age, as well as for the country of origin of migrants.

The paper addresses also a number of econometric issues. Demand pull effects on migration at industry level require appropriate instruments. Moreover, there might be a set of additional unobserved factors that affect both patent production and migration at the industry level. Also, the use of Labour Force Surveys can generate measurement errors. Our identification strategy employs longitudinal data at the industry country level and is based on two different instrumental variable strategies: the first relies on the adaptation of the common procedure used in the literature, devised by Card (2001); the second exploits the availability of internal instruments, that is lags in the endogenous variables (system-GMM: Blundell and Bond, 1998). Panel estimations then test whether innovation increases in the three countries as their respective immigrant skilled workforce develops.

The results suggest that native and immigrant skilled workforce increases innovation output with elasticities of respectively 0.3 and 0.09. We discuss the issue of over-education showing that the contribution of skilled migrants is larger in those industries where they are employed in occupations that require the level of education they have attained and we discuss some differences across the industries of our three countries in the impact of skilled migrant on innovation.

Secondly we find that the impact of skilled immigrants varies according to the characteristics of the sectors in which they are employed. Skilled migrants have a stronger impact in the sectors with a high level of R&D intensity. We also find that when a sector has a high level of Foreign Direct Investments (FDIs) the contribution of skilled immigrants is twice as large as the impact in sectors with low FDIs. Skilled immigrants have also a stronger effect on innovation in sectors that are more open to international trade.

Lastly we control for the role of ethnic diversity and show that the impact of skilled immigrants is higher in sectors with a higher level of ethnical diversity. In addition the positive effect on innovation is stronger for European migrants than for third-countries nationals.

The article is structured as follows. The next section explains the conceptual framework, surveys the recent literature and presents the research hypothesis. The third section introduces the empirical model, the data set and discusses identification issues. The fourth section presents the results. The final section presents the conclusion, discusses some policy implications and makes suggestions for further research.

2. High Skilled Migration and Innovation

High skilled migration increased rapidly in the OECD countries in recent years. A recent OECD report shows an increase of 13 million in the total number high skilled migrants from 2001 to 2011, reaching a total number equal to 31 million

(OECD, 2015). In addition the share of the highly educated among all migrants rose to 30% in 2010/11 from 25% in 2000/01. This is due to an overall growing number of tertiary educated persons worldwide and to a relatively higher propensity to migrate of skilled workers. The recent literature has studied whether high skilled migration stimulates innovation activities in destination countries[1]. Typically two sets of explanations are put forward. The first one argues that skilled migration contributes *directly* to research activities and innovation (e.g. Kerr and Lincoln, 2010; Hunt, Gauthier-Loiselle, 2010; Chellaraj et al., 2008; No and Walsh, 2010; Stephan and Levin, 2001; Bosetti et al. 2015) and the second one underlines the role of ethnic and cultural diversity (e.g. Ottaviano and Peri 2012; Østergaard et al. 2011; Ozgen et al. 2012; Niebhur, 2010; Nathan and Lee, 2013; Nathan, 2015; for a survey Kemeny, 2017).

2.1 The direct impact of skilled migration on innovation

This literature is mainly based on the US experience where in 2015 the share of tertiary educated is 48% among migrants and, at the same time, 30% among US-born population (Batalova and Fix, 2017). In addition skilled migrants are more likely to work in S&E occupations and display higher entrepreneurial and inventive abilities. Skilled migrants provide improved knowledge assets and greater availability of skills (otherwise unavailable) and this is generated by a combination of immigrant policies (e.g. temporary visa such as H-1B) and self-selection on superior technical and engineering skills. The empirical literature suggests that the overall effect on innovation is indeed positive because there is no crowding-out effects on natives.

In line with this literature the first objective of our paper is to test for Europe the direct impact of skilled migration on innovation and compare it to the impact of natives. In the European context we focus on France, Germany and the UK that are the most important destination countries. The European context has also important specificities because a substantial amount of the international mobility is within the continent and, in parallel, European countries have the lowest share of high skilled immigrants over total immigrants (De la Rica et al. 2014; OECD – DIOC database). In addition, compared to the US, immigration policy is less selective on technical and engineering skills. Nonetheless net high skilled migration rates remain positive for the UK, France, and Germany; in UK and France immigrants are on average more educated than the natives (De la Rica et al. 2014). In this context, it is difficult to say ex-ante whether the direct innovation impact of skilled migration in Europe is positive, or whether it is stronger than the one of the natives. To our knowledge our paper is the first attempt to analyze the impact of high skilled migration on innovation at the industry level. Some studies have provided evidence on European countries using individual data on inventors or aggregated data at regional or country level. The evidence is mixed although in the majority of cases tends to suggest a positive effect[2] in particular for France, Germany and UK which are the most attractive countries in the EU for high skilled workers.

[1] The number of papers studying migration and innovation increases rapidly. A set of recent excellent surveys provides a full coverage of the topic: Breschi et al. (2016); Lissoni (2017); Rashidi-Kollmann and Pyka (2016); Kemeny (2017); Kerr (2016).

[2] Bosetti et al. (2015), using a panel of twenty European countries, find that skilled migrants contribute positively to the number of patents and citations of scientific publications. Gagliardi (2015) finds that the share of skilled migrants within a UK province has a positive impact on the innovative performances of firms in that specific province. At the same time Ozgen et al. (2012) find that in a sample of 170 European regions the share of immigrants does not lead to a higher number of patent applications. Using data on Italian provinces, Bratti and Conti (2014) do not find that skilled migration has any effect on patent production. Zheng and Ejermo (2015) using individual data on Swedish foreign-born inventors confirm that the positive effect of skilled migrants in Europe is less clear-cut, since they find that in Sweden immigrant inventors do not outperform natives in terms of the number of patent applications submitted.

H1: the contribution of skilled migrants, at industry level, to innovation in France, Germany and UK is positive. It is stronger than the one of the natives if a self-selection and an education effect prevails.

An additional factor that could play an important role in the overall contribution of skilled migrants to innovation is over-education, or skill mismatch. As shown by the existing literature, skilled migrants are more likely to be employed in occupations that require a lower level of education than the one that they have attained, with respect to natives (Chiswick and Miller, 2010; Dustman and Glitz, 2011) This applies also to European labor markets (Nieto et al., 2015) and is often due to the limited portability of their human capital, or to discrimination in the host country labor markets. Overeducation[3] typically weakens the contribution of skilled immigrants to firms' productivity (Mahy et al., 2015) and, in the case of skilled immigrants, is also likely to weaken their contribution to the innovative performances of the companies for which they end up working. Accordingly it is reasonable to assume that the level of overeducation will have a negative impact on the impact of skilled immigrants on innovation:

H2: the contribution of skilled immigrants to innovation is higher in sectors characterized by a low level of overeducation relatively to sectors with a high level of overeducation.

We also exploit sectoral heterogeneity to study the conditions under which it is more likely to observe a positive impact of skilled migration on innovation. In particular we consider (i) R&D intensity, the degree of firms' internationalization in terms of (ii) international trade and presence of (iii) Foreign Direct Investments (FDIs).

In the US skilled migrants are more likely to be employed than natives in occupations related to Science and Engineering (S&E). S&E knowledge is more codified and transfers more easily across countries and it does not rely on location specific institutional or cultural knowledge, like politics, law or medicine. We expect therefore that the impact of skilled migrant on innovation is stronger in high tech sectors (where S&E knowledge is more important) relative to other fields where it's key to master the local language or have a larger social capital[4] (Chiswick and Taengnoi, 2007; Sparber, 2010).

H3: the contribution of skilled immigrants to innovation is higher in sectors characterized by a high level of R&D intensity relatively to sectors with a low level of R&D intensity.

The contribution of foreign workers to the innovative process also depends upon the level of international involvement of the companies they work for. Companies that are active in international markets, either through the establishment of

[3] Here overeducation means only "vertical" overeducation, i.e. a worker has a level of education that is too high for the competences required by the worker's own job. On the contrary we do not focus on horizontal overeducation, i.e. working in an occupation that would require a different type of educational field, since this phenomenon is less relevant for immigrant workers (Nieto et al., 2015). In addition However, as emphasized by Hanushek and Woessmann (2011) what matters is the quality of education and the cognitive skills. Measuring skills with the years of schooling is not precise also if lower educated workers are high-skilled in certain specific tasks or activities.

[4] A possible explanation can also take into account the characteristics of the knowledge base. If the knowledge base is more "analytical", i.e. formal, codified, close to scientific research, it builds also on a larger epistemic network that is by its own nature more global (Knorr-Cetina, 1999; Martin, 2013). International movement of skilled labour gives a greater contribution to innovation because there is a larger and common knowledge base. Alternatively if the knowledge is more tacit, hence more embodied in individuals and in locally established routines among "communities of practice", the contribution of foreign workers to the innovative process might be smaller, due to a lack of knowledge related to local markets and networks.

foreign subsidiaries or through export and import operations, are more likely to adopt management practices that can boost the contribution of individuals with a different cultural background (Breschi et al., 2017). Multinational companies can exploit the innovative potential of foreign workers for several reasons. First of all, their organizational structure is by definition geared towards exploiting the benefits of a culturally diverse labour force. The existing literature on the "global mindset" of management in multinationals stresses precisely the importance of using cultural diversity for competitive advantage (Bartlett and Ghoshal 1989; Levy et al, 2007). Hence with respect to domestic firms the management of multinational companies is likely to have a higher ability to exploit the innovative potential of employees from different nationalities. Moreover multinational companies often rely on the international mobility of their employees (assigned expatriation) in order to transfer knowledge trans-nationally within their network of subsidiaries (Edström and Galbraith, 1977; Minbaeva and Michailova, 2004; Caligiuri, 2016). Foreign expatriates can hence contribute better to the innovative processes of these companies, because they provide specific bits of knowledge that are strategically important for the development of new products and processes in that specific subsidiary. Finally multinational companies, thanks to their status and higher productivity, are also better able at recruiting talents internationally: it is hence likely that they recruit high quality foreign workforce which ultimately leads to a higher innovative contribution. According to this view the presence of FDIs is likely to increase the probability that foreign workers can contribute to the innovation process. The presence of FDIs is not homogeneously spread across the different sectors of the manufacturing, hence we expect that:

H4: the contribution of skilled immigrants to innovation is higher in sectors characterized by a high presence of FDIs relatively to sectors with a low level of FDIs.

Similarly, also companies that compete in international markets through the export of their products and through participation to international value chains are likely to be better able at benefiting from the innovative contribution of foreign employees. First of all, as shown by the literature on global production networks (Sturgeon, 2002), firms involved in international value chains typically employ international technological standards with high levels of modularity which allow them to easily interact with different foreign partners, hence they are likely to use platform technologies that are well known also outside of the national borders and to which also foreign workers can contribute actively. Moreover for these companies the need to interact with foreign customers and foreign suppliers typically leads to organizational structures and management practices that enhance the contribution of individuals from different cultural backgrounds. Finally, as shown by recent empirical evidence: immigrant employees positively impact the export performances, in their own country of origin, of the companies they work for (Hiller, 2013; Lodefalk, 2016). The level of internationalization of firms varies a lot across sectors, with some sectors having a very high level of openness to trade and some others seeing a large prevalence of domestic firms. We hence expect that:

H5: the contribution of skilled immigrants to innovation is higher in sectors characterized by a high level of openness to trade relatively to sectors with a low level of openness to trade.

2.2. Ethnic Diversity and Innovation

There is large literature that studies the impact of cultural and ethnic diversity on innovation[5]. Ethnic diversity is typically used as a proxy of cultural diversity and measured at the firm or regional level. Diversity at the firm level enhances innovation because it broadens the firm's knowledge base, allows for new knowledge combinations, enhancing problem solving and the generation of new ideas (Østergaard et al. 2011; Parrotta et al. 2014; Kemeny and Cooke, 2017; Kemeny, 2017). At the same time diversity can create barriers to interaction and communication and lead to conflicts and lack of common action. If competences and experiences are very much disconnected, innovative learning is also more difficult[6].

At the regional level ethnic diversity could enhance positive externalities based on cultural diversity and creativity, complementarities in the labour market conditions and demand. A growing set of papers analyses the European context with mixed results (Niebhur, 2010; Ozgen et al., 2012; Bratti and Conti, 2014). In these papers the effect of skilled migrants is typically measured on the innovative performance of the country/region/province in which they are resident. However, a geographical approach cannot distinguish between the effect of immigrants that directly contribute to innovation because they work in innovative sectors and the effect of immigrants that work in other sectors in the same region. For this reason we provide a different view analyzing diversity at the industry level.

In the specific European context we can also assume that skilled immigration within the EU could be more effective. Within the EU we expect less institutional "frictions" in the labour market that can be related to factors such as the type of contracts. Especially in the case of international mobility, frictions can be related also to the specific inter-governmental agreements on the freedom to move and work across nation states. The gradual implementation of the freedom of movement for workers in the European Union (which in the early 90's already involved all citizens of the EU-15) is likely to provide an easier job-match for EU nationals with respect to non EU individuals. So we expect that within the EU skilled labor mobility produces a better match of jobs and task specialization and a process of learning by hiring that fills high skilled labor shortages in specific sectors.

H6: the contribution of skilled immigrants to innovation is higher in sectors characterized by high level of cultural diversity relatively to sectors with a low level of cultural diversity

H7: the impact of skilled migrants on innovation is higher if their country of origin is within the EU

3. Model, Methodology and Data

3.1 Model

Unlike the previous literature that uses mostly country, regions or provinces, our unit of analysis is the manufacturing sector. Our empirical model adapts Furman *et al.* (2002) that studies the innovative capacity of countries. According to standard endogenous growth models (Romer, 1990) the rate of technological progress is given by:

$$\dot{A}_t = \delta(A_{t-1}^{\beta}\, \mathrm{H}_{\mathrm{t\text{-}1}}^{\gamma}) \qquad (1)$$

[5] We do not consider here the large literature on diversity and productivity (see Kemeny, 2017; Fassio et al. 2017).

[6] The dominance of specific ethnic groups or specific forms of ethnic segmentation can be particularly problematic (Alesina and La Ferrara, 2005; Caselli and Coleman, 2013)

The sustainable rate of technological progress at time t ($\dot{A}_t$) depends upon the stock of accumulated knowledge A_{t-1} and by an ideas generation input (H_{t-1}), which operates according to a standard Cobb-Douglas production function. This particular specification assumes some complementarity between inputs, so that the marginal impact on innovation of the inputs increases in the level of all of the other factors. At industry level expanding Eq. (1) we obtain:

$$\dot{A}_{it} = \delta(A_{it-1}^{\beta}\ R\&D_{it\text{-}1}^{\gamma}\ L_{it-1}^{\varphi}\ \mathrm{X}_{it\text{-}1}^{\theta}) \tag{2}$$

We test whether the annual flow of patents ($\dot{A}_{i,t}$) (weighted by citations) in year t and sector i is explained by lagged yearly expenditures in Research and Development ($R\&D_{i,t-1}$) and a lagged measure of the openness to trade of a specific sector ($X_{i,t-1}$) that is the volume of exports plus imports per unit of production in sector i at time $t-1$. This controls for aggregate competitiveness effects. Following equation (2), we also control for the stock of patents in the previous year ($A_{I,t-1}$), which measures the stock of prior ideas and prior research[7]. The main focus of the paper is on the role of human resources in innovation. We use the lagged human capital characteristics ($L_{i,t-1}$) in that specific sector i. It is important to underline that we decompose the human capital variable by age, education and ethnicity. In doing so we assume imperfect substitutability of different labour factors as in Ottaviano and Peri (2012).

The dependent variable is the number of forward citations received by the patents in the four years after the priority date.[8] We model our production function as a Cobb Douglas and we take logs to estimate the elasticity of each of the different inputs. We lag each independent variable by one year[9] as follows:

$$\ln \dot{A}_{it} = d + \beta \ln A_{it-1} + \gamma \ln R\&D_{it-1} + \sum_{k} \phi_k \ln L_{it-1}^{k} + \theta \ln X_{it-1} + \alpha_i + \lambda_t + \varepsilon_{it} \tag{3}$$

The employment L is divided into k different components, according to ethnicity, education and age; α_i is the time-invariant fixed effect of each sector, λ_t denotes a common time trend (that we proxy with time dummies) and ε_{it} is the idiosyncratic shock occurring at time t in sector i. The analysis covers sixteen industries (two digits NACE[10]) in the manufacturing sector, from 1994 to 2005 and three countries: France, Germany and the UK. As a consequence subscript i refers to the country-sector pair, which is our observational unit in the panel. Table (1) provides the list with the definition of the variables.

[7] We take into account the rate of decay of the accumulated knowledge and we use the perpetual inventory method. We have patent data from 1986 onwards, that is 8 years before the beginning of the period under analysis (1994). We use data before 1990 to calculate the average growth rate in each sector, then we divide the levels in 1990 by the sum of the average growth rate between 1986-1990 and the depreciation rate, in order to build the first stocks in 1990. A standard assumption in the literature is that the depreciation rate used to construct the measured stock is between 10% and 20%. We have built different stocks using a depreciation rate of 10%, 15% and 20% (for a discussion see Pakes and Schankerman, 1984, and Hall and Mairesse, 1995). We have built also the patent stock without depreciation as in Furman et al. (2002). In the regressions below we display the results with a 10% depreciation rate, but results do not change using the knowledge stocks with different depreciation rates.

[8] We use the number of forward citations received by each patent, instead of the simple number of patents, in order to select only patents with economic value (for a thorough explanation see Section 3.3)

[9] We acknowledge that the lag could be longer, but considering that we are using the priority date of patents, and that the R&D and labour force time series are quite persistent, we believe that one lag is a correct compromise in order to maintain a sufficient number of observations.

[10] The 2-digits level of sectoral aggregation was preferred to 3 or 4 digits in order to have a statistically reliable measure of the number of migrants (especially highly educated ones) in all sectors. Indeed lower levels of sectoral disaggregation would risk in increasing errors in the measurement of migrants, considering that already at the 2-digit level the number of migrants in some sectors is quite low.

Equation (3) assumes that the coefficients ϕ_k are the same in all sectors. In order to test the hypothesis outlined above (from H2 to H6), we remove this restriction and allow these coefficients to vary according to a set of industry level variables (over-education, R&D intensity, FDIs, openness to trade and diversity).

[Table 1 about here]

3.2 Identification strategy

In order to estimate equation (3) we need to address a number of econometric issues that might affect our coefficients of interest. Our main concerns are directed towards the correct identification of the effect of the labour variables, and in particular of migrant workers, on patent production. The decision to move to a specific country is, in most of the cases, a strategic decision that depends on the specific dynamics of the sectors in which migrants will work. This demand-pull effect, if not accounted for, is likely to affect our estimates. Moreover, it is likely that patent productivity shocks in a given sector have differentiated effects according to workers' skills and education. Indeed, an increase in the overall number of patents in a sector indicates a gradual shift of firms towards higher levels of technological sophistication. According to the vast literature on biased technological change (Acemoglu, 2002) technical change is more likely to exert a positive effect on the demand for educated workers, while it might have a negative effect on the demand for unskilled ones. In this respect the choice to lag by one year all the independent variables in equation (3) represents a first step in addressing this problem. But it is not likely to solve it completely.

A second problem is generated by other unobserved factors, which might affect both patent-productivity at the sectoral level and the decision of migrants to move to a specific national sector. For example a high-tech multinational that starts a green field investment in a given country is likely to affect both the production of new patents in a given sector and the flow of skilled migrants that come to work in that same sector. Again these factors would lead to problems of omitted variables bias due to both time-invariant and time-varying unobserved heterogeneity. Finally possible measurement errors may appear in the number of migrant workers. The use of Labor Force Surveys data should allow us to take into account sampling errors, through the use of population weights. However, the probability of incurring random measurement errors in national statistics on the labor force is not irrelevant, especially for data on migrant workers. This might lead to attenuation bias problems in the estimation of the coefficients of interest (Aydemir and Borjas, 2011).

To address these issues, our starting point is a fixed-effects Ordinary Least Squares estimation that accounts for time-invariant unobserved heterogeneity denoted by α_i in equation (3). However the fixed effects estimator is consistent under the unrealistic assumption of strict exogeneity between the covariates and the sector-specific idiosyncratic productivity shock ε_{it}. Wintoki *et al.* (2012) focus specifically on the direction of the bias of the fixed effects estimator when strict exogeneity is violated and find that when the explanatory variable is negatively correlated with past values of the dependent variable the fixed effects estimator will have an upward bias. A positive correlation of the explanatory variable with past shocks of the dependent variable will, conversely, lead to a downward bias in the fixed effects estimator. In the case of patents the demand for educated workers is positively correlated with past shocks of patent productivity, while the opposite might occur for unskilled workers. Therefore, we expect a downward bias of the fixed effects estimator for educated workers and, possibly, an upward bias for unskilled workers.

In addition fixed effects estimators fail to account for the unobserved factors that might occur during the period of observation (as in the example of multinationals' brand new investments) and which might also induce a bias in the coefficients.

In order to address these issues, we implement two different instrumental variable strategies: the first relies on the use of external instruments, according to a common procedure used in the literature and first devised by Card (2001), while the second exploits the availability of internal instruments, that is lags in the endogenous variables. We implement both strategies since they have advantages and drawbacks: the use of external instruments is well suited to our empirical setting, but it relies on specific behavioural assumptions by immigrants which may or may not apply. On the contrary the use of internal instruments does not require these specific assumptions. Rather, it is better suited to large samples with a high number of observations.

External instruments

The first instrumental variable strategy relies on the well-known identification strategy first implemented by Card (2001). This methodology takes advantage of the fact that migrants of a certain nationality tend to move to locations where other people of the same nationality had already settled. Therefore, by using the original distribution of nationalities at the beginning of the period of observation and the exogenous migration flows, it is possible to create fictional flows of migrants to be used as external instruments. This is possible because these flows are strongly correlated with the endogenous stocks of migrants, and, at the same time, they are also uncorrelated with the shocks of the dependent variable. Indeed the aggregate flows of immigrants from specific countries of origins are unlikely to be correlated with specific sectoral shocks in the destination country. For our empirical design we do not exploit the fact that migrants tend to move to areas where people of the same nationality have already settled, as in Card (2001). Rather, we take advantage of the fact that migrants often work in the same economic activities in which their compatriots are already active. The validity of this identification strategy rests on the hypothesis that the network effect, or better the effect of the "migratory chain" on the new inflows of migrants is not only limited to location effects. Rather, the "migratory chain" extends also to the sector of employment. Indeed the community of origin acts as a placing agency, reducing the cost of finding a job in the sectors in which the migrants from a specific country of origin are already concentrated (Ellis and Wright, 1999). Frequently job engagement is already found before the arrival of the co-nationals.

For skilled and unskilled migrants, we create these "supply-push" levels of migrants workers in each sector (see the Appendix A1 for the details of the computation). The procedure allows us to create a new variable both for skilled and unskilled migrants: these two new variables are likely to be exogenous to the patent productivity shocks. These new stocks are used as external instruments for the real stocks of high and middle-low educated migrants in equation (3) in an IV setting with a two-stage least squares estimator. If our hypotheses hold these "supply-push" stocks should be correlated with the actual stocks of migrants in each sector; but at the same time they should not be correlated with the patent shocks.

Internal instruments

The second instrumental variable strategy relies instead on the use of internal lags in the endogenous variables as suitable instruments: we use the Blundell and Bond (1998) GMM-SYS estimator. The GMM-SYS estimator accounts for the violation of the strict-exogeneity condition, which can greatly affect the reliability of fixed effects estimates. Moreover, differently from the exactly-identified Card-like IV strategy based on external instruments, the GMM-SYS estimator allows us to test for the exogeneity of the instruments, since the use of several lags in the endogenous variables allows for an over-identified specification. Finally the GMM-SYS allows us to instrument, as well, the number of skilled and unskilled native workers, since these variables are, also, likely to be endogenous.[11]

In equation (3) we consider the labour variables (both migrants and natives) as endogenous, that is, correlated with past and present values of the error term, while we consider all the other control variables as exogenous.[12] We will then estimate equation (3) instrumenting the endogenous variables L^k with their own lags. A possible shortcoming of the GMM-SYS estimator is that it is better suited for large samples of individuals, while in our sample the number of sectors in the three countries is limited. This may lead to the problem of instruments over-fitting (Roodman 2009), due to the high number of instruments with respect to the number of observations, which decreases the reliability of the Hansen test on the exogeneity of the internal instruments. For this reason, in our estimates, we limit, as much as possible, the number of lags used as instruments, employing only those that are most informative. Furthermore, we implement the procedure suggested by Roodman (2009) in order to reduce the overall number of instruments, by collapsing, into one single instrument, all the lags used as instruments.

Lastly, the adoption of internal instruments is also able to address the problems related to measurement errors. Indeed, if measurement error is free of serial correlation (and we believe this would be the case in our context), the panel dimension of the data deals with attenuation bias, precisely because it provides internal instruments. Griliches and Hausman (1986) show that the use of fixed effects (within estimators) can amplify the problems due to measurement error in panel studies. They also show that the best strategy to overcome this problem is the use of internal instruments.[13]

3.3 Data

We take advantage of an original dataset which combines data on innovation, as proxied by patents, and information on the characteristics of the labour force (migration, age and education) at the sectoral level. Measuring innovation and technical change is a daunting challenge since innovation is a multi-faceted phenomenon and knowledge creation does not always leave a paper trail. One of the most popular indicators of innovation is the number of patents applications at industry or country level (e.g. Furman *et al.* 2002; Malerba et al. 2013). We use patent applications at the European Patent Office (EPO) because we analyse three European countries. In addition international patent applications at the EPO are costly and, therefore, we select inventions with relevant market potential (Deng, 2007)[14]. Finally we use an international

[11] We use the GMM-SYS estimator instead of the Arellano and Bond (1991) GMM-DIF, which also uses lags of the endogenous variables as suitable instruments, because labour variables are usually quite persistent. When time series are persistent the GMM-SYS specification is to be preferred (Blundell and Bond, 1998). See also Appendix A2 for a detailed discussion.

[12] We also considered relaxing this restriction, treating also the controls as endogenous. See later in Section 4.2., footnote 24.

[13] We acknowledge that in the case in which measurement errors displayed instead some serial correlation, i.e. they are not random, the advantage of the GMM-SYS estimator would be less significant in this respect.

[14] Patent indicators have many limitations that have to be taken into account. Many inventions are not patented. Even if patents are increasingly used by companies, the evidence provided by many surveys of R&D managers indicates that, in many sectors, patents are not considered the major source of profit from new products and processes (e.g. Cohen *et al*., 2000). This depends upon the nature of the technologies. As a consequence, companies have a significantly different propensity to patent across different sectors of economic activity. Finally, like R&D measures, patents tend to be a better proxy for the technological activities of large firms. Small firms tend to have a lower propensity to patent because – all other things being equal – the use of intellectual property requires high fixed costs of

patent office to offer a homogeneous database which allows cross-country comparisons and is less distorted by country-specific institutional or policy changes.

The technological and economic value of patents varies considerably and many patents have low economic and technological value, while a few of them are extremely valuable. Patent citations are then used to correct this problem and to measure the economic and technological value of a patent. The dependent variable is the number of citations received by the patents applied at the EPO in the four years after the application. For all three countries the complete details of patents and patent citations are derived from PATSTAT (see Appendix B). The conversion of the International Patent Classification to NACE sectors is provided by Schmoch *et al.* (2003). Almost all patents are assigned to the manufacturing sectors and are assigned to countries using the address of the inventors and fractional counting.

The information concerning human capital (level of education, country of origin, age) is retrieved through the aggregation at the sectoral level of micro data from the national Labour Force Surveys for the UK and France and from the Microcensus in Germany. Since we are only interested in individuals with an occupation, we only included in our sample employed individuals (both for immigrants and natives). For the UK and French data we have used the country of birth to identify immigrant workers. For the German data instead immigrant status is defined as the holding of a non-German citizenship, since this is the only information available in the Microcensus. because of the relatively low naturalization rate, due to the long eligible period to become a German citizen and to the limits on the holding of double nationality, the discrepancy in the definition of immigrant status with France and the UK is not likely to affect the analysis. Appendix B provides an extensive description of the data. R&D expenditure and trade data by sectors are provided by the STAN database (OECD), data on Foreign Direct investments is provided by the OECD Dataset on Inward Activity of Multinationals by Industrial Sector. The list of the countries of origin of the migrant workforce is in Appendix A5. Table (1) provides the full list of the variables.

4. The empirical analysis

4.1 Descriptive Evidence

Table (2) and (3) display the main characteristics of the database in the three countries in two sub-periods at the beginning and at the end of the period considered (1994-2005): the number of patents and citations *per* worker, the share of immigrants and the share of workers aged 35 years or younger (40 for Germany). Table (2) refers to all manufacturing sectors, Table (3) shows the data just for high-tech sectors (See the Table A3 in the Appendix for the classification of sectors). Patents and patent citations *per* employee are higher in the high tech sector. The number of citations decreases substantially in the second period due to the right-end truncation bias[15]. In the manufacturing sectors considered the share of young workers remarkably decreases over time, mainly because of the decreasing share of young workers among the non-tertiary educated. The share of tertiary educated, instead, is on the rise particular in the UK and France.

implementation and scale (e.g. Patel and Pavitt, 1994). It follows that the size distribution of firms may have an important effect on the aggregate count of patents at the national level.

[15] Recent cohorts of patents are less likely to be cited then the older ones, because the pool of potentially citing patents is smaller. See Bacchiocchi and Montobbio (2010) for the analysis of the truncation bias in patent citations in different patent offices.

The overall share of immigrant workers in manufacturing sectors is falling slightly in Germany, where it is about 12% of the overall employment; on the contrary, in the UK and France the share of immigrants increases, respectively, from 6% to almost 8% and from 2% to 4%. Note that the share of tertiary-educated immigrants is growing in all countries: in the UK, where it already accounted for 1.2% of the labour force in 1994-1996, it doubles during the period of observation and reaches 2.4% in 2003-2005. Also in France, where the shares of highly-educated migrants are substantially lower (0.3% in 1994-1996), the percentage doubles reaching 0.7% in 2003. In Germany the growth is slightly less high (from 0.7% to 1.1%).

[Table 2 and 3 about here]

Table (2) shows an increase in the number of EU27-nationals immigrants in France and the UK. In the UK this is primarily due to the growth of tertiary educated EU27-nationals (mainly young highly-educated immigrants from Eastern Europe). In Germany instead the share in EU27-nationals is quite stable over time, but there is an increase in the share of tertiary educated EU-nationals.

In Table (4) we show the number of patents and citations *per* employee, as well as the share of immigrants, broken down by sectors. It highlights once more the great heterogeneity in the production of patents at the sectoral level: high tech sectors like Office, Accounting and Computing Machinery display more than 10 patents for 1000 workers, compared to 0.2 in the Textile sector. The share of immigrant workers is high in the Textile and Automotive sector, but it mainly consists of low and middle educated workers. On the contrary tertiary-educated immigrants are more numerous in Office, Accounting and Computing Machinery, as well as in the Chemicals and Pharmaceutical sectors and Radio, Television and Communication Equipment. The share of European Union workers is quite constant across all sectors (around 3-4% of the labor force); on the contrary, the share of tertiary-educated EU nationals is substantially higher in all high tech sectors.

[Table 4 about here]

4.2 Econometric results

Table (5) reports the descriptive statistics for each of the variables used in the estimations. We have sixteen two-digit sectors for twelve years in France (1994-2005) fourteen two-digit sectors for twelve years in UK (1994-2005) and sixteen two-digit sectors for ten years in Germany (1996-2005). [16] In Table (6) we turn to the estimation of equation (3) using data

[16] For the UK we lack data on R&D expenditures in two sectors (Manufacture of wood products and cork; Manufacture of paper and paper products) therefore we can only apply our model to fourteen sectors. Our original sample consists hence of 520 observations: 192 observations in France, 168 observations in UK and 160 observations for Germany. In the estimation we use one year lag and therefore we lose sixteen observations in France and Germany and fourteen in the UK (46 overall), which correspond to the first year of each time-series. Furthermore, in France, in the first years of observation in some sectors with a low number of employees (Wood and products of wood and cork, Paper and paper products, Office Accounting and Computing Machinery) there are no foreign workers at all, so we cannot retrieve information on the average age of foreign workers: therefore, we lose fifteen observations in France. This also happens (for only one observation) both in UK and Germany. Overall, and obviously discounting the 'lost' observations, we have 161 observations for France, 143 observations for Germany and 153 observations for the UK, which sums up to 457 observations that will be used in our estimates.

from all countries, including time dummies to account for the common time trend. All variables are in logs and each covariate is included with a lag of one year. Our specifications include controls for openness to trade, expenditures in R&D, and the cumulated stock of patents.

[Table 5 about here]

In Table (6) we measure the effect of all the labour force (*E*) together and then we distinguish between tertiary educated (*E_Tedu*) and low-middle educated workers (*E_noTedu*). The GMM-SYS estimators in columns (2) and (4), which properly account for the possible endogeneity of the labour force, show that the coefficient of all those employed is negative and significantly different from zero. In column (4), when we distinguish between high and low educated workers however, we find that, as expected, the two have different effects on innovation: tertiary educated workers display positive and significant effects, while middle-low educated workers have a negative and significant effect. With respect to the other control variables in the model, the results show a negative effect on the average age of workers, especially for non-educated workers, and positive and significant coefficients for R&D expenditures and the stock of knowledge: the openness to trade is, meanwhile, negative and significant. This result is composed of a negative effect of import intensity, which is a signal of low competitiveness, and conversely of a positive effect of export intensity, which indicates higher competitiveness.[17] The AR(1) and AR(2) tests confirm the goodness of our model specification,. moreover, the heteroskedasticity-robust Hansen test fails to reject the null-hypothesis of strength and exogeneity of the lagged instruments in use.

[Table 6 about here]

In Table (6) we also report the coefficients obtained with fixed effects estimators in columns (1) and (3) to check whether the direction of the bias of these estimators is consistent with our expectations. According to Wooldridge (2002) and Wintoki *et al*. (2012) we should expect a downward bias for educated workers (positive correlation with past shocks of the dependent variable) and an upward bias for unskilled ones (negative correlation with past shocks). Indeed, looking at the results of column (5) we find that the fixed effects estimator displays a downward bias in the coefficient of educated workers, with respect to the GMM-SYS estimates, and an upward bias in the coefficient of non-educated workers.

In Table (7) we distinguish between the native and immigrant workforce and within each of these subsets we discriminate between tertiary educated and non-tertiary educated employees. Our specifications include time dummies, all the additional controls (R&D expenditures, stock of citations, openness to trade) and, finally, the average age of each of the four identified groups of workers (highly-educated natives, highly-educated immigrants, low educated natives and low educated immigrants): all the coefficients of the control variables are reported in the Appendix in Table (A6). In Table (7) we only show the estimated coefficients of the labour variables. In column (1) we report the coefficients obtained with a

[17] Results are available from the authors upon request.

fixed effects estimator: the estimated coefficients of the labour variables are likely to be affected by endogeneity, therefore we report them only as a benchmark. In columns (2) and (3) we show the results of a Two Stage Least Squares (2SLS) instrumental variable estimation in which we use, as external instruments, the supply-push stocks of high and low educated migrants, following our modified version of Card (2001). We first instrument only middle-low educated migrants (*E_noTedu_mig*) with their supply-push stocks and then we instrument only highly-educated migrants (*E_Tedu_mig*) with the supply-push stocks of highly-educated migrants. The results in column (2), in which we instrument only middle-low educated migrants, show that this category of migrants has a negative and significant effect. As in Table (6) when we instrument non educated workers we find the coefficient becomes even more negative, in line with the hypothesis of an upward bias in fixed effects estimates (Wintoki *et al.*, 2012).[18] In column (3), instead, we adopt the same specification, but this time we instrument the highly-educated migrants with their supply-push stocks: in this case the predictive power of the instrument is extremely low, contrary to the case of low educated migrants we cannot rely on this identification strategy for this category of migrant workers: the instrument is extremely weak. We interpret these results as an empirical test of the behavioural assumptions behind our estimation strategy: while for low-educated workers it seems that the presence of immigrants from a certain country in a given sector attracts new migrants from abroad to the same sector, in the much more recent and lower scale case of highly-educated workers this is not the case. For highly-skilled migrants the market signals are more efficient than ethnic networks. Indeed, Card's strategy is originally devised to account for the behaviour of mainly low-skilled migrants entering the United States (in Card's study immigrants were mostly Hispanics from Mexico and South America and had, on average, two or three of education less than natives).

[Table 7 about here]

To address endogeneity we also implement a GMM-SYS estimator. First we address the issue of the correct choice of the lag specification of the variables. Since our data has a limited number of observational units and a quite large number of years, we are very parsimonious with respect to the number of lags used as instruments, to avoid the problem of over-fitting instruments (Roodman, 2009). Moreover we test whether the Blundell-Bond (1998) GMM-SYS estimator is more appropriate than the Arellano-Bond (1991) estimator. In the Appendix (A2) we show that the GMM-SYS estimator is indeed better suited for our variables than the GMM-DIF. Moreover, as shown in the Appendix (A2) we adopt a specific procedure aimed at identifying the correct lags to be used as internal instruments in the following estimates.

On the basis of the findings of Table (A2) we estimate equation (3) with a GMM-SYS in which we use, as instruments, only the lags that are found to be useful for each labour variable. Moreover, we apply the procedure suggested by Roodman (2009) to further decrease the number of instruments.[19] Since we have good reasons to believe that natives may not be

[18] When we look at the first stage statistics in the lower part of Table (8) we see that the external instrument is a good predictor of the levels of non-educated migrants. The Angrist and Pischke F-statistic shows that the instrument is not weak. The Hausman test on the endogeneity of the instrumented variables cannot reject the null-hypothesis of exogeneity, though the p-value of the test is relatively low, which casts some doubts on the real exogeneity of the variable.

[19] In Tables (6) and (7) we show one-step standard errors, since in small samples with a large number of instruments (due to a large T) standard errors in two-step GMM tend to be severely downward biased (Windmeijer, 2005). However, as a further robustness check, we also performed GMM-SYS estimations with the two-step procedure and found that the significance of our results was not affected at all.

strictly exogenous, in column (5) we also instrument highly-educated natives (*E_Tedu_nat*) and low-educated natives (*E_noTedu_nat*) with their own lags. Highly-educated migrants (*E_Tedu_mig*) have a positive and significant coefficient, while non educated migrants (*E_noTedu_mig*) have a negative and significant coefficient. Also highly-educated natives are positive and significant, while non educated native display a negative and significant effect: again when we endogenize the labour variables we find that, with respect to the fixed effects results, the coefficient for educated workers increases in size, while it decreases for unskilled ones. [20] In the Appendix in Table (A6) we also display the coefficients of the other control variables.

The estimated elasticities show that highly-educated migrants have a positive effect on innovation in the three European countries analysed, but their effect is smaller than educated natives: it stands, in fact, at only one third than that of natives. A 1% increase in the number of highly-educated natives leads to a 0.3% increase in the number of citation-weighted number of patents, whereas a 1% increase in the number of highly-educated migrants leads to slightly less than 0.1% increase in the dependent variable. In Figure 1 we also plot the marginal effects of highly educated migrants and natives. The marginal effects depend upon the size of the variables (*E_Tedu_nat* and *E_Tedu_mig*). In Figure 1 we hold all the other variables constant at their mean values. We show that the effect is generally larger for educated natives. The marginal effect of educated immigrants is larger only when the size of the labour force is rather small. With respect to H1 these results suggest that in Europe, differently from the United States, immigrants do not have a direct effect on the innovation that is stronger than the one of the natives. Although on average in UK, Germany and France there are more graduates among immigrants than among natives, and despite self-selection phenomena, these effects are weaker than in the United States. Probably in the United States there is also a greater demand for skills and immigration policy favours a better job-matching by operating efficiently both on the demand side and on the supply side of the labour market.

[Figure 1 about here]

In Figure 2 we compute the overall effect of skilled migrants and natives on patent production by sector distinguishing between countries of destination. The elasticities that we have estimated in Table 7 indicate the percentage change of patent production for a 1% increase of each type of skilled workers (natives or immigrants). We hence calculate the total percentage change in some sectors (we focus in particular on high-tech sectors) and each country of destination of the number of, respectively, skilled natives and skilled immigrants occurred in the period 1994-2006 (1996-2006 for Germany) and multiply it by the two estimated elasticities. This allows us to identify the contribution of each of the two types of workers to the overall growth of patents in the period. Figure 2 shows that when we consider the sectoral aggregate effect on patent production the gap between migrants and natives is often reduced and in some cases the contribution of skilled migrants is even higher. This is due to the fact that, while the elasticities of skilled migrants are lower, the overall growth

[20] A further concern in our estimation strategy relates to the possible endogeneity of the other covariates of the model, in particular we suspected that the R&D expenditures, the stock of citations and the openness to trade might not be strictly exogenous. In additional analyses (the results are available upon request to the authors) we instrumented with their own suitable lags also these three variables and found that the results were extremely robust: in particular the coefficients and significance of our variables of interest (natives and immigrants) were not affected. On the basis of these results we chose to adopt the more parsimonious specification presented in Tables (7) and (8), since increasing the number of instruments in the model lowers the reliability of the Hansen test concerning the correct choice of the instruments.

of migrants is much higher than the growth of skilled natives. We find that in France the contribution of migrants is generally higher than in the UK and in Germany. In the majority of French high-tech sectors migrant's contribution is higher than that of natives (with the exception of the manufacture of motor-vehicles and other transport equipment). In the UK migrants' contribution is usually slightly lower or equal to that of natives. In Germany skilled migrants are the only category of workers with a positive contribution to patent production in the chemical as well as in the machinery and equipment sectors.

[Figure 2 about here]

In Table 8 we check whether the contribution of skilled immigrants depends on the overall level of over-education within the sector in which they are employed. We exploit occupational data from the Labour Force Surveys in France and the UK, as well as from the MicroCensus in Germany and we aggregate it at the sectoral level (using ISCO-code occupations) to create an index of overeducation among tertiary educated immigrants. The index is built as the ratio of the number of tertiary educated immigrants in a sector over the number of immigrants employed in managerial, professional, scientific or technical occupations (we use the Standard Classification of Occupations -ISCO-88- by International Labour Office, 1990[21]). It captures overeducation because it identifies the relative number of skilled immigrants that are employed in occupations that do not require their level of education. We compute the average level of the overeducation index for the total number of observations (for the period 1994-2006) and then we classify each sector in each country as a sector with high or low overeducation according to the fact that the average of the sector is higher or lower than the average over all observations. We then interact the coefficient of skilled immigrants in equation (3) with two self-excluding dummies: one equal to one for sectors with high overeducation and one equal to one for sectors with low overeducation. The results in column (1) of Table 8 show that the contribution of skilled immigrants is only positive and significantly different from zero in sectors with low levels of overeducation, confirming hypothesis H2.

[Table 8 about here]

The effect of R&D intensity

According to hypothesis H3 we expect that the higher the level of codification of the knowledge used in a sector, the higher will also be the innovative contribution of foreign employees. In column (2) of Table (8) we interact the coefficient of skilled immigrants with two self-excluding dummies. A dummy equal to 1 for the observations belonging to sectors with high R&D intensity and another dummy which is equal to 1 for the observations belonging to sectors with low R&D intensity, using the standard OECDE classification of high and low tech sectors.[22] In Table 8 we check whether the

[21] The ISCO occupations at the 1-digit level used to define high qualified occupations are the following: ISCO1 (Legislators, senior officials and managers), ISCO2 (Professionals) and ISCO3 (Technicians and associate professionals).
[22] We follow the standard OECD classification of sectors (Hatzichronoglou, 1997), that is based on the levels of R&D intensity, to distinguish between sectors with high R&D intensity, as opposed to sectors with low R&D intensity (see the Appendix A3 for details).

contribution of skilled immigrants is different in the two types of sectors. In line with H3 we find that the coefficient of skilled migrants is only significantly different from zero in the sectors with high level of R&D intensity. Moreover its coefficient is 40% higher than the coefficient found for the total economy (see Table 7).

The effect of internationalization

In columns (3) and (4) of Table 8 we test hypothesis H4 and H5. In column (3) we use OECD data on Inward Activity of Multinationals by Industrial Sector on the number of workers employed in foreign multinationals (i.e. Foreign Direct Investments) to calculate the share of workers in foreign multinationals over the total labour force in a sector. We first compute the average level of FDIs for all the observations in the sample. Once we have identified this threshold we check whether in each national sector the share of employed in foreign multinationals is below or above this average (for each sector we use the average level for the years 1994-2007). The advantage of this methodology is that we do not assume common patterns among countries, since the same sectors might display a higher or lower presence of FDIs in Germany, France or the UK (See Appendix A4 for the details). On the basis of this procedure we can interact skilled immigrants with a dummy that indicates a high presence of FDI in a sector and a dummy indicating low FDIs presence in the sector. While high-tech sectors generally show a larger presence of FDIs, there is substantial variation between the intensity of FDIs and the intensity of R&D. Many high-tech sectors display a low intensity of FDIs (see the chemical sector and the machinery and equipment sector in Germany, as well as the manufacture of transport equipment in France and the UK), while there are low-tech sectors with a large presence of foreign multinationals (for example the rubber and plastics, or the basic metals sectors in France, as well as the pulp and paper industry in the UK). The results show that when a sector has a high level of FDIs the contribution of skilled immigrants is almost two times higher than in sector with low FDIs, and only in the former case the coefficient is significantly different from zero: all in all, this confirms the validity of H4.

In column (4) of Table (8) we follow a similar procedure to identify sectors with high or low openness to trade: also in this case we compute the average level of openness to trade for all the observations and then we classify each national sector as a sector with high or low openness to trade. Then we estimate a similar model in which the skilled immigrants are interacted with two dummies indicating the level of openness of each sector. In line with H5 we find that in sectors with high openness to trade the contribution of skilled immigrants is higher and significantly different from zero.

The effect of diversity

Finally in Table (9) we test the validity of our hypotheses about the impact of ethnic diversity among the skilled workforce. Following the established literature on diversity (Kemeny, 2017), we compute the diversity of nationalities (including the natives) within each national sector using the Shannon entropy index:

$$Shannon_i = -\sum_{r=1}^{R} s_{ri} \text{ x } \ln(s_{ri})$$

More specifically the sectors with high R&D intensity correspond to the set of medium-high tech and high tech sectors, according to the OECD classification, while the sectors with low R&D intensity correspond to the OECD set of low-tech and mid-low tech sectors.

where s is the proportion of skilled workers in sector i born in country r, and R is the number of different countries of origin in sector i. In order to be sure that our measure does not depend on the level of aggregation of countries, we calculate the diversity index using 8 geographical areas[23] or, alternatively, 27 different countries of origin (see the full list of countries of origin in table A4 in the Appendix). We calculate the average diversity at the sectoral level and standardize the diversity index with respect to each country average, since the overall level of diversity differs substantially across countries.[24] Finally we divide the sectors in two groups based on whether the average diversity is above or below the country average.

In columns (1) and (2) of Table (9) shows that the estimated impact of skilled immigrants on innovation is higher in sectors with higher levels of diversity. This result holds regardless of the way we measure diversity This confirms Hypothesis H6 according to which higher cultural diversity boosts the contribution of skilled immigrants.

[Table 9 about here]

In column (3) we test whether the innovative contribution of skilled immigrants depend upon their country of origin. In particular we distinguish between migrants coming from European or non-European countries[25]. Indeed recent works (e.g. Moguérou, Di Pietrogiacomo, 2008) show that skilled migration in Europe consists mainly of European citizens moving from one country to another, exploiting their right to move freely across European borders[26]. Table (9) shows the estimated effects of skilled immigrants using GMM-SYS. The results in column (3) show that tertiary-educated foreign workers have a positive effect on innovation and the estimated coefficient of the European workers is more than twice as large as the coefficient of the non Europeans, providing evidence for the validity of hypothesis H7.

5. Concluding comments

This paper analyses the effects of international skilled migrants on innovation activities in UK, Germany and France. Previous studies have mainly focused on the estimation of the effect of skilled migrants on innovation at the individual or firm level or at the regional or country level, emphasizing the potentially positive role of ethnic and cultural diversity. In this paper we combine, at the sectoral level, the French and UK Labour Force Surveys, the German Microcensus and the European patents and citations database (PATSTAT) to estimate the effect of the employment of native and migrant workers on innovation. We adopt an empirical strategy based on industries that has the advantage (relative to using

[23] We use the following geographical areas: 1) Africa, 2) North America, 3) Central and South America, 4) Middle East and Central Asia, 5) South and Eastern Asia, 6) Eastern Europe, 7) Western Europe, and 8) Oceania

[24] This is due to the fact that the diversity index (by including also natives) proxies also the overall share of immigrants in a sector. Since this shares is substantially different among the three countries (with Germany as a upper bound and France as a lower bound) using an average of diversity across countries would mean that all German sectors would results in high diversity sectors and all French sectors would result as low-diversity sectors.

[25] In our analysis the set of European countries includes also some countries which are not inside the European Union, such as Norway, Switzerland, Bosnia-Herzegovina, Serbia, Montenegro and Albania. We made this decision as some national statistical offices aggregated workers coming from a contiguous set of countries, so in some cases we could not distinguish between, say, a Slovenian (inside the European Union) and a Bosnian (outside the European Union).

[26] This is a key difference with respect to other countries such as the United States, where migrants come from very different world regions (Latin and Central America, as well as Asia and Europe).

regions or provinces) that it provides information on whether immigrants (in particular skilled immigrants) are really employed in the patenting sectors. In addition we exploit differences across industries to analyze under which conditions it is more likely that skilled immigrants contribute to the European innovation systems. We complement and add to previous studies in a number of respects.

The first result is the estimation for Europe of how the skilled migrants differ from skilled natives in their impact on innovation, using different sets of internal and external instruments. Using an innovation production function at the industry level we control for age, level of education, countries of origin, R&D, knowledge stock and openness to trade. We show that a highly-educated labour force has a positive impact on innovation. This holds not only for the highly-educated natives but also, with a smaller coefficient, for high skilled migrants. In particular a 1% increase in the number of educated natives leads to a 0.3% increase in the citation-weighted number of patents, a 1% increase in the number of highly-educated migrants leads to a slightly less than 0.1% increase in the citation-weighted number of patents.

Secondly we test the size of the impact of native and migrant workers on innovation conditioning our results to different characteristics of the manufacturing sectors (in different countries). The overall idea is that skilled migrants give a greater contribution to a country innovation system in sectors that are more R&D intensive, with larger presence of multinational corporations and more open to trade. In line with our hypotheses we find that the skilled migrants have a stronger impact in the sectors with a high level of R&D intensity. In this case the estimated coefficient is also 40% higher than the one calculated for the total economy. Our results show also that when a sector has a high level of FDIs the contribution of skilled immigrants is twice as large as the impact in sectors with low FDIs. In addition skilled immigrants have also a stronger effect on innovation in sectors that are more open to international trade.

We interpret these results with the view that skilled migrants participate and influence innovation in those R&D activities where knowledge is more codified and transfers more easily across countries and it does not rely on location specific institutional or cultural knowledge. In addition they can be integrated more easily in the innovation process in those companies that are active in international markets, either through the establishment of foreign subsidiaries or through export and import operations. These companies are more likely to adopt management practices that can boost the contribution of individuals with a different cultural background.

This raises the question on how migration policies can shape the European innovation system. Our evidence suggests that innovation in Europe would benefit from policies that favor the entrance of potential workers with tertiary education, or with advanced degrees, such as masters' degrees and doctorates in these areas. In fact national policies (at least, in the Netherlands, Sweden and the UK) and the European Blue Card (used extensively in Germany) try to facilitate the recruitment of highly-skilled workers.

At the same time our results raise the question on whether Europe is able to attract top level high skilled worker and proficiently participate to the global competition for talents. On the one side EU attractiveness depends on factors that are not affected by migration policies like living standards, the welfare system, the entrepreneurial environment and taxes. Institutional quality and governance effectiveness could also increase Europe's attractiveness for highly-qualified migrants. On the other side migration policies designed to simply increase the number of highly skilled workers could miss the target in terms of innovation impact. Our results suggest that an innovation-oriented migration policy should be able to adjust to labour demand from industries with high R&D intensity, high levels of FDIs and open to the international markets. In this respect the European Blue card or some national schemes (e.g. Netherland) follow this line and offer easy access to highly skilled foreigners in demand.

Finally we consider the countries of origin of the skilled workforce and analyze ethnic diversity, measuring how many nationalities there are and how they are distributed at the industry level. We find that the impact of skilled immigrant is higher in sectors with a higher level of ethnical diversity. In addition the positive effect on innovation is stronger for European migrants than for third-countries nationals. Our results support the view that a diverse environment enhances the role of skilled immigrants and that skilled labor mobility within the EU produces a better match of jobs and task specialization and a process of learning by hiring that fills high skilled labor shortages in specific sectors.

This result, on the one hand, suggests that in a diverse environment skilled migration has a higher impact on innovation. This is the case when diversity is calculated both at country level and at level of larger geographical areas. On the other hand the positive effect of skilled worker movements within the EU show that the innovative contribution comes predominantly from skilled migrants that come from less distant cultural backgrounds. This suggests that in the short run favoring intra EU mobility is a first best solution to increase innovation, while increasing diversity with third national countries could be also beneficial to improve the contribution of skilled migrants to the European innovation system.

We believe that future work should explore more in depth the impact of labor demand on skilled immigrants and how the nature and composition of demand affect the contribution of skilled migrants to European industry. A precise understanding of the mechanisms that link demand and supply of skilled labor in Europe would facilitate a closer coordination between immigration and innovation policy. More attention needs to be paid to skilled migration in high tech sectors and to the mobility of skilled workers with specific skills within MNCs (e.g. Foley and Kerr, 2013). The impact on destination countries of the movement of people within companies remains rather unexplored in the literature of migration and innovation. Finally our results suggest that policy initiatives to attract skilled workers in Europe should be framed also as part of the general effort to favour the internal circulation of scientists and engineers, as suggested by Lissoni (2017). In this respect more research is needed to understand in Europe not only the innovation impact of selective policies to attract foreign talents but also the impact of these policies on skilled workers mobility within the EU.

References

Acemoglu, D., (2002), Directed Technical Change, *Review of Economic Studies*, 69(4), 781-809.

Alesina A. La Ferrara E. (2005) Ethnic Diversity and Economic Performance. *Journal of Economic Literature.* Vol. 43, No. 3 (Sep., 2005), pp. 762-800.

Arellano M., S. Bond. (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, 58, 277-297.

Aydemir, A., Borjas, G.J. (2011). Attenuating bias in measuring the wage impact of immigration *Journal of Labor Economics,* 29 (1), 69–112.

Bacchiocchi, E., Montobbio F. (2010), International knowledge diffusion and home-bias effect. Do USPTO & EPO patent citations tell the same story?, *Scandinavian Journal of Economics*, 112 (3), 441-470

Bartlett, C.A. and Ghoshal, S. (1989) *Managing Across Borders: The Transnational Solution*, Harvard Business School Press: Boston.

Batalova J. Fix M. (2017); New Brain Gain: Rising Human Capital among Recent Immigrants to the United States. Migration Policy Institute. Washington DC

Becker, G., (1975), *Human Capital. A Theoretical and Empirical Analysis with Special Reference to Education*, University of Chicago Press, Chicago.

Blundell, R. and S. Bond (1998), Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics*, 87, 115-143.

Bosetti, V., Cattaneo, C., Verdolini, E., (2015), Migration of Skilled Workers and Innovation: A European perspective, *Journal of International Economics,* 96(2), 311–322.

Bratti, M., Conti C., (2014), The Effect of (Mostly Unskilled) Immigration on the Innovation of Italian Regions, *Institute for the Study of Labor (IZA) DP No. 7922*.

Breschi, S., Lissoni, F., Noumedem Temgoua, C. (2016). Migration and innovation: A survey of recent studies. In R. Shearmur, C. Carrincazeaux, & D. Doloreux (Eds.), Handbook on the geography of innovation (pp. 382–398). Cheltenham: Edward Elgar.

Breschi, S., Lissoni, F., & Miguelez, E. (Forthcoming 2017). Foreign-origin inventors in the US: Testing for diaspora and brain gain effects. Journal of Economic Geography. doi:10.1093/jeg/lbw044

Caligiuri, P., Bonache, J. (2016) Evolving and enduring challenges in global mobility, Journal of World Business, 51, 1, 127-141.

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration, *Journal of Labor Economics*, 19(1), 22-64.

Caselli, F. Coleman, W. J. (2013), On the Theory of Ethnic Conflict. Journal of the European Economic Association, 11: 161–192. doi:10.1111/j.1542-4774.2012.01103.x

Chellaraj, G., Maskus, K., A. Mattoo, (2008), The contribution of skilled immigrations and international graduate students to U.S. innovation. *Review of International Economics* 16(3), 444–62.

Chiswick, B. R. Taengnoi, S. (2007), Occupational Choice of High Skilled Immigrants in the United States. *International Migration*, 45: 3–34. doi:10.1111/j.1468-2435.2007.00425.x

Chiswick, B. Miller, P. (2010), The Effects of Educational-Occupational Mismatch on Immigrant Earnings in Australia, with International Comparisons, *International Migration Review*, 44(4), 869–898.

Cohen, W.M., Levinthal, D.A., 1990. Absorptive capacity: a new perspective of learning and innovation. Administrative Science Quarterly 35 (1), 128–152.

Cohen, W. M., Nelson, R. R. and J. P. Walsh, (2000), Protecting their intellectual assets: appropriability conditions and why U.S. manufacturing firms patent (or not), *NBER Working Paper 7552*.

Constant, A., (2005), Immigrant Adjustment in France and Impacts on the Natives in K.F. Zimmermann (ed.), "European Migration: What Do We Know", *Oxford, OUP*, and *IZA DP. 2004/8063.*

De la Rica, Sara, Albrecht Glitz and F. Ortega, (2014) "Immigration in Europe: Trends, Policies and Empirical Evidence", in B. Chiswick and P. Miller, eds, Handbook of the Economics of International Migration, Elsevier

Deng, Yi, (2007), Private value of European patents, *European Economic Review*, 51, 1785–1812.

Dustman, C. Glitz, A. (2011), Migration and Education, *Handbook of the Economics of Education*, 4, 327-439.

Edstrom, A., Galbraith, J. (1977). Transfer of managers as a coordination and control strategy in multinational organizations. *Administrative Science Quarterly*, 22: 248–263.

Ellis, M., Wright, R., (1999), The Industrial Division of Labor among Immigrants and Internal Migrants to the Los Angeles Economy, *International Migration Review*, 33(1), 26-54

Fassio C., Kalantaryan S., Venturini A. (2017); High-skilled Immigration and Innovation, in High-Skilled Migration: Drivers and Policies, M. Czaicha (ed.), Oxford University Press. Forthcoming.

Foley, C. F., Kerr, W. R. (2013). Ethnic innovation and U.S. multinational firm activity. Management Science, 59(7), 1529–1544. doi:10.1287/mnsc.1120.1684

Furman, J., Porter, M.E., Stern, S., (2002), The determinants of national innovative capacity, *Research Policy* 31, 899–933.

Gagliardi L., (2015), Does Skilled Migration Foster Innovative Performance? Evidence from British Local Areas, *Papers in Regional Science,* 94(4): 773-794.

Griliches, Z., Hausman, J. A., (1986), Errors in variables in panel data*, Journal of Econometrics*, 32, 93-118.

Hall, B. H., Mairesse, J., (1995), Exploring the relationship between R&D and productivity in French manufacturing firms. *Journal of Econometrics*, 65(1), 263–293.

Hanushek Eric A. Woessmann L., (2011). How much do educational outcomes matter in OECD countries?, Economic Policy, 26(67), 427-491.

Hatzichronoglou, T. (1997), Revision of the high-technology sector and product classification, *STI Working Papers, OECD/GD*, 97, 216.

Hiller, S. (2013) Does immigrant employment matter for export sales? Evidence from Denmark. *Review of World Economics*, 149:369–394

Hunt, J., Gauthier-Loiselle M., (2010), How much does immigration boost innovation?, *American Economic Journal: Macroeconomics*, 2(2), 31-56.

Kemeny T. (2017); Immigrant diversity and economic performance in cities. International Regional Science Review 40 (2), 164-208.

Kemeny, T. Cooke, A. (2016) Urban immigrant diversity and inclusive institutions. *Economic Geography*, 93:3, 267-291.

Kerr, W., Lincon, W.F., (2010), The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention, 28(3), 473-508, *Journal of Labour Economics*.

Kerr, W. R. (2016). US high-skilled immigration, innovation, and entrepreneurship: Empirical approaches and evidence. In C. Fink & E. Miguelez (Eds.), The international mobility of talent and innovation: New evidence and policy implications (pp. 193– 221). Cambridge: Cambridge University Press.

Knorr-Cetina, K. (1999) Epistemic Cultures. How the Sciences Make Knowledge. Cambridge: Harvard University Press.

ILO (1990) ISCO-88: International standard classification of occuaptions, Technical Report, ILO Geneva

Levy, O., Beechler, S., Taylor, S., & Boyacigiller, N. A. (2007). What we talk about when we talk about 'global mindset': Managerial cognition in multinational corporations. *Journal of International Business Studies*, 38(2): 231–258.

Lissoni F. (2017); International migration and innovation diffusion: an eclectic survey. Regional Studies, forthcoming. https://doi.org/10.1080/00343404.2017.1346370

Lodefalk, M. (2016). Temporary expats for exports: micro-level evidence*. Review of World Economics*, 152 (4), 733-772.

Mahy, B., Rycx, F. Vermeylen, (2015) Educational Mismatch and Firm Productivity: Do Skills, Technology and Uncertainty Matter? *G. De Economist*, 163: 233.

Malerba F., Mancusi M.L., Montobbio, F. (2013)."Innovation, international R&D spillovers and the sectoral heterogeneity of knowledge flows, *Review of World Economics (Weltwirtschaftliches Archiv)*, 149(4), 697-722.

Martin R. (2013); Differentiated knowledge bases and the nature of innovation networks. European Planning Studies 21 (9), 1418-1436

Minbaeva, D. B., Michailova, S. (2004) Knowledge transfer and expatriation in multinational corporations: The role of disseminative capacity, *Employee Relations*, Vol. 26 Issue: 6, pp.663-679.

Moguérou P., Di Pietrogiacomo M.P., (2008), Stock, Career and Mobility of Researchers in the EU, *JRC Scientific and Technical Reports*, European Commission.

Nathan, M. Lee, N. (2013). Cultural Diversity, Innovation, and Entrepreneurship: Firm-level Evidence from London. *Economic Geography*, 89 (4), 367-394.

Nathan, M. (2015) Same difference? Minority ethnic inventors, diversity and innovation in the UK, *Journal of Economic Geography*,15, 129-168.

Niebuhr, A., (2010), Migration and innovation. Does Cultural Diversity matter for regional R&D Activity?, *Papers in Regional Sciences*, 89(3), 563-585.

Nieto, S. Matano, A., Ramos, R. (2015) Educational mismatches in the EU: immigrants vs natives, *International Journal of Manpower*, 36(4), 540-561.

No, Y., Walsh, J. P., (2010), The importance of foreign-born talent for US innovation, *Nature biotechnology*, 28(3), 289-291.

OECD (2015), Connecting with Emigrants: A Global Profile of Diasporas 2015, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264239845-en

Østergaard, C.R.; Timmermans, B.; Kristinsson, K. (2011) Does a different view create something new? The effect of employee diversity on innovation. In: Research Policy, Vol. 40, No. 3, 04.2011, p. 500-509

Ottaviano, GM., Peri, G., (2012), Rethinking the Effects of Immigration on Wages, *Journal of the European economic association*, 10(1), 152-197.

Ozgen, C., Nijkamp, P., Poot, J., (2012), Immigration and Innovation in European Regions, in P. Nijkamp, J. Poot J and M. Sahin (eds.) “Migration Impact Assessment: New Horizons”, Edward Elgar.

Ozgen, C., Nijkamp, P., and Poot, J. (2017) The elusive effects of workplace diversity on innovation. Papers in Regional Science, 96: S29–S49. doi: 10.1111/pirs.12176.

Pakes, A., M. Schankerman, (1984). The Rate of Obsolescence of Knowledge, Research Gestation Lags, and the Private Rate of Return to Research Resources; in Griliches Z. (ed.), *R&D Patents & Productivity*, Chicago: University of Chicago Press.

Parrotta, P., Pozzoli, D. & Pytlikova, M. (2014); The Nexus Between Labor Diversity and Firm’s Innovation. *Journal of Population Economics,* 27: 303. https://doi.org/10.1007/s00148-013-0491-7

Patel, P., K. Pavitt, (1994), The Nature and Economic Importance of National Innovation Systems, *STI Review*, No. 14, OECD, Paris.

Rashidi-Kollmann, S., Pyka, A. (2016). Migration and innovation: A survey. In A. Pyka, Y. Kuştepeli, & D. Hartmann (Eds.), International innovation networks and knowledge migration: The German–Turkish nexus (pp. 89–116). Oxford: Routledge.

Roodman, D. (2009), A Note on the Theme of Too Many Instruments, *Oxford Bulletin of Economics and Statistics*, 71, 135–158.

Romer, P. M., (1990), Endogenous Technological Change, *Journal of Political Economy*, 98(5), 71-102.

Schmoch, U., Laville, F., Patel P., Frietsch R. (2003), Linking Technologies to industrial Sectors. *Final report to the European Commission*, DG Research.

Sparber, C. (2010); Racial diversity and aggregate productivity in US industries: 1980–2000. Southern Economic Journal, 75(3), p. 829

Stephan, P., Levin, S., (2001), Exceptional contributions to US science by the foreign-born and foreign-educated, *Population Research and Policy Review*, 20(1), 59-79.

Sturgeon, T., J. (2002) Modular production networks: a new American model of industrial organization, *Industrial and Corporate Change*, 11, 3, 451–496, Windmeijer, F., (2005), A finite sample correction for the variance of linear efficient two-step GMM estimators, *Journal of Econometrics*, 126(1), 25-51.

Wintoki, M.B., Linck, J.S., Netter, J.M., (2012), Endogeneity and the dynamics of internal corporate governance, *Journal of Financial Economics*, 105, 581–606.

Wooldridge, J.M.,(2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge.

Zheng, Y. and Ejermo, O. (2015), How Do the Foreign-born Perform in Inventive Activity? Evidence from Sweden, *Journal of Population Economics,* 28 (3), 659-695.

TABLES

Table 1. Description of the variables

VARIABLE	DESCRIPTION	SOURCE
Logcit	Log of the 4-years citation-weighted patents.	PATSTAT – EPO Database
R&D	log of R&D expenditures (in PPP 2005 dollars)	OECD, BERD-SAN Database
Stock_cit	log of the stock of citations-weighted patents, created using a perpetual inventory method (depreciation rate set at 10%)	PATSTAT – EPO Database
Open	log of the openness to trade. (Import + Export)/Value added	OECD-STAN database
E	log of total employment	Labour force Surveys for France and UK. Microcensus for Germany
E_Tedu	log of employees with tertiary education. In the UK we consider as tertiary-educated those workers that left school when they were older than 21 years old. In France we consider tertiary educated those workers who obtained a degree that is beyond that of the "baccalaureat general". In Germany tertiary education corresponds to at least 3 years of tertiary education.	Labour force Surveys for France and UK. Microcensus for Germany
E_noTedu	log of employees without tertiary education	Labour force Surveys for France and UK. Microcensus for Germany
E_Tedu_nat	log of native employees with tertiary education	Labour force Surveys for France and UK. Microcensus for Germany
E_Tedu_mig	log of immigrant employees with tertiary education. In each of the Labour Force Surveys (and the Microcensus for Germany) we considered as immigrant/foreigner any worker whose nationality is different from that of the country in which he or she is working.	Labour force Surveys for France and UK. Microcensus for Germany
E_noTedu_nat	log of native employees without tertiary education	Labour force Surveys for France and UK. Microcensus for Germany
E_noTedu_mig	log of immigrant employees without tertiary education	Labour force Surveys for France and UK. Microcensus for Germany
E_Tedu_mig EU	log of immigrant employees with tertiary education holding the nationality of a European country (for the case of Germany) or born in a European country (for the case of France and UK).	Labour force Surveys for France and UK. Microcensus for Germany

E_Tedu_mig NOEU	log of immigrant employees with tertiary education holding the nationality of a non-European country (for the case of Germany) or born in a non-European country (for the case of France and UK).	Labour force Surveys for France and UK. Microcensus for Germany
Avg_age	log of the average age of the total employment	Labour force Surveys for France and UK. Microcensus for Germany
Avg_age_Tedu	log of the average age of employees with tertiary education	Labour force Surveys for France and UK. Microcensus for Germany
Avg_age_nat	log of the average age of the native employees	Labour force Surveys for France and UK. Microcensus for Germany
Avg_age_Tedu_nat	log of the average age of native employees with tertiary education	Labour force Surveys for France and UK. Microcensus for Germany
Avg_age_mig	log of the average age of the immigrant employees	Labour force Surveys for France and UK. Microcensus for Germany
Avg_age_Tedu_mig	log of the average age of immigrant employees with tertiary education	Labour force Surveys for France and UK. Microcensus for Germany

Table 2. Patent and Human capital aggregate statistics

	UK		FRANCE		GERMANY	
	1994-1996	2003-2005	1994-1996	2003-2005	1996-1998	2003-2005
Patents/Citations (per 1000 employee)						
Patents	0.91	1.64	1.42	2.09	2.28	3.08
Citations	1.54	0.73	2.04	0.78	3.21	1.54
Share of young workers	0.44	0.35	0.40	0.37	0.41	0.34
Tertiary-educated	0.05	0.07	0.08	0.11	0.04	0.03
Non-tertiary-educated	0.39	0.28	0.33	0.27	0.37	0.31
Share of tertiary educated	0.08	0.14	0.14	0.20	0.10	0.11
Share of immigrants	0.064	0.079	0.027	0.042	0.127	0.121
Tertiary-educated	0.012	0.024	0.003	0.007	0.008	0.011
Non-tertiary-educated	0.052	0.055	0.024	0.035	0.110	0.103
EU nationals	0.021	0.024	0.012	0.020	0.062	0.063
EU nationals tertiary-educated	0.004	0.007	0.002	0.004	0.004	0.006
Share of immigrant employment in the whole economy	0.071	0.096	0.025	0.047	0.086	0.092

We classify as "young" workers that are younger than 35. The share of young workers and tertiary educated includes also immigrant workers. See Table (1) for a precise definition of "tertiary-educated workers" and "immigrant workers". All data refer only to the manufacturing sectors except for the last row which refers instead to the whole economy.

Table 3. Patent and Human capital aggregate statistics in High Tech sectors.

	UK		FRANCE		GERMANY	
	1994-1996	2003-2005	1994-1996	2003-2005	1996-1998	2003-2005
Patents/Citations (per 1000 employee)						
Patents	1.67	2.86	2.79	4.23	3.74	4.88
Citations	2.98	1.31	4.20	1.62	5.53	2.47
Share of young workers	0.45	0.34	0.38	0.37	0.41	0.34
Tertiary-educated	0.06	0.09	0.11	0.14	0.05	0.05
Non-tertiary-educated	0.38	0.26	0.27	0.23	0.35	0.30
Share of educated	0.12	0.18	0.21	0.28	0.15	0.16
Share of immigrants	0.061	0.078	0.021	0.035	0.118	0.113
Tertiary-educated	0.016	0.030	0.004	0.011	0.012	0.016
Non-tertiary-educated	0.045	0.048	0.017	0.024	0.098	0.090
EU nationals	0.020	0.024	0.011	0.018	0.060	0.060
EU nationals tertiary-educated	0.005	0.009	0.003	0.008	0.006	0.008

We classify as "young" workers that are younger than 35. The share of young workers and tertiary educated includes also immigrant workers. See Table (1) for a precise definition of "tertiary-educated workers" and "immigrant workers" and Table A3 in the Appendix for the definition of high-tech sectors.

Table 4. Patents and migrant shares by sector

		Patents/Citations (per 1000 employee)		Immigrants				
Industry	ISIC REV. 3.1	*Patents*	*Citations*	*Share of immigrants*	*Tertiary educated*	*Non-tertiary educated*	*EU nationals*	*EU nationals tertiary educated*
Food Products, Beverages And Tobacco	15-16	0.12	0.09	0.07	0.01	0.06	0.03	0.003
Textiles And Textile Products, Leather And Footwear	17-19	0.21	0.16	0.12	0.01	0.11	0.04	0.003
Wood And Products Of Wood And Cork	20	0.14	0.06	0.05	0.00	0.05	0.03	0.002
Pulp, Paper, Paper Products, Printing And Publishing	21-22	0.38	0.38	0.08	0.01	0.07	0.04	0.004
Chemicals And Pharmaceuticals	24	4.63	5.86	0.06	0.02	0.04	0.03	0.008
Rubber And Plastics Products	25	1.54	1.12	0.08	0.01	0.08	0.03	0.002
Other Non-Metallic Mineral Products	26	1.11	0.90	0.06	0.01	0.05	0.03	0.003
Basic Metals	27	0.68	0.42	0.09	0.01	0.08	0.03	0.002
Fabricated Metal Products, exc. Machinery. and Equip.	28	0.56	0.38	0.07	0.01	0.07	0.03	0.002
Machinery And Equipment, Nec	29	2.90	2.20	0.06	0.01	0.05	0.03	0.004
Office, Accounting And Computing Machinery	30	10.57	7.12	0.08	0.04	0.04	0.03	0.017
Electrical Machinery And Apparatus, Nec	31	1.73	1.33	0.07	0.01	0.05	0.03	0.005
Radio, Television And Communication Equipment	32	6.80	6.52	0.07	0.02	0.04	0.03	0.010
Medical, Precision And Optical Instruments	33	6.10	5.58	0.06	0.01	0.04	0.03	0.006
Motor Vehicles, Trailers And Semi-Trailers	34	1.63	1.90	0.10	0.01	0.08	0.04	0.004
Other Transport Equipment	35	0.79	0.48	0.05	0.01	0.04	0.02	0.006

Table 5. Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max	Observations
logcit	5.355	1.607	0.405	8.677	457
R&D	20.219	1.525	16.132	23.374	457
open	-0.154	0.606	-1.412	1.631	457
stock_cit	7.821	1.527	2.943	10.739	457
E	12.461	0.657	9.834	14.052	457
E_Tedu	10.379	0.703	8.503	12.030	457
E_noTedu	12.268	0.725	8.849	13.826	457
E_Tedu_nat	10.280	0.723	8.439	11.957	457
E_noTedu_nat	12.192	0.707	8.849	13.717	457
E_Tedu_mig	7.697	0.995	4.691	9.826	457
E_noTedu_mig	9.445	1.161	4.940	11.889	457
E_Tedu_mig EU	6.360	2.141	0	9.210	457
E_Tedu_mig NOEU	6.453	2.471	0	9.302	457
avg_age	3.681	0.033	3.546	3.764	457
avg_age_Tedu	3.644	0.073	3.483	3.859	457
avg_age_Tedu_nat	3.643	0.076	3.476	3.867	457
avg_age_Tedu_mig	3.649	0.125	2.996	4.135	457
avg_age_noTedu	3.686	0.037	3.544	3.787	457
avg_age_noTedu_nat	3.687	0.037	3.545	3.783	457
avg_age_noTedu_mig	3.706	0.093	3.332	4.060	457

We have 16 two-digit sectors for 12 years for France (1994-2005), 14 two-digit sectors for 12 years for the UK (1994-2005) and 14 two-digit sectors for 10 years for Germany (1996-2005). Our original sample, thus, consists of 520 observations: 192 observations in France, 168 observations in the UK and 160 observations for Germany. Because of the one year lag chosen for our estimation, we lose 16 observations in France and Germany and 14 in the UK (46 overall), which correspond to the first year of each time-series. Furthermore, especially in France, in the first years of observation for some small and high tech sectors there were no foreign workers at all, so we can't retrieve information on the average age of foreign workers: therefore we lose those observations (15 observations in France). This also happens, for only one observation, both in the UK and Germany. Net of these missing observations, overall we have 161 observations for France, 143 observations for Germany and 153 observations for the UK, which sums up to 457 observations that are used in our estimates.

Table 6. Baseline model: skills.

Variables	(1) OLS	(2) GMM-SYS	(3) OLS	(4) GMM-SYS
E_{t-1}	-0.180	**-0.441****		
	(0.167)	**(0.222)**		
E_Tedu_{t-1}			-0.015	**0.468****
			(0.091)	**(0.220)**
E_noTedu_{t-1}			-0.123	**-0.886*****
			(0.185)	**(0.323)**
avg_age_{t-1}	-2.429**	-3.526***		
	(1.088)	(1.236)		
$avg\_age\_Tedu_{t-1}$			-0.400	-0.187
			(0.664)	(0.471)
$avg\_age\_noTedu_{t-1}$			-2.133**	-2.880**
			(0.980)	(1.272)
$R\&D_{t-1}$	0.301***	0.305***	0.289***	0.261***
	(0.084)	(0.076)	(0.083)	(0.092)
$stock\_citations_{t-1}$	0.134	0.183**	0.131	0.133
	(0.160)	(0.072)	(0.163)	(0.085)
$open_{t-1}$	-0.587***	-0.920***	-0.552***	-0.878***
	(0.195)	(0.258)	(0.185)	(0.303)
Constant	9.458*	17.213***	9.526*	16.314***
	(4.940)	(5.561)	(5.440)	(5.924)
Time effects	YES	YES	YES	YES
Fixed effects	YES	YES	YES	YES
Observations	457	457	457	457
Number of id	46	46	46	46
R-squared	0.790	-	0.791	-
AR(1) p-value		0.002		0.003
AR(2) p-value		0.546		0.508
Hansen test		0.649		1.828
Hansen test p-value		0.723		0.767

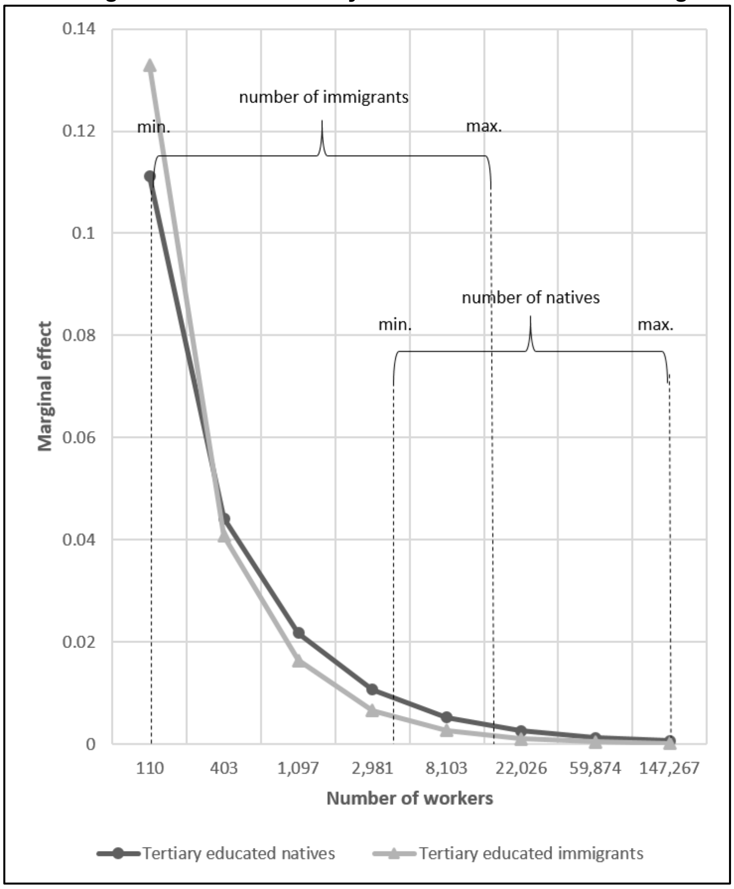
The dependent variable is the log of citations received in the last 4 years. All variables are in logs and are 1-year lagged. In columns (1) and (3) OLS estimators are implemented. In columns (2) and (4) (one-step) robust GMM-SYS estimators are used. All models include time, country and industry dummies. In the GMM models the endogenous variables are *E, E_edu, E_noedu*. The collapse option (Roodman, 2009) is implemented for the GMM-style instruments in levels and in differences, only the first two lags of the endogenous variables are used. Standard errors are clustered at the country-industry level, *** p<0.01, ** p<0.05, * p<0.1.

Table 7. Skills and Ethnicity

Variables	(1) OLS *all exog*	(2) IV -CARD *E_noedu_im endog*	(3) IV -CARD *E_edu_im endog*	(4) GMM-SYS *E_noedu_im & E_edu_im endog*	(5) GMM-SYS *all labor endog*
E_Tedu_mig$_{t-1}$	**0.036***	**0.037***	-1.551	**0.067***	**0.089****
	(0.021)	**(0.022)**	(1.362)	**(0.038)**	**(0.045)**
E_noTedu_mig$_{t-1}$	-0.048	**-0.212****	0.081	-0.170	**-0.338***
	(0.052)	**(0.103)**	(0.162)	(0.283)	**(0.201)**
E_Tedu_nat$_{t-1}$	-0.013	0.045	0.173	-0.082	**0.292****
	(0.083)	(0.099)	(0.329)	(0.107)	**(0.141)**
E_noTedu_nat$_{t-1}$	-0.129	-0.100	0.981	0.199	**-0.752****
	(0.195)	(0.186)	(1.044)	(0.270)	**(0.354)**
Other controls	YES	YES	YES	YES	YES
Time effects	YES	YES	YES	YES	YES
Fixed effects	YES	YES	YES	YES	YES
First stage		*lognosk_im*	*logsk_im*		
E_Tedu_mig_card$_{t-1}$		-	0.085		
		-	(0.076)		
E_noTedu_mig_card$_{t-1}$		**0.857*****	-		
		(0.147)	-		
Angrist-Pischke F test of excl. instr:		33.97	1.25		
p-value		0.000	0.269		
Hausman endog. test p-value		0.142	0.060		
Observations	457	451	448		457
Number of id	46	45	45		46
R-squared	0.795	-	-		-
Num. instruments					47
AR(1) p-value					0.001
AR(2) p-value					0.578
Hansen test p-value					0.426

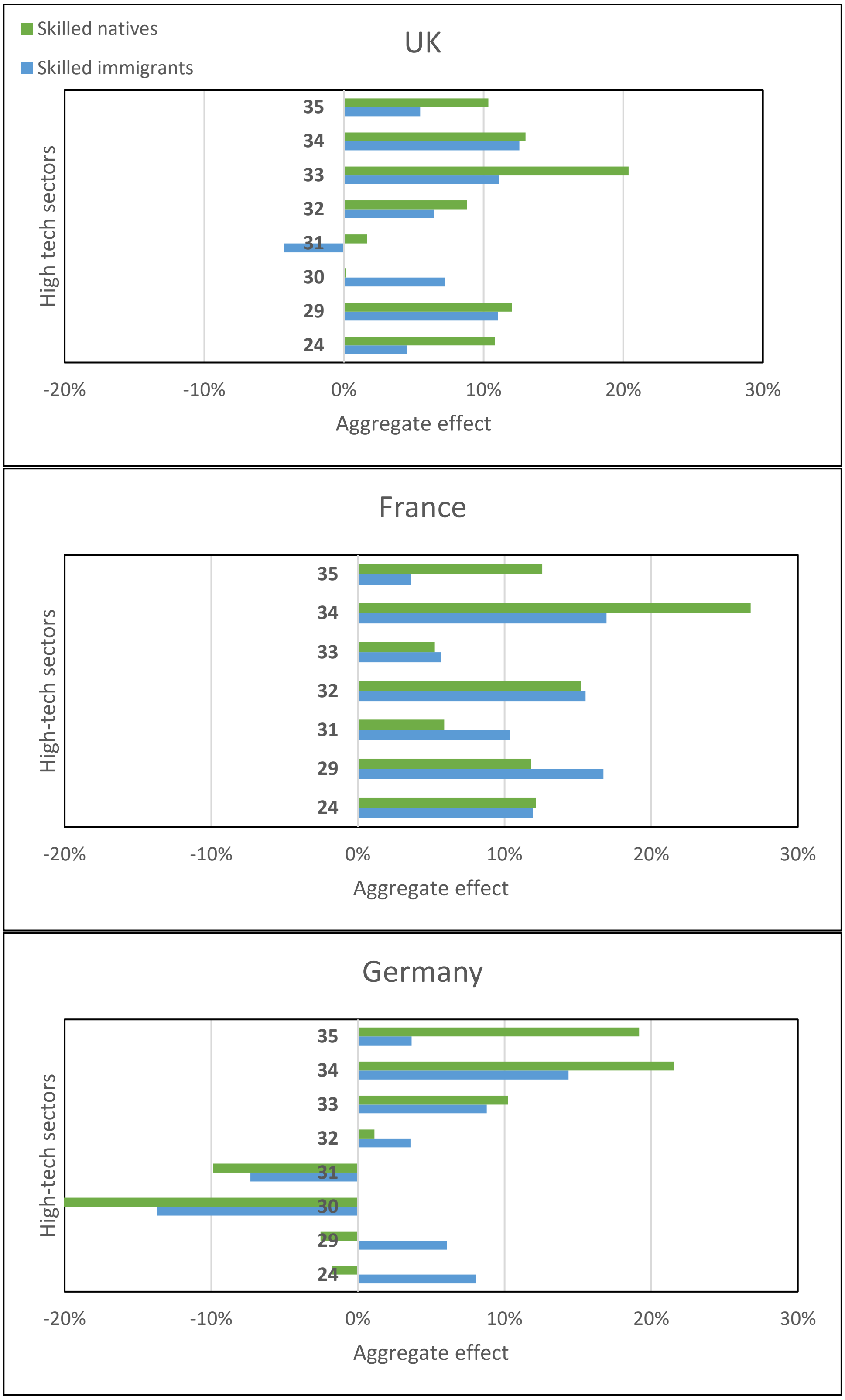
The dependent variable is the log of citations received in the last 4 years. All variables are in logs and are 1-year lagged. All models include additional control variables, as displayed in Table (A2). All models include time, country and industry dummies. In column (1) the OLS estimator is implemented. In columns (2) and (3) two-stage least squares estimators are implemented. In the panel below first-stage coefficients are reported. The Angrist-Pischke test of excluded instruments reports the probability that excluded instruments in columns (2) and (3) are weak, the Hausman test reports the probability that the instrumented variables are endogenous. In columns (4) and (5) (one-step) robust GMM-SYS estimators are used. The collapse option (Roodman, 2009) is implemented for the GMM-style instruments in levels and in differences. In column (4) the endogenous variables are *E_Tedu_mig*, *E_noTedu_mig*. On the basis of the results in Table (9) *E_Tedu_mig* is instrumented with one and two year lags, while *E_noTedu_mig* is instrumented with two and three years lags. In column (5) the endogenous variables are *E_Tedu_mig*, *E_noTedu_mig*, *E_Tedu_nat*, *E_noTedu_nat,*. On the basis of the results in Table (9) both *E_Tedu_nat* and *E_noTedu_nat* are instrumented with one year lags, while *E_Tedu_mig*, *E_noTedu_mig* are instrumented with the same lags as in column (4). In columns (4) and (5) all the additional control variables are considered as exogenous. Standard errors are clustered at the country-industry level, *** p<0.01, ** p<0.05, * p<0.1.

Figure 1. Marginal effects of tertiary educated natives and immigrant workers.



Note: the marginal effects refer to column (5) of Table (8) holding all other variables constant at their mean.

Figure 2. The aggregate effect of skilled migrants and skilled natives on patent production in high-tech sectors between 1994 and 2006



Note: The data are computed by interacting the coefficients estimated in table (7) by the average increase of skilled natives and skilled immigrants in the period 1994-2006 (1996-2006 for Germany).

Table 8. Skilled migrants: Overeducation, R&D, FDI and Openness to Trade

	(1)	(2)	(3)	**(4)**
	GMM -SYS	GMM -SYS	GMM -SYS	GMM -SYS
the contribution of skilled immigrants (*E_Tedu_mig*)	**Overeducation**	**R&D intensity**	**Foreign Direct Investments**	**Openness to Trade**
sectors with high levels of	0.069	**0.128****	**0.122*****	**0.141****
	(0.046)	**(0.059)**	**(0.042)**	**(0.060)**
sectors with low levels of	**0.084****	0.039	0.075	0.059
	(0.039)	(0.067)	(0.049)	(0.044)
Other labour variables	YES	YES	YES	YES
Other controls	YES	YES	YES	YES
Time effects	YES	YES	YES	YES
Fixed effects	YES	YES	YES	YES
Observations	457	457	439	457
Number of id2	46	46	44	46
AR(1) p-value	0.002	0.002	0.001	0.001
AR(2) p-value	0.775	0.828	0.566	0.625
Hansen test p-value	0.120	0.724	0.249	0.106

The dependent variable is the log of citations received in the last 4 years. All variables are in logs and are 1-year lagged. All models include additional control variables. All models include time, country and industry dummies. In all models (one-step) robust GMM-SYS estimators are used. In the GMM-SYS estimates all labor variables are considered as endogenous, while the control variables (age, R&D, openness to trade and the stock of citations) are considered as exogenous. The collapse option (Roodman, 2009) is implemented for the GMM-style instruments in levels and in differences, only the first two lags of the endogenous variables are used. Standard errors are clustered at the country-industry level, *** p<0.01, ** p<0.05, * p<0.1

Table 9. Skilled migrants: Diversity and country of origin

	(1)	(2)		(3)
the contribution of skilled immigrants (*E_Tedu_mig*)	GMM -SYS **diversity (geographical areas)**	GMM -SYS **diversity (individual countries)**	the contribution of skilled immigrants	GMM -SYS **EU vs NOEU**
sectors with higher than average share of	**0.131****	**0.107****	*E_Tedu_mig EU*	**0.036*****
	(0.058)	**(0.048)**		**(0.010)**
sectors with lower than average share of	0.067	0.061	*E_Tedu_mig NOEU*	**0.014***
	(0.047)	(0.038)		**(0.008)**
Other labour variables	YES	YES	Other lab. variables	YES
Other controls	YES	YES	Other controls	YES
Time effects	YES	YES	Time effects	YES
Fixed effects	YES	YES	Fixed effects	YES
Observations	457	457	Observations	457
Number of id2	46	46	Number of id2	46
AR(1) p-value	0.001	0.002	AR(1) p-value	0.003
AR(2) p-value	0.773	0.905	AR(2) p-value	0.897
Hansen test p-value	0.117	0.183	Hansen test p-value	0.965

The dependent variable is the log of citations received in the last 4 years. All variables are in logs and are 1-year lagged. All models include additional control variables. All models include time, country and industry dummies. In all models (one-step) robust GMM-SYS estimators are used. In the GMM-SYS estimates all labor variables are considered as endogenous, while the control variables (age, R&D, openness to trade and the stock of citations) are considered as exogenous. The collapse option (Roodman, 2009) is implemented for the GMM-style instruments in levels and in differences, only the first two lags of the endogenous variables are used. Standard errors are clustered at the country-industry level, *** p<0.01, ** p<0.05, * p<0.1.

Appendix A1
The procedure for external instruments

The procedure to create the supply-push stocks of migrants to be used as external instruments is as follows: sticking to the original notation of Card (2001), for each destination country (France, Germany and the UK), we compute the flow ΔMig_{ot} of new migrants from a specific area of origin *o* (we use eight large geographic groups[27]) in year *t*. Then for each destination country we computed the distribution of migrant workers from a specific area of origin in the different sectors of the economy at the beginning of our period of observation.[28] For each sector and for each area of origin we calculated the share λ_{oj}, where *j* indicates the sector in which they are active:

$$\lambda_{oj} = \frac{Mig_{oj94}}{Mig_{o94}}$$

In order to distinguish between skilled and unskilled migrants we calculated for each year *t* the fraction τ_{ogt} of all new immigrants from a specific country of origin *o* that have a specific type *g* of education (either high or middle-low education) as follows:

$$\tau_{ogt} = \frac{\Delta Mig_{ogt}}{\Delta Mig_{ot}}$$

For each sector *j* in each destination country, the supply-push flow of new migrants from a specific country of origin *o* with education *g* is equal to:

$$\Delta Mig\_instr_{ojgt} = \Delta Mig_{ot} * \lambda_{oj} * \tau_{ogt}$$

These supply-push flows of new migrants are aggregated over countries of destinations (differentiated by the two types of education) to obtain the supply-push stocks of total migrants of a specific type of education in sector *j* at time *t*.

[27] Following D'Amuri and Peri (2014) we use the following eight zones of origin: Africa, North America, Central and South America, Middle East and Central Asia, South and Eastern Asia, Eastern Europe, Western Europe, and Oceania.
[28] This corresponds to 1994 for France and the UK and 1996 for Germany.

Appendix A2
The choice of GMM-SYS and the choice of the correct lags

In order to check the appropriateness of the GMM-SYS estimator for our specific variables (as opposed to the GMM-DIF), in Table (A2) we test the predictive power of lagged levels and lagged differences of each of the labour variables. In the upper panel of Table (A2) we present the first stage results of a fixed effects 2SLS estimation of equation (3) in levels, in which we instrument separately each of the four labour variables with their lagged differences. On average the results show that lagged differences have a good predictive power, as shown by the significance of the coefficients. However, we find that for educated migrants the first, second and third lagged difference can be used as suitable instruments, while for non-educated migrants only the second and third lagged differences are relevant. When we check for educated natives we find that only the first lagged difference is significant, while for non-educated natives the first and second lagged differences are significant. The lower panel of Table (A2) tests whether when we transform equation (3) in first-differences, lagged levels of the endogenous labour variables are good instruments. In line with our expectations we find that lagged levels are not sufficiently powerful instruments for the variables in differences, due to the persistency of the labour variables (Blundell and Bond, 1998): almost all the lagged levels are insignificant in the first stage, with the exception of educated natives, in which, instead, the one and two-years lagged levels are significant. These results confirm that the GMM-SYS specification is legitimised by the relevance of lagged differences as instruments for the equation in levels.

Table A2. First-stage on the lag specification

	(1)	(2)	(3)	(4)
	specification in levels			
Regressors	*Tedu_natives*	*Tedu_migrants*	*noTedu_natives*	*noTedu_mig*
Δx_{t-1}	**0.275*** **	**0.170** **	**0.410*** **	0.101
	(0.078)	**(0.082)**	**(0.153)**	(0.069)
Δx_{t-2}	0.033	**0.211*** **	**0.368** **	**0.218*** **
	(0.087)	**(0.072)**	**(0.155)**	**(0.078)**
Δx_{t-3}	0.051	**0.181*** **	0.145	**0.183*** **
	(0.090)	**(0.054)**	(0.147)	**(0.056)**
F-statistics	3.939	5.847	4.171	4.802
Hausman test p-value	0.400	0.846	0.317	0.005
Hansen test p-value	0.466	0.985	0.850	0.100
Observations	304	296	304	300

	(1)	(2)	(3)	(4)
	specification in first differences			
Regressors	*Tedu_natives*	*Tedu_migrants*	*noTedu_natives*	*noTedu_mig*
x_{t-2}	**-0.184*** **	0.009	-0.052	0.124
	(0.062)	(0.078)	(0.076)	(0.087)
x_{t-3}	**0.196** **	-0.010	0.087	-0.082
	(0.077)	(0.085)	(0.117)	(0.075)
x_{t-4}	-0.006	-0.048	-0.013	0.047
	(0.061)	(0.0666)	(0.077)	(0.064)
F-statistics	3.54	0.411	2.974	1.628
Hausman test p-value	0.558	0.860	0.088	0.933
Hansen test p-value	0.657	0.422	0.149	0.014
Observations	302	295	302	299

The estimates in the upper panel report the results from a first-stage instrumental variable estimation of equation (3) in levels. Each of the columns reports the results of first stage estimates in which only one of the endogenous variables is instrumented with its own lags in differences. The estimates in the lower panel report the results from a first-staqe instrumental variable estimation of equation (3) in differences. Each of the columns reports the results of first stage estimates in which only one of the endogenous variables is instrumented with its own lags in levels. The F-statistics refer to the first-stage estimation. The Hansen test reports a test of over-identifying restrictions on the goodness of the instruments (the null-hypothesis is that instruments are valid). The Hausman test checks for the exogeneity of the instrumented variable in equation (3), the null hypothesis is that the instrumented regressor is exogenous. Standard errors are clustered at the country-industry level, *** p<0.01, ** p<0.05, * p<0.1.

Appendix A3
Classifying sectors on the basis of the intensity of R&D, FDI, Openness to Trade and Diversity

In the following table we classify sectors on the basis of their R&D intensity. In Table X instead we classifiy sectors on the basis of their average levels of FDI, openness to trade and diversity.

Table A3. Definition of high tech and low tech sectors

Low tech sectors	
15-16	Food products, beverages and tobacco
17-19	Textiles, textile products, leather and footwear
20	Wood and products of wood and cork
21	Paper and paper products
25	Rubber and plastics products
26	Other non-metallic mineral products
27	Basic metals
28	Fabricated metal products, except machinery and equipment
High tech sectors	
24	Chemicals and chemical products
29	Machinery and equipment, nec
30	Office, accounting and computing machinery
31	Electrical machinery and apparatus
32	Radio, television and communication
33	Medical, precision and optical instruments
34	Motor vehicles, trailers and semi-trailers
35	Other transport equipment

Table A4. Classification of sectors on the basis of the intensity of FDI, Openness and Diversity

	LEVEL OF FDI (in %)					
	FRANCE		GERMANY		UK	
	AVERAGE (ACROSS COUNTRIES)					
	21.64					
sector						
15-16	13.50	low	4.81	low	18.92	low
17-19	14.71	low	6.21	low	8.04	low
20	9.40	low	3.03	low	5.44	low
21	.		.		22.85	high
24	44.25	high	16.85	low	37.10	high
25	30.92	high	11.14	low	19.11	low
26	26.49	high	7.49	low	17.30	low
27	40.85	high	11.36	low	19.75	low
28	14.41	low	4.90	low	9.44	low
29	39.49	high	9.86	low	25.73	high
30	59.30	high	32.44	high	44.89	high
31	32.62	high	7.84	low	26.95	high
32	35.67	high	23.92	high	36.35	high
33	22.71	high	11.73	low	24.04	high
34	27.18	high	11.99	low	56.26	high
35	21.11	low	25.69	high	21.27	low

	OPENNESS TO TRADE					
	FRANCE		GERMANY		UK	
	AVERAGE (ACROSS COUNTRIES)					
	100.10					
sector						
15-16	42.26	low	38.11	low	38.51	low
17-19	121.29	high	197.36	high	149.20	high
20	39.38	low	37.37	low	45.54	low
21	71.99	low	81.52	low	63.29	low
24	99.91	low	105.26	high	104.65	high
25	61.08	low	62.40	low	50.35	low
26	36.32	low	38.90	low	33.62	low
27	90.89	low	86.15	low	105.68	high
28	27.43	low	34.81	low	32.17	low
29	101.51	high	79.58	low	106.08	high
30	328.22	high	309.77	high	217.86	high
31	92.66	low	64.32	low	107.83	high
32	123.23	high	197.73	high	206.35	high
33	89.34	low	112.38	high	125.99	high
34	91.05	low	79.48	low	125.17	high
35	90.31	low	162.44	high	112.24	high

	DIVERSITY INDEX (*using countries of origins*)					
	FRANCE		GERMANY		UK	
	AVERAGE (BY COUNTRY)					
	0.146		0.431		0.659	
sector						
15-16	0.096	low	0.476	high	0.810	high
17-19	0.254	high	0.562	high	0.989	high
20	0.121	low	0.291	low	0.359	low
21	0.092	low	0.458	high	0.688	high
24	0.166	high	0.423	low	0.650	low
25	0.134	low	0.436	high	0.594	low
26	0.080	low	0.276	low	0.579	low
27	0.129	low	0.348	low	0.500	low
28	0.117	low	0.433	high	0.678	high
29	0.138	low	0.389	low	0.677	high
30	0.243	high	0.435	high	0.808	low
31	0.141	low	0.473	high	0.651	low
32	0.182	high	0.525	high	0.778	high
33	0.147	high	0.430	low	0.560	low
34	0.142	low	0.535	high	0.712	high
35	0.147	high	0.415	low	0.507	low

	DIVERSITY INDEX (*using 6 geographical areas*)					
	FRANCE		GERMANY		UK	
	AVERAGE (BY COUNTRY)					
	0.133		0.374		0.551	
sector						
15-16	0.091	low	0.421	high	0.660	high
17-19	0.227	high	0.498	high	0.855	high
20	0.117	low	0.259	low	0.349	low
21	0.087	low	0.416	high	0.587	high
24	0.141	high	0.358	low	0.502	low
25	0.124	low	0.382	high	0.524	low
26	0.078	low	0.249	low	0.505	low
27	0.129	low	0.311	low	0.421	low
28	0.107	low	0.371	low	0.591	high
29	0.122	low	0.318	low	0.541	low
30	0.216	high	0.383	high	0.630	high
31	0.129	low	0.394	high	0.543	low
32	0.162	high	0.439	high	0.636	high
33	0.135	high	0.376	high	0.457	low
34	0.134	high	0.439	high	0.600	high
35	0.125	low	0.377	high	0.418	low

Table A5. List of countries of origin used for the diversity index

Europe	
1	Austria
2	France
3	Germany
4	Greece
5	Italy
6	Netherlands
7	Portugal
8	Spain
9	UK
10	Poland
11	Romania
12	Former Yugoslavian republics
13	Other eastern European countries (HU CZ SK)
14	Small Western European countries (BE DK FI IE SE LU MT CY)
15	European countries outside of the EU: BG AL NOR CH
Africa and Maghreb	
16	Morocco
17	Other African Countries
Middle East and Central Asia	
18	Turkey
19	Middle East (Israel, Sirya, Iraq, Lebanon, Saudi Arabia, UAE)
20	Iran and Pakistan
21	Russia and former Soviet republics
South and Eastern Asia	
22	Indochine (Vietnam, Laos, Thailand, Cambodia)
23	India and Bangladesh
24	China, Korea, Japan, Malaysia and Indonesia
Central & South America	
25	All central and south American countries
North America	
26	US and Canada
Oceania	
27	Australia and New Zealand

Table A6. Full coefficients of the regressions in Table (7)

	(1) OLS	(2) IV -CARD	(3) IV -CARD	(4) GMM-SYS	(5) GMM-SYS
Variables	*all exog*	*E_noTedu_im endog*	*E_Tedu_im endog*	*E_noTedu_im & E_Tedu_im endog*	*all labor endog*
E_Tedu_mig$_{t-1}$	**0.036***	**0.037***	-1.551	**0.067***	**0.089****
	(0.021)	**(0.022)**	(1.362)	**(0.038)**	**(0.045)**
E_noTedu_mig$_{t-1}$	-0.048	**-0.212****	0.081	-0.170	**-0.338***
	(0.052)	**(0.103)**	(0.162)	(0.283)	**(0.201)**
E_Tedu_nat$_{t-1}$	-0.013	0.045	0.173	-0.082	**0.292****
	(0.083)	(0.099)	(0.329)	(0.107)	**(0.141)**
E_noTedu_nat$_{t-1}$	-0.129	-0.100	0.981	0.199	**-0.752****
	(0.195)	(0.186)	(1.044)	(0.270)	**(0.354)**
avg_age_Tedu_mig$_{t-1}$	0.021	-0.010	-0.381	-0.044	-0.101
	(0.113)	(0.106)	(0.513)	(0.187)	(0.187)
avg_age_noTedu_mig$_{t-1}$	0.283	0.402	-0.359	0.235	0.315
	(0.198)	(0.286)	(0.878)	(0.162)	(0.214)
avg_age_Tedu_nat$_{t-1}$	-0.341	-0.421	-2.235	0.126	-0.292
	(0.639)	(0.647)	(2.196)	(0.424)	(0.444)
avg_age_noTedu_nat$_{t-1}$	-2.502**	-2.348**	-1.732	-2.011**	-4.170**
	(1.016)	(1.059)	(3.453)	(0.851)	(1.652)
R&D$_{t-1}$	0.279***	0.291***	0.907	0.234***	0.312***
	(0.081)	(0.082)	(0.644)	(0.064)	(0.105)
stock_citations$_{t-1}$	0.118	0.089	-0.287	0.275***	0.160
	(0.170)	(0.175)	(0.433)	(0.076)	(0.098)
open$_{t-1}$	-0.526**	-0.576***	-0.247	-0.607***	-1.017***
	(0.197)	(0.199)	(0.425)	(0.217)	(0.304)
time effects	YES	YES	YES	YES	YES
fixed effects	YES	YES	YES	YES	YES
First stage		*lognosk_im*	*logsk_im*		
E_Tedu_mig_card$_{t-1}$		-	0.085		
		-	(0.076)		
E_noTedu_mig_card$_{t-1}$		0.857***	-		
		(0.147)	-		
Angrist-Pischke F test of excl. instr:		33.97	1.25		
p-value		0.000	0.269		
Hausman endog. test p-value		0.142	0.060		
Observations	457	451	448	457	457
Number of id	46	45	45	46	46
R-squared	0.795	-	-		
num. instruments				44	47
AR(1) p-value				0.002	0.001
AR(2) p-value				0.758	0.578
Hansen test p-value				0.170	0.426

E_Tedu_mig $_{t-1}$*, lags used (1, 2)*		
Hansen test excluding group: chi2	0.793	0.306
Difference (null H = exogenous): chi2	0.096	0.530
E_noTedu_mig $_{t-1}$*, lags used (2, 3)*		
Hansen test excluding group: chi2	0.166	0.534
Difference (null H = exogenous): chi2	0.213	0.274
E_Tedu_nat $_{t-1}$*, lags used (1, 1)*		
Hansen test excluding group: chi2		0.296
Difference (null H = exogenous): chi2		0.629
E_noTedu_nat $_{t-1}$*, lags used (1, 1)*		
Hansen test excluding group: chi2		0.516
Difference (null H = exogenous): chi2		0.287

The dependent variable is the log of citations received in the last 4 years. All variables are in logs and are 1-year lagged. All models include additional control variables, as displayed in Table (A4). All models include time, country and industry dummies. In column (1) the OLS estimator is implemented. In columns (2) and (3) two-stage least squares estimators are implemented. In the panel below first-stage coefficients are reported. The Angrist-Pischke test of excluded instruments reports the probability that excluded instruments in columns (2) and (3) are weak, the Hausman test reports the probability that the instrumented variables are endogenous. In columns (4) and (5) (one-step) robust GMM-SYS estimators are used. The collapse option (Roodman, 2009) is implemented for the GMM-style instruments in levels and in differences. In column (4) the endogenous variables are *E_Tedu_mig*, *E_noTedu_mig*. On the basis of the results in Table (10) *E_Tedu_mig* is instrumented with one and two year lags, while *E_noTedu_mig* is instrumented with two and three years lags. In column (5) the endogenous variables are *E_Tedu_mig*, *E_noTedu_mig*, *E_Tedu_nat*, *E_noTedu_nat,*. On the basis of the results in Table (10) both *E_Tedu_nat* and *E_noTedu_nat* are instrumented with one year lags, while *E_Tedu_mig*, *E_noTedu_mig* are instrumented with the same lags as in column (4). In columns (4) and (5) all the additional control variables are considered as exogenous. Standard errors are clustered at the country-industry level, *** p<0.01, ** p<0.05, * p<0.1.

Appendix B - Data description

Patents data come from the PATSTAT-KITES database

PATSTAT (EPO Worldwide PATent STATistical Database) is a patent database, run by the European Patent Office (EPO) developed in cooperation with the World Intellectual Property Organisation (WIPO), the OECD and Eurostat. PATSTAT provides raw patent data coming from around 90 patent offices worldwide, including, of course, the most important and largest ones such as the European Patent Office (EPO) and the United States Patent and Trademark Office (USPTO). The data set includes the full set of bibliographic variables concerning each patent application. PATSTAT is provided in a raw format. Data coming from PATSTAT has, therefore, been thoroughly elaborated by KITES (Bocconi University: http://db.kites.unibocconi.it/) to produce a clean and harmonized database. Data processing consisted mainly in a thorough work of cleaning and standardization rough information provided by the EPO. The aggregation of patent technological classifications (so called IPC classes) into NACE Rev. 1 fields follows Schmoch *et al.* (2003)[29]

UK Labour Force Survey

The British Quarterly Labour Force Survey (QLFS) is a quarterly sample survey of households living at private addresses in Great Britain. The QLFS is conducted on a quarterly basis and aims to obtain a sample of around 60,000 households every quarter. Since 1992 respondents are interviewed in five successive waves, thus approximately a fifth of the sample in each quarter will contain individuals from each of the five waves. Every quarter one wave of approximately 12,000 leaves the survey and a new wave enters. The rotational element to the QLFS creates an 80 percent overlap between quarters and thus 20 percent of the sample enter and exit the survey each quarter.

The survey contains data on among other variables: employment and self-employment; full-time and part-time employment; second jobs; average age; economic activity; occupations and industry sectors and education.

French Labour Force Survey

The French Labour Force Survey was launched in 1950 and applied in 1982 as an annual survey. Redesigned in 2003, the survey is a continuous survey providing quarterly results. The survey covers private households in metropolitan France. It includes a part of the population living in collective households, persons who have family ties with private households. Participation in the survey is compulsory. The resident population comprises persons living in the French metropolitan territory.

The household concept used is that of the 'dwelling household': a household means all persons living in the same dwelling. It may consist of a single person or of two families living under the same roof.

The survey provides longitudinal data on households and individuals. Persons average aged fifteen years or over are interviewed. Data refer to the number of persons who were working during the survey week including employees, self-employed as well as family workers. Data include persons who have a job but are not at work due to illness (less than one year), vacation, labour dispute, educational leave, etc.

German Microcensus

The Microcensus provides official statistics for the population and the labour market in Germany. The Labour Force Survey of the European Union (EU Labour Force Survey) forms an integral part of the Microcensus. The Microcensus supplies statistical information in a detailed subject-related and regional breakdown on the population structure, the economic and social situation of the population, families, consensual unions and households, on employment, job search, education/training and continuing education/training, the housing situation and health. The German Microcensus includes 1% of the resident population in the former West Germany, and is a large, representative, random sample containing comprehensive information on individual and household characteristics.

References of Appendix A and B

D'Amuri, F., Peri, G., (2014), Immigration, Jobs and Employment Protection: Evidence from Europe before and during the Great Recession, *Journal of the European Economic Association*, 12(2), 432–464.

Schmoch, U., Laville, F., Patel P., Frietsch R. (2003), Linking Technologies to industrial Sectors. Final report to the European Commission, DG Research.

[29] ftp://ftp.cordis.europa.eu/pub/indicators/docs/ind_report_isi_ost_spru.pdf