OECD Regional Development Papers

Migration and regional innovation in Australia

This paper provides evidence on the impact of international migrants on regional innovation. The study combines administrative individual-level data covering all Australian residents with data on intellectual property rights applications such as patents, trademarks, and design rights. The analysis uses a standard shift-share instrument based on past migrant settlements to identify the causal effects of migration on innovation. Its four main findings are the following: First, on average, a one percentage point increase in the regional employment share of higher-educated migrants relative to total employment leads to a 4.8% rise in regional patent applications in the medium run (five years). Second, while migrants of all skill and education levels have a positive impact on patenting, those in scientific occupations have the largest effect. Third, regions with lower levels of patenting benefit relatively more from increases in migration compared to those with higher patenting levels. Fourth, there is no effect of migration on trademarks or design rights applications.

JEL codes: R10, O34, R23
Keywords: Australia, innovation, regions

Contact:
Cem Özgüzel- cem.ozguzel@oecd.org
Jasper Hesse- jasper.hesse@oecd.org

ABOUT THE OECD

The OECD is a multi-disciplinary inter-governmental organisation with member countries which engages in its work an increasing number of non-members from all regions of the world. The Organisation's core mission today is to help governments work together towards a stronger, cleaner, fairer global economy. Through its network of specialised committees and working groups, the OECD provides a setting where governments compare policy experiences, seek answers to common problems, identify good practice, and co-ordinate domestic and international policies. More information available: www.oecd.org.

ABOUT OECD REGIONAL DEVELOPMENT PAPERS

Papers from the Centre for Entrepreneurship, SMEs, Regions and Cities of the OECD cover a full range of topics including regional statistics and analysis, urban governance and economics, rural governance and economics, and multi-level governance. Depending on the programme of work, the papers can cover specific topics such as regional innovation and networks, sustainable development, the determinants of regional growth or fiscal consolidation at the subnational level. OECD Regional Development Papers are published on http://www.oecd.org/cfe/regional-policy.

This paper is authorised for publication by Lamia Kamal-Chaoui, Director, Centre for Entrepreneurship, SMEs, Regions and Cities, OECD.

Executive summary

Migrants are an integral part of the Australian population. In 2021, almost three out of ten Australian residents (29%) were born abroad, positioning Australia as the country with the third-highest migrant share among OECD countries. On average, foreign-born residents in Australia are highly educated and well-integrated into the labour market. Their level of education exceeds the average educational attainment of the native population.

The arrival of higher-educated migrants into the labour supply of Australian regions increases patenting activity. On average, a one percentage point increase in the regional employment share of higher-educated migrants relative to total employment leads to a 4.8% rise in regional patent applications in the medium run (five years). In contrast, the inflow of higher-educated migrants does not affect other innovation measures, such as trademarks and design rights.

The positive influence on innovation is most significant for migrants in scientific occupations and previously less innovative regions. Although the inflow of migrants across all education levels positively impacts patent applications, those in scientific occupations have the most pronounced effect. Furthermore, while migration positively affects all regions in Australia, the benefits are most pronounced in less innovative regions.

The overall positive impact of migration on regional innovation reaffirms the significant contribution of migrants to the Australian economy. The analysis suggests that the inflow of higher-educated and higher-skilled migrants, facilitated by the selective migration policies, positively influenced Australia's innovation activities.

Acknowledgements

This paper was prepared by the OECD Centre for Entrepreneurship, SMEs, Regions and Cities (CFE) led by Lamia Kamal-Chaoui, Director. The work was conducted as part of the OECD Regional Development Policy Committee's program of work with financial support from the Centre for Population at the Australian Treasury.

Cem Özgüzel, Economist at CFE, coordinated the preparation of the paper under the guidance of Ana Moreno-Monroy, Head of the Statistics and Territorial Analysis Unit (CFE) and the supervision of Rüdiger Ahrend, Head of the Economic Analysis, Data and Statistics Division at CFE. Gabriel Chaves Bosch, Jasper Hesse, and Cem Özgüzel (all CFE) drafted the paper. The authors are thankful to Riccardo Crescenzi (London School of Economics) and Giovanni Peri (University of California, Davis) for helpful feedback. Thanks are also due to Nadim Ahmad, Lars Ludolph and Michelle Marshallian (all CFE), Ana Damas de Matos (OECD Directorate for Employment, Labour and Social Affairs, ELS), and Ben Westmore (OECD Economics Department, ECO) for their valuable comments. Eric Gonnard (CFE) provided statistical support. Pilar Philip (CFE) prepared the paper for publication.

The paper also benefited from valuable comments by Patrick Fazzone IV, Rodrigo Rodrigues, Ian South, and Kasey Stanfield (all Centre for Population), Joe Castellino, Kathleen Cross, Linda Velzeboer (Department of Infrastructure, Transport, Regional Development, Communications and the Arts). Anna Maria Mayda (Georgetown University), Ceren Özgen (University of Birmingham) and Riccardo Turati (Autonomous University of Barcelona) provided guidance throughout the project as scientific advisors. The report benefited from analytical input by Aneeq Sarwar (Economic Society Australia). The Secretariat appreciates the feedback provided by the delegates during the 45th session of the Working Party on Territorial Indicators (WPTI) on 14 November 2023.

Michel Beine (University of Luxembourg) and Christopher Parsons (University of Western Australia) kindly shared valuable data on the historical settlement of migrants across Australia. IP Australia provided data on intellectual property rights applications, and the Australian Bureau of Statistics (ABS) provided access to the Multi-Agency Data Integration Project (MADIP), which made the analysis possible.

This paper is the fourth output of a multi-annual collaboration between the Australian Centre for Population and the OECD. The first working paper of the project provides a detailed descriptive overview of migrants in Australia and the context of productivity and the labour market (OECD, 2023[1]). The second examines the contribution of migrants to regional labour productivity differences in Australia (OECD, 2023[2]). The third paper evaluates the impact of migration on regional labour markets (OECD, 2023[3]).

Table of contents

FIGURES

TABLES

1 Introduction

With 29% of the population born abroad as of 2021 (OECD, 2022[4]), Australia has the third-highest migrant share among OECD countries.[1] Migrants can contribute to national economic growth through various channels, for instance, by bringing new skills and ideas, as well as fostering innovation. Yet, evidence on the potential contribution of migration to innovation remains scant.

This paper measures the impact of migration on regional innovation across Australian Statistical Area 4 (SA4) regions, focusing on higher-educated migrants who are most likely to innovate (Hunt, 2011[5]; Hunt and Gauthier-Loiselle, 2010[6]; Bernstein et al., 2022[7]).[23] It uses comprehensive administrative individual-level data covering all Australian residents and detailed information on various measures of innovation to assess the causal influence of higher-educated migrants on innovation across Australian regions from 2011 to 2018. The analysis employs a difference-in-difference design comparing innovative activities in regions that received more higher-educated migrants with those that received less higher-educated migrants before and after the arrival of the migrants. As migrants often select their location based on economic opportunities, which can be linked to innovative activities, the study employs an instrumental variable strategy to identify the causal effect of migration on regional innovation.

This paper uses patents, trademarks, and design rights applications as innovation indicators to provide estimates that align with the current migration literature.[4] Innovative activities – which are multifaceted and intricate - tend to cluster geographically, resulting in significant disparities among regions within countries. It is thus essential to employ diverse measures of intellectual property (IP) to capture various aspects of the broader concept of innovation instead of relying on patents that, although widely used in the literature (Ozgen, 2021[8]), may lead to skewed representations as they are concentrated in large cities. Unlike patents, which typically capture innovations in STEM industries, trademarks typically capture innovations

[1] The terms “migrants” and “foreign-born” are used interchangeably throughout this paper. Individuals born outside of their country of residence are considered migrants. Unlike citizenship, this criterion does not change over time, it is not subject to country differences in legislation, and it is thus adequate for international comparisons. In Australia, migration is primarily measured as Net Oversea Migration (NOM), which refers to the net increase or decrease in the Australian population resulting from immigration to and emigration from Australia, irrespective of the individual’s country of birth or nationality.

[2] The term “higher-educated” describe individuals who have obtained at least a college degree. Annex D provides details on the definition.

[3] The Australian Bureau of Statistics (ABS) has designed Statistical Areas 4 (SA4) regions by considering a range of criteria that balance various factors. The two main criteria are population size and commuting patterns. As a result, the 88 SA4 regions considered in this analysis cover the whole of Australia and represent all regional labour markets. SA4 regions exhibit a functional characteristic in terms of capturing labour supply and demand. Throughout this paper, the term “region” refers to SA4 regions unless indicated otherwise.

[4] The analysis follows previous literature (Hunt and Gauthier-Loiselle, 2010[6]) and uses applications rather than granted IP rights because the most immediate step after development of an innovation, typically, is to apply for its IP right.

in the retail or services sector, while design rights are more relevant in manufacturing.[5] Jointly analysing these measures provides a more comprehensive view of regional innovation activities in Australia.[6]

The paper makes four findings:

- **Higher-educated migrants have a positive effect on patent applications across Australian regions.** Concretely, a one percentage point increase in the regional employment of higher-educated migrants relative to total employment leads, on average, to a 4.8% rise in regional patent applications in the medium run (five years).
- **All migrant groups contribute to the positive effect on patent applications.** Migrants of all education levels have a positive impact on patent applications. While the effect is smaller for non-higher-educated migrants, it is most pronounced for higher-educated migrants, particularly those in scientific occupations.
- **The positive impact of migrants on patenting activity is more significant in regions that initially had lower levels of innovation.** Although all regions benefit from the arrival of higher-educated migrants, the increase in patenting activity is notably larger in regions with initially lower levels of innovation. Moreover, migration benefits all regions irrespective of their regional income level or population density.
- **Migrants do not affect trademarks or design rights applications.**[7] The analysis finds no statistically significant regional effect of migration on trademarks or design rights applications.

The paper is organized as follows. The next section presents a brief review of the relevant migration literature. Section 3 presents the data used in the analysis and provides descriptive statistics. Section 4 introduces the empirical strategy, Section 5 presents the results, and Section 6 concludes.

---

[5] Patents and design rights represent novel innovations introduced to the world, while trademarks register innovation that is new to the local market or individual firms (Jensen and Webster, 2009[21]).

[6] To ensure comparability with previous research (Crown, Faggian and Corcoran, 2020[18]; Blit, Skuterud and Zhang, 2020[23]; Hunt and Gauthier-Loiselle, 2010[6]) and due to lack of data at the regional level, this paper does not follow the OSLO definition of innovation developed by the OECD and Eurostat (OECD/Eurostat, 2018[41]).

[7] This finding is consistent with previous research finding migrant-owned firms are more innovative in terms of more R&D spending and patenting activity but no are not more likely to apply for trademarks or copyrights in the US (Brown et al., 2020[29]).

2 Literature

Migration can boost innovation through multiple channels. A recent study finds that migrants in the US generate substantial positive externalities on innovation. While only 16% of all innovators in the US are foreign-born, they contribute to 23% of the national innovation output (Bernstein et al., 2022[7]). Yet, beyond their direct impact on innovation, migrants can also indirectly influence innovation by complementing the efforts of individuals and firms in their host region through increased diversity and positive externalities (Ozgen, 2021[8]; Perez-Silva, Partridge and Foster, 2019[9]). However, the impact of migration on innovation might vary across countries as higher-skilled migrants settle unevenly (Kerr et al., 2016[10]).

Extensive evidence from OECD countries suggests that the impact of migration on regional innovation (like the native population) depends on the characteristics of the migrant population, particularly education composition. Recent studies find that an increase in the presence of higher-educated migrants has a strong and positive effect on regional innovation in US counties (Hunt and Gauthier-Loiselle, 2010[6]), English local authorities (Gagliardi, 2015[11]), French districts (Mayda, Orefice and Santoni, 2022[12]), and, more generally, European countries (Bosetti, Cattaneo and Verdolini, 2015[13]). However, when assessing the effect of an overall increase in the number of migrants, existing evidence is more nuanced, with studies finding no effect in Italian provinces (Bratti and Conti, 2018[14]) or an effect limited to already innovative municipalities across OECD countries (OECD, 2022[15]). Moreover, migration can also reduce patenting activities if firms shift towards labour-intensive activities and reduce their investment in capital following the arrival of lower-skilled migrants in the region (Imbert et al., 2022[16]).[8] Increased innovation is also closely linked to labour productivity, as Box 2.1 illustrates.

Less is known about the impact of migrants on innovation in Australia, as findings from other countries may not hold for Australia given differences in industrial composition and national institutional settings, as well as characteristics and presence of migrants. For instance, recent evidence shows that university-educated migrant men from the same country have higher expected employment rates and weekly earnings in the US than in Australia (Clarke, Ferrer and Skuterud, 2019[17]), which might incentivise most innovative migrants to choose the US over Australia. Furthermore, it remains unclear to what extent the Australian context can effectively leverage the expertise brought by migrants to support innovation. Existing evidence from Australia shows that foreign-born graduates have a positive impact on patents but not on trademarks or design rights (Crown, Faggian and Corcoran, 2020[18]). However, this effect might not be translatable to other migrant groups that attained their education abroad. Hence, the overall impact of arriving migrants with different education levels on innovative activities in Australia remains an open question (Jensen, 2014[19]). Lastly, it is uncertain whether potential positive effects of migrants depend on specific regional characteristics or apply to all regions.

This study looks to fill that gap and provide evidence at the regional level. It utilises individual-level data covering all Australian residents and intellectual property rights information to provide a precise analysis of the regional impact of migration and a nuanced understanding of the contribution of migrants based on

---

[8] These findings are obtained in the context of Chinese prefectures and following an episode of large-scale internal rural-to-urban migration. From a theoretical perspective, a similarly large-scale international migration inflow can also have similar effects on the economy.

their educational levels or occupations. Additionally, it sheds light on uneven effects across regions, a factor often overlooked in such analyses.

Box 2.1. Does innovation affect labour productivity?

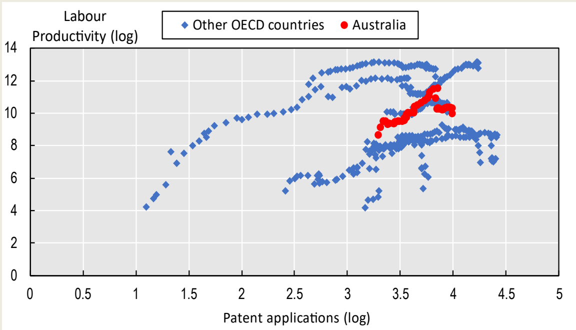
Innovation affects productivity in various ways. On the one hand, it can increase the efficiency of workers, also known as labour productivity (LP) (Sarwar, 2022[20]). On the other hand, innovation can improve the output of capital per unit of labour, leading to an increase in total factor productivity (TFP). Therefore, increasing innovation is key to boosting both LP and TFP.

In fact, innovation, proxied by patent applications, is positively associated with labour productivity. Panel A of Figure 1 illustrates the positive relationship between labour productivity (vertical axis) and patent applications (horizontal axis) in other OECD countries (blue markers) and Australia (red markers) over 47 years (1970-2017). The link between patent applications and labour productivity is stronger in Australia than in the US, Canada, or Türkiye. Panel B of Figure 1 presents the correlation between patent applications and productivity using an ordinary-least-square (OLS) regression, which controls for time-invariant characteristics of each country using country fixed-effects. There is a clear positive association between patent applications and productivity, although the strength of this relationship varies across countries. In Australia, a 1% increase in patent applications is associated with a 7% increase in labour productivity.

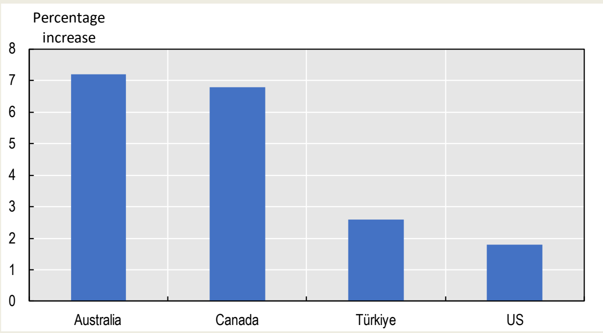
Figure 1. Patent applications and labour productivity

Scatterplot of labour productivity and patent applications (left panel) and impact of patent applications on labour productivity (right panel) across selected OECD countries

Panel A: Patent applications have a strong positive relationship with labour productivity



Panel B: An increase in patent applications leads to an increase in labour productivity



Note: The figure presents a scatterplot of patent applications and labour productivity, both in logarithm (left panel) and the coefficients from a regression of labour productivity on patent applications (right panel). Data include Australia, Canada, Denmark, Japan, South Korea, New Zealand, Norway, Türkiye, and the US between 1970 and 2017 (47 years). The regression from the right panel controls for labour quality, economic stability, and country-time fixed effects.
Source: Data are from Sarwar (2022[20]).

3 Data and preliminary evidence

This section elaborates on the data sources and presents descriptive analysis and preliminary evidence on migration and innovation in Australian regions. It starts by describing the data sources and construction of the analysis sample. Next, it provides an overview of migration and innovation across Australian regions, discussing their relationship in detail.

Data sources and sample construction

Data sources

This study uses rich individual-level administrative panel data provided by the Australian Bureau of Statistics (ABS). The Multi-Agency Data Integration Project (MADIP) dataset compiles information from various ministries related to health, education, government payments, income and taxation, employment, population demographics, migration, as well as Census data. It contains 27.1 million individual records, covering all Australian residents who interacted with the social security system, paid income tax, or engaged with the health system at any point between 2006 and 2020.

The Intellectual Property Government Open Dataset (IPGOD) dataset, maintained by IP Australia, provides information on intellectual property (IP) rights applications used to construct regional innovation measures. This dataset offers detailed geolocated information on the number of patents, design rights, and trademark applications. Each of these measures captures distinct dimensions of innovation. Patents and design rights encapsulate innovations that are new to the world. Trademarks represent innovations that are not new to the world but are new either to the firm or the regional market. These measures display a high correlation with other firm-level indicators of innovation, such as survey-based measures of perceived innovation and R&D expenditure (Jensen and Webster, 2009[21]). Annex A provides a comprehensive description of both data sources.

Constructing the analysis sample

The primary analysis is conducted at the SA4 level, a classification developed by the ABS based on regional labour demand and supply data to reflect local labour market areas.[9] The individual-level residence information is aggregated at the regional level and yearly frequency.

The analysis uses the MADIP dataset to generate socio-economic variables. It restricts the sample to individuals aged 15-64 who were employed at some point between 2011 and 2018 and whose income was above the minimum threshold for a tax declaration. Following the literature, the analysis excludes workers in the public, health, agriculture, and mining sectors, as these sectors are generally less likely to employ

[9] Australia is disaggregated into 89 SA4 regions with a population between 100 000 and 500 000. Following the OECD territorial grid, the SA4 "Other territories" is excluded from the analysis, resulting in a total of 88 SA4 regions.

migrants or apply for design rights, patents, or trademarks. The resulting sample contains nearly 26 million individual-year observations over the 2011 to 2018 period.

Matching the sample with the 2016 Census allows to obtain education information. The primary variable of interest is the inflow of higher-educated migrants, defined as the net change in higher-educated migrant workers relative to the total number of workers in a region, defined based on individual-level residence information.[10] Higher-educated workers are workers with at least a college degree.[11] Additionally, the analysis constructs relevant controls at the regional level, such as industry shares (defined as the share of workers working for one-digit industry groups), the share of higher-educated native workers, and the logarithm of the population count.

IPGOD microdata provides regional innovation measures, such as the number of annual applications for patents, trademarks, and design rights. The dataset provides detailed information about each IP rights application, including the involved parties, their location, and their roles. The analysis considers all patent applications in Australia with at least one applicant located in Australia regardless of their country of birth or nationality. In the years of the analysis, firms submitted about 95% of all patent applications. In cases where applications involve applicants from different locations, the paper assigns a count of one to each location. However, the analysis also conducts robustness tests using fractional counts across locations.[12]

Additional datasets

Measuring the causal impact of migration on innovation requires using instrumental variables based on past settlement patterns of migrants, as explained in Section 4. Historical Census data from 1981 provide reliable information on the migrant distribution across Australian regions. The historical census, adjusted to 2016 borders by the ABS, provides information on the total employed population in 1981 disaggregated by country of birth, industry of employment, and highest post-school qualification, where tertiary education matches with the Census definition. Section 4 provides more details on the construction of the instrument.

---

[10] Lacking information on migrants' workplace locations, the analysis assumes that migrants work in the same region where they reside.

[11] Annex C shows that using a higher-skilled definition based on occupations rather than education does not alter the results, as the two definitions are highly correlated. The analysis considering higher-skilled migrants is not limited to information based on the Census 2016. Hence, it also confirms that the results are not driven by the exclusion of migrants without education information.

[12] The fractional count allocates, for each party in an IP rights application, a share equal to one divided by the total number of applicants for each applicant’s location. It gives a lower weight to innovation happening in subsidiary firms if multi-party applications include parent companies which are typically located in denser metropolitan areas.

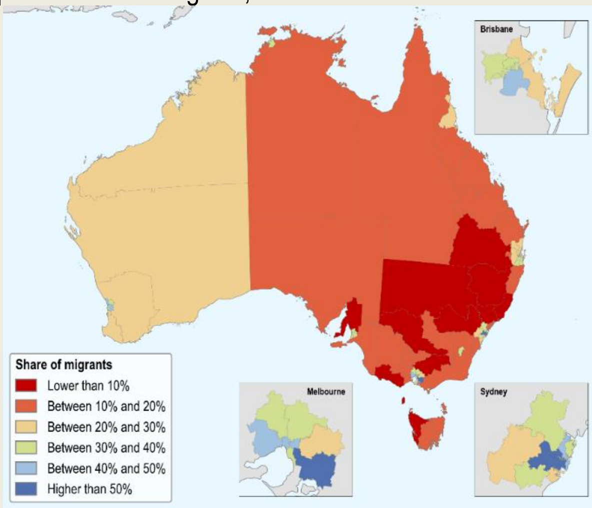
Box 2. The geographical distribution of migrants

Australia is one of the largest migrant-receiving countries in the OECD. In 2021, Australia had the third-highest share of migrants (29%) among OECD countries, after Luxembourg (49%) and Switzerland (30%). This share is substantially higher than the migrant share in Canada (21%), Germany (16%), the UK (14%), and the US (14%) (OECD, 2023[22]). Moreover, the share of migrants in Australia increased by six percentage points from 23% in 2000. Over the same period, the migrant share across the total OECD increased by just four percentage points from 10% to 14%.

The presence of migrants has a pronounced regional dimension in Australia. About 82% of all Australian migrants concentrate in large and midsize metropolitan areas, such as Brisbane, Melbourne, Perth, and Sydney, compared to 66% of natives. Consequently, only 18% of the migrant population lives in non-metropolitan areas, compared to almost one-third (33%) of natives. As a result, migrants constitute a high share of the population in large metropolitan areas (40%) such as Brisbane, Melbourne, Perth, and Sydney. Similarly, in midsize metropolitan regions, the migrant share is around 29%. In non-metropolitan areas, however, less than one-fourth of the population is born abroad, with some regions in the southeast exhibiting values of less than 10% (OECD, 2023[1]).

Figure 2. The presence of migrants varies across Australian labour markets

Share of the foreign-born population across regions, 2016

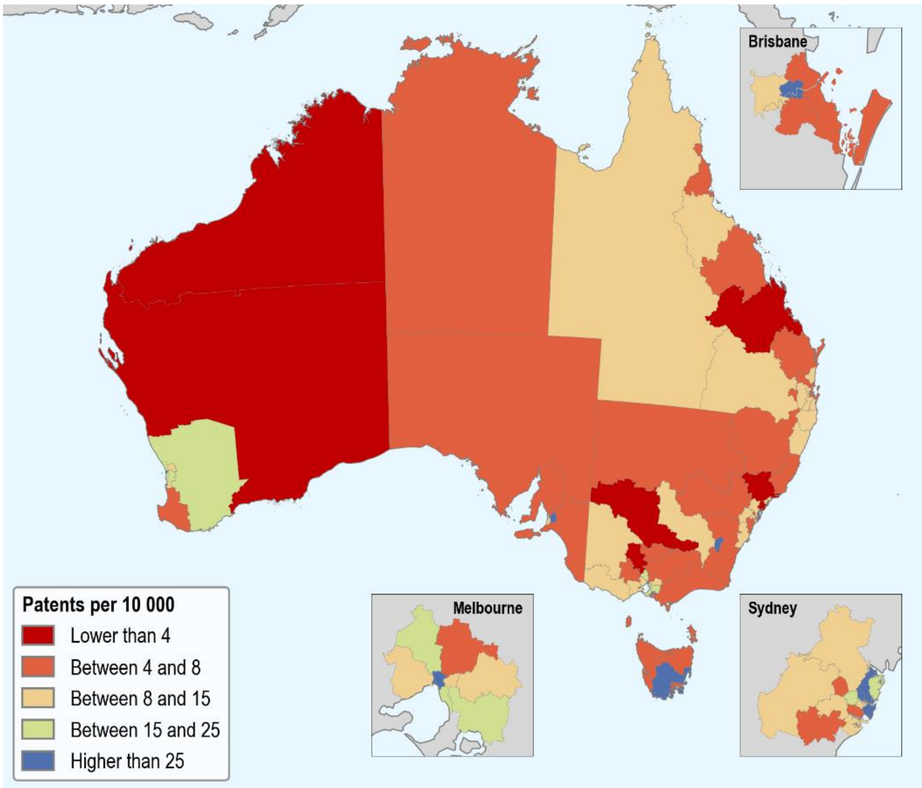


Note: The figure presents the share of foreign-born among the working-age population (15-64 years) in Australia disaggregated by regions. Data are for 2016.
Source: OECD calculations based on the Australian Census of Population and Housing 2016 accessed via ABS Census TableBuilder (accessed May 2022).

Preliminary evidence: Innovation and migration at the regional level Patenting activities in Australia exhibit a substantial geographical dimension. Figure 3 illustrates the number of patent applications per 10 000 workers across Australian regions in 2018. Patenting activity per workers varies widely, ranging from around one in Murray to above 70 in districts of Sydney, Melbourne, and Perth. The major cities Adelaide, Brisbane, Melbourne, Perth and Sydney are, on average, the most innovative, with almost 19 patents per 10 000 workers. This value is almost double the value of the remaining regions.

Figure 3. Geographical distribution of patenting activity across Australian regions

Number of patents per 10 000 workers in Australian regions, 2018



Note: The figure presents the cumulative number of patents per 10 000 workers in Australia disaggregated by regions. Data are for 2018.
Source: OECD calculations based on MADIP and IPGOD (accessed July 2023).

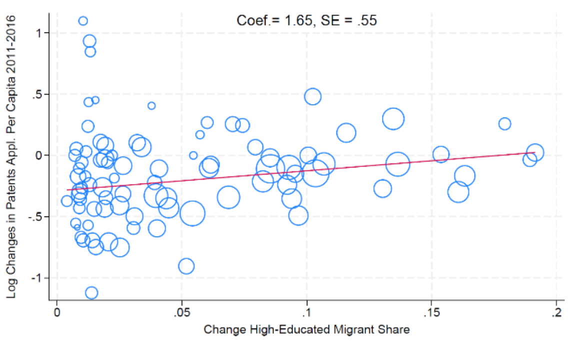
The regional presence of higher-educated migrants only shows a positive association with patents but not with trade markets and design rights. Figure 4 provides scatterplots illustrating the correlations between higher-educated migrant shares and IP rights applications across Australian regions. While correlations provide insights into the relationship between migration and regional innovation, they do not reveal the true impact of migration on regional innovation. As indicated by the red slope line, a positive relation exists between the increases in the share of higher-educated migrants among the total workforce and increases in patent applications per worker. At face value, the correlation would imply that a one percentage point increase in the share of higher-educated migrants among the workforce correlates with an average 1.65% increase in patent applications per worker. Conversely, the flat slope line suggests a lack of a clear association for trademarks and design rights.

However, these figures do not necessarily indicate a causal relationship between migration and regional innovation. As migrants may settle in already innovative urban areas, these associations may be correlational and not causal. The following section details the empirical strategy employed in this paper to overcome this issue.

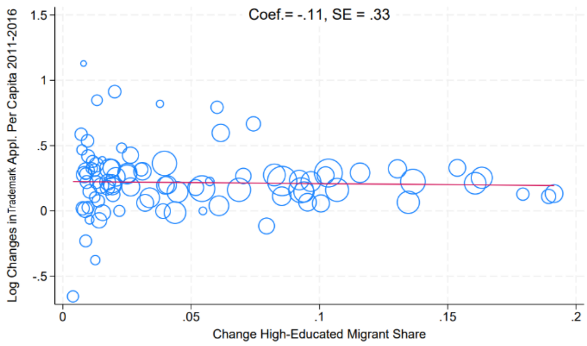
Figure 4. Correlations between changes in IP rights applications and higher-educated migrant shares

Changes in IP rights applications and share of the higher-educated foreign-born population across regions, 2011-2016

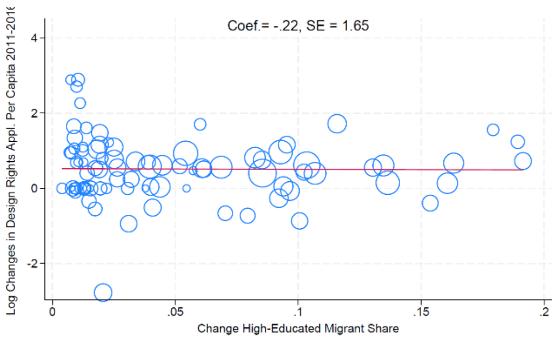
Panel A: Patents



Panel B: Trademarks



Panel C: Design Rights



Note: The figures present scatterplots showing the correlations of changes in the logarithm of IP rights applications per worker(vertical axis) and changes in higher-educated migrant shares across Australian regions (horizontal axis) between 2011 and 2016. Panel A refers to patents, Panel B to trademarks, and Panel C to design rights. Each circle represents a region, and its size corresponds to its population in 2011. Lines of best fit are calculated from OLS regressions weighted by population in 2011. Estimated slope coefficients and robust standard errors are provided on top.
Source: OECD calculations based on the MADIP and IPGOD 2022 data (accessed July 2023).

Summary statistics based on the sample

Table 1 shows the unweighted summary statistics of the main variables used in the analysis, including changes from 2011 to 2016 across regions. On average, the regional share of migrants in the workforce is 24%, and the regional share of workers with college degrees is 21% (comprising 8% migrants and 13% natives). However, notable regional disparities exist. For example, the regional share of higher-educated migrants in the workforce ranges from 1% to 30%. Similarly, substantial differences exist in IP rights applications, with some regions showing minimal innovation while others have application rates per worker ten times higher than the average. Across regions, trademarks are the most common, followed by patents. In contrast, design rights are less common, reflecting their specific nature, primarily used by the manufacturing industry.

Panel B of Table 1 presents the changes in the shares of higher-educated migrants and natives from 2011 to 2016 across regions and the changes in IP rights applications per worker. The growth in the number of higher-educated migrants is stronger than that of higher-educated natives both in levels and in relative terms due to a substantial inflow of higher-educated migrants between 2011 and 2016 and a smaller initial higher-educated migrant population. Moreover, there is a slight decrease in patent applications per worker, while trademark applications per worker experience a substantial increase. Moreover, design rights applications show a small increase.

Table 1. Descriptive statistics (unweighted)

Variable	Mean	Std. Dev.	Min	Max
Panel A: Levels, yearly variables (2011-2018)				
Migrant share	0.239	0.134	0.065	0.607
Share of higher-educated migrants over total workers	0.076	0.070	0.009	0.318
Share of higher-educated natives over total workers	0.125	0.610	0.049	0.309
Number of patents applications per 10 000 workers	25.28	25.47	0	214.19
Number of trademarks applications per 10 000 workers	147.57	153.24	0	1322.20
number of design rights applications per 10 000 workers	6.94	10.07	0	98.37
Panel B: Relative changes, 2011-2016				
Change in share of higher-educated migrants over total workers	0.045	0.043	0.003	0.174
Change in share of higher-educated natives over total workers	0.038	0.025	0.009	0.129
Change in number of patents applications per 10 000 workers	-3.65	11.37	-40.42	62.44
Change in number of trademarks applications per 10 000 workers	42.76	54.44	-65.73	250.04
Change in number of design rights applications per 10 000 workers	5.39	10.61	-18.50	80.58

Note: The table presents descriptive statistics on regional measures of innovation and migration. Panel A contains yearly variables across 88 regions from 2011 to 2018, amounting to a total of 704 observations. In all variables, the denominator is the total (migrant and native) employed population. The mean value is calculated across all 704 year-region observations and is not weighted by population size. Panel B contains changes from 2011 to 2016, divided by the employed population in each region in 2011. Each statistic in Panel B is calculated in a cross-section of 88 regions. Variables per 10 000 workers are per-worker variables multiplied by 10 000. In both panels, the column "Mean" presents the mean value of the sample, while the columns "Min" and "Max" display the value of the region with the lowest and highest changes, respectively.
Source: OECD calculations based on MADIP and IPGOD (accessed July 2023).

4 Empirical strategy

This section presents the empirical strategy used to estimate the impact of migration on innovation in Australian regions. First, it explains the employed empirical model. Second, it discusses the empirical challenges in measuring the causal impact of migration on regional innovation and explains the methods used to address them.

Empirical model

The analysis adopts a difference-in-difference approach to measure the impact of migration on regional innovation, comparing regions that received more higher-educated migrants with those that received less.[13] The strategy employs a first-differences regression model which consists of measuring the changes over time to eliminate the influence of certain time-invariant location-based (e.g., regional infrastructure or population density) and group-based characteristics (e.g., age, sex). Accounting for such time-invariant characteristics eliminates any factor specific to these places or groups that may affect the relationship between migration and regional innovation.

Following the literature, the estimation uses five-year changes to capture medium-run effects (Hunt and Gauthier-Loiselle, 2010[6]):

$$\Delta \log Y_{r,t+5} = \beta_0 + \beta_1 \Delta Migrants^{HE}_{r,t} + \beta_2 X_{r,t} + \alpha_t + \Delta\varepsilon_{r,t} \quad (1)$$

Where $\Delta \log Y_{r,t+5}$ is the logarithm change in the number of applications in year $t+5$ and region r per employed population at baseline year t for each period.[14] Following the literature, in the baseline, the dependent variable (patent, trademark, and design rights applications) measures the changes in the IP rights applications not in the same year (i.e., time *t*) but after five years (i.e., time *t+5*) to allow time between the migration-induced change in labour inputs and IP rights applications (Hunt and Gauthier-Loiselle, 2010[6]).[15] Recent evidence for patents argues that the influence of newly arriving migrants reaches its peak on patent applications four years after arrival (Blit, Skuterud and Zhang, 2020[23]). Robustness checks

---

[13] In the migration literature, this is known as a spatial correlations approach (Dustmann, Schonberg and Stuhler, 2016[37]).

[14] Following the literature, variables are divided by baseline employed population, preventing estimates from capturing changes due to native mobility (Card and Peri, 2016[38]).

[15] The results are robust to using contemporaneous rather than one-year leaded changes in IP rights applications. For patents, estimated effects are smaller although still significant, consistent with the notion that innovation takes time to materialise.

further include annual and triannual changes, as well as the effect on cumulative patents, i.e., $\log \sum_{j=t+1}^{t+d} Y_{r,j}$ over three and five years (*d=3,5*).

$\Delta Migrants_{r,t}^{HE}$ is the change in the headcount of higher-educated employed migrants in year t over the employed population at baseline. X_{rt} contains baseline characteristics, including shares of workers in one-digit industries, the change in higher-educated native workers, and the logarithm of the population; α_t is a vector of time fixed-effects that absorbs the effect of any shock that might affect all Australian regions each year. β_1 is the coefficient of interest indicating the effect of a change in migration on the outcome variables, namely patent, trademark, and design rights applications. The analysis clusters the standard errors at the regional level to account for possible within-region correlation of random disturbances.

Endogeneity of the migrant share

Assessing the causal impact of migration on regional innovation poses an empirical challenge because of a possible "omitted variable bias", whereby other regional factors that are not accounted for in the analysis influence migrants' location choices and regional innovation. Migrants generally prefer to live in cities (OECD, 2022[15]), but innovation also clusters in cities to benefit from agglomeration economies (Carlino and Kerr, 2015[24]). Moreover, measurement error in recording those migrants who innovate - resulting from using changes in higher-educated or higher-skilled individuals as measures - could potentially weaken the relationship between migration and innovation.

To address both challenges, the empirical strategy combines the difference-in-differences method with instrumental variables. The analysis uses an instrument based on the past settlements of the migrant population (also known as the shift-share instrument), which is the most extensively used instrumental variable in the migration literature (Jaeger, Ruist and Stuhler, 2018[25]). The idea is that newly arrived migrants tend to settle in places where they can find migrants from their own country of origin. This way, the instrument predicts the inflow and settlement of migrants driven by networks rather than economic factors that may be driving innovative activity (Bartik, 1991[26]). The instrument uses information on the location of migrants who have arrived in earlier years to predict where the new migrants will settle, including information on pre-existing migrant enclaves and the number of newly arrived migrants at the national level by country of origin.

Building the shift-share instrument follows these steps (see Annex B for further details):

1. Split the migrant population of 1981 into 60 countries or regions of origin (See Table 5 (Annex B) for a detailed list).
2. Calculate the regional distribution (settlement pattern) of each origin using the 1981 Census.
3. For each country and region of origin, predict the presence of higher-educated migrants in each region and year using the total annual migrant population by origin between 2011 and 2018 (i.e., the shift) and the regional distribution of migrants by origin in 1981 (i.e., the share).
4. Sum up predicted settlements of higher-educated migrants across countries and regions of origin to obtain the predicted total number of higher-educated migrants living in a given region and year.
5. Apply the same process to obtain the predicted number of natives to avoid the instrument capturing the mobility response of natives due to the migrant inflow.
6. Use the predicted higher-educated migrant population to compute the predicted increase in higher-educated migrants over the total predicted number of migrants and natives at baseline in each region and year.

Validity of the instrumental variable strategy

The main identification assumption is that the instrumental variable affects regional innovation only through its impact on the increase of higher-educated migrants. This assumption is a combination of the instrumental variable being associated with the migrant increase (instrument relevance) and not being associated with other factors determining innovation (instrument exogeneity). In the context of shift-share instruments, recent literature has shown that instrument exogeneity can be satisfied from either exogeneity of the aggregate "shifts" (Borusyak, Hull and Jaravel, 2022[27]) or the baseline "shares" (Goldsmith-Pinkham, Sorkin and Swift, 2020[28]). This study relies on identification based on the exogeneity of the baseline shares, which means that the initial settlement of migrants across regions in 1981 is not correlated with persistent omitted factors that could also determine regional innovation.[16]

Three tests provide evidence that the instrument used in the analysis satisfies the exogeneity condition. First, initial shares of migrants in 1981 should be uncorrelated with regional characteristics in 1981 (Goldsmith-Pinkham, Sorkin and Swift, 2020[28]). Table 6 in Annex B shows that the shares of the top origin nationalities that drive most of the variation during the 2011-2018 period, i.e., India, China, Philippines, and Korea[17], as well as the instrument built upon these shares, are not associated with a set of regional characteristics reflecting regional labour markets characteristics and industry composition. Second, inflows of higher-educated migrants prior to the study period should not influence current changes in innovation (Jaeger, Ruist and Stuhler, 2018[25]). By including past inflows of higher-educated migrants in the main estimation equation, in Annex B shows that the results are unchanged. Consequently, the main estimates are driven by current migration inflows rather than by the long-term effects of past migration inflows. Finally, the effect of higher-educated migrants on innovation could be explained by a spurious correlation between predicted higher-educated migrants flows by the instrument and pre-existing trends in innovation. Annex B shows that the instrument has no significant correlation with previous trends in IP applications.

These tests provide evidence in favour of the exogeneity of the baseline shares in 1981. In turn, the instrument built upon this share is likely to be uncorrelated with other characteristics that could be driving the impact on changes in regional innovation during the 2011-2018 period. These tests are a necessary condition for the instrument to impact regional innovation only through its effect on actual inflows of higher-educated migrants.

[16] The number of high-educated migrants increased as of 2005 due to the reforms in the migration policies (Nguyen and Parsons, 2018[39]). While the dramatic increase also creates an exogenous "shift", this study relies on the exogeneity of the "shares".

[17] These are the four nationalities that contribute the most to the increase in migration during the 2011-2018 period.

5 Results

This section presents the results in two steps. Firstly, it examines the impact of migration on regional innovation. Secondly, it delves into the uneven effects due to different types of migrants and variations across regions.

Average effect of migration on innovation across regions

The effect of migration on innovation may not materialise immediately, as innovation often takes time to develop. To capture the varying horizons of innovation, the regressions consider 5-year, triannual, or annual differences for each outcome. Table 2 provides the results from estimating the main empirical model, using either patents, trademarks, or design rights as outcomes. While Panel A presents Ordinary Least Squares (OLS) estimates, Panel B provides instrumental variable estimates using Two Stage Least Squares (2SLS), which addresses potential endogeneity concerns.

Higher-educated migrants boost regional patenting activity with no impact on other intellectual property rights

Migration is positively associated with patenting applications but not with trademarks or design rights. The OLS and 2SLS estimates indicate a significant and positive impact of the inflow of higher-educated migrants on regional patent applications per worker.[18] On average, a one percentage point increase in the regional employment of higher-educated migrants relative to total employment leads to a 4.8% rise in regional patent applications in the medium run (five years). Similar effects are also observed in the shorter run (one or three years). [19] In contrast, the effect of migration on regional trademarks and design rights applications remains insignificant for both estimates and regardless of the timeframe. Moreover, increases in the number of higher-educated native individuals do not appear to influence any type of IP rights.[20]

The arrival of higher-educated migrants between 2011 and 2016 lifted regional patent applications by around 20%. Between 2011 and 2016, the average Australian region experienced a 4.5 percentage point increase in employment due to higher-educated migration, as shown in Table 1. A back-of-the-envelope

[18] The instrumental variable is a strong predictor of the main endogenous variable, conditional on the exogeneity of the instrument (see Annex B). This is confirmed by a consistently high first-stage F-statistic exceeding 10. The small and statistically insignificant differences between IV and OLS estimates suggest that the omitted variable bias and measurement error are minimal.

[19] The point estimates increase slightly when exploiting annual changes (from 4.8% to 6.6%). Yet, as the confidence intervals at any conventional significance level overlap, the estimates are not statistically different.

[20] The lack of significant effects for natives are possibly driven by two factors. Firstly, the proportion of higher-educated natives is already substantial with limited variation across regions and time. Secondly, higher-educated migrants tend to live in places where natives are also higher educated (OECD, 2022[15]). These factors make it challenging to empirically measure the distinct effects of higher-educated migrants and natives on innovation simultaneously.

calculation based on the five-year impact of migration on innovation (Column 1 in Panel B, Table 2) suggests that higher-educated migration during this period contributed to an approximate increase in patenting activity of 20%. This implies that if the baseline number of patent applications per 10 000 workers is 33, the presence of migrants increased this value to 39.

The increase in patenting but not in trademarks or design rights aligns with previous findings. For Australia, previous research finds that an increase in Temporary Graduate visa holders between 2007 and 2014 led to more patent applications per worker, while trademarks or design rights were not significantly impacted (Crown, Faggian and Corcoran, 2020[18]). Similarly, in the case of the US, previous research finds an impact of migrants on innovation and patenting in high-tech sectors but no effect on trademarks (Brown et al., 2020[29]). The inflow of higher-educated migrants between 2011 and 2016 in Australia predominantly results from the arrival of migrants in STEM industries that are more likely to patent. This potentially explains the significant regional effect on patterns and the absence of any effects on other innovation measures.

The positive effect of migration on regional innovation holds against different robustness tests. Annex C presents various robustness checks assessing the sensitivity of results and validating the empirical strategy. Overall, the magnitude and statistical significance of the coefficients are consistent with the baseline results.

Table 2. The average effect of migration on regional innovation

Estimated average effect of migration inflows on regional patents, trademarks, and design rights at the regional level

	Patents			Trademarks			Design rights		
	5 years	3 years	1 year	5 years	3 years	1 year	5 years	3 years	1 year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: OLS									
Δ HE Migrants	3.572**	3.838***	4.917*	-1.066	-0.397	1.931	-0.766	-1.568	9.122
	(1.242)	(1.825)	(2.523)	(0.651)	(0.849)	(1.136)	(3.743)	(3.969)	(5.344)
Δ HE Natives	0.701	-0.941	-3.529	-1.826	-1.454	-1.494	-2.374	12.036*	-2.287
	(2.375)	(2.839)	(3.134)	(1.563)	(1.912)	(1.831)	(7.178)	(6.954)	(8.510)
R^2	0.211	0.082	0.052	0.054	0.077	0.189	0.066	0.204	0.142
Panel B: 2SLS									
Δ HE Migrants	4.754**	6.351***	6.564**	-1.459	-1.919	-0.245	0.754	8.725	14.621*
	(1.726)	(2.832)	(3.264)	(1.225)	(2.671)	(2.403)	(5.727)	(7.235)	(8.241)
Δ HE Natives	0.576	-1.542	-4.603	-1.785	-1.091	-0.075	-2.534	9.575	-5.873
	(2.340)	(2.909)	(3.744)	(1.540)	(2.255)	(2.487)	(6.878)	(6.471)	(8.935)
N	88	176	616	88	176	616	88	176	616
F-stat	27.1	25.3	43.1	27.1	25.3	43.1	27.1	25.3	43.1

Note: The table presents results based on estimating Equation 1 using as outcomes either Patents (Columns 1-3), Trademarks (Columns 4-6) and Design rights (Columns 7-9) applications. Panel A uses OLS as an estimator, while Panel B estimates IV results using a 2SLS estimator. The dependent variable is expressed as log changes in IP applications per worker, using either 5, 3, or 1-year differences. The independent variable is the increase in employment due to higher-educated migrants, where the measure of high education is tertiary education (at least with a college degree). The columns present different time intervals for both the dependent and independent variables. Columns 1, 4 and 7 represent five-year changes (2011-2016), Columns 2, 5 and 8 represent three-year changes (2011-2014 and 2014-2017), and Columns 3, 6, and 9 represent annual changes. No intervals overlap. All specifications control for baseline shares of higher-educated population, log of population and industry shares. All specifications are weighted by the number of employed natives in the region at the baseline year. Time fixed-effects are applied to account for time-varying shocks affecting the entire country, except for five-year changes due to collinearity. Standard errors, in parentheses, are clustered at the regional level in all specifications. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively. The analysis considers 88 regions, yielding 616 observations over five years.
Source: OECD calculations based on MADIP and IPGOD (accessed July 2023).

Uneven effects across workers and places

The impact of migration on innovation can differ based on the characteristics of migrants and the conditions of the regional economy. Firstly, the analysis delves into the impact of distinct groups of migrants, categorised by skill levels and educational attainments. Secondly, the analysis tests whether the effects vary across different Australian regions, depending on their existing levels of innovation, population density, and income.

Migrants of all education groups contribute to patenting across Australian regions

Estimating the effect of different migrant subgroups based on their education and occupation allows a broader understanding of the impact of migration on innovation. Table 3 compares the baseline estimates (Panel A) to alternative definitions of the migration inflow. Panel B focuses specifically on changes in higher-skilled migrants, employing occupation-based skill assignments instead of education levels. Panel C restricts the migrant inflow to scientists using a subset of occupations more likely to be involved in patenting activities (for detailed information, see Annex D). Finally, Panel D includes all migrants, irrespective of their educational background or occupation. Consistent with prior studies (Hunt and Gauthier-Loiselle, 2010[6]), all specifications use the instrument based on 1981 settlement patterns.

In all panels, the findings consistently demonstrate positive effects on patent applications with similar magnitudes across various specifications and timeframes. This suggests that the definition of higher-educated workers does not substantially influence the results, and migrants of all skill and education groups contribute to the patenting activity. Moreover, the effect is most pronounced when focusing only on migrants employed as scientists. Similar to the baseline findings, there are no discernible effects on trademarks or design rights. These patterns further affirm the reliability of the migration measure and the empirical approach. Finally, Table 10 in Annex E shows that even when excluding scientists, the effect of all migrants or higher-educated migrants on patenting activity is still significant, thus showing that migrant scientists are not driving the results.

Table 3. Uneven effects of migrants with different skills and education backgrounds

	Patents			Trademarks			Design rights		
	5 years	3 years	1 year	5 years	3 years	1 year	5 years	3 years	1 year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Baseline - Higher-educated (at least college)									
Δ HE Migrants	4.754***	6.351**	6.564**	-1.459	-1.919	-0.245	0.754	8.725	14.621*
	(1.726)	(2.832)	(3.264)	(1.225)	(2.671)	(2.403)	(5.727)	(7.235)	(8.241)
F-stat	27.1	25.3	43.1	27.1	25.3	43.1	27.1	25.3	43.1
Panel B: Higher-skilled occupations									
Δ HS Migrants	4.916***	6.407**	5.367*	-1.532	-2.539	-0.251	1.912	6.994	12.402
	(1.862)	(2.733)	2.878	(1.216)	(2.395)	(2.161)	(6.316)	(7.058)	(9.059)
F-stat	38.7	33.3	67.1	38.7	33.3	67.1	38.7	33.3	67.1
Panel C: Scientists									
Δ Scientist Migrants	12.050***	15.375**	17.197**	-3.828	-5.862	-2.175	2.868	14.111	46.042**
	(4.239)	(6.428)	(7.640)	(2.790)	(5.446)	(5.897)	(14.882)	(16.276)	(21.865)
F-stat	31.9	26.6	37.8	31.9	26.6	37.8	31.9	26.6	37.8
Panel D: All migrants									
Δ Migrants	1.550***	2.135**	2.454**	-0.407	-0.781	0.032	0.769	2.167	4.993
	(0.597)	(0.891)	(1.148)	(0.390)	(0.829)	(0.831)	(2.029)	(2.472)	(3.342)
F-stat	33.1	30.4	77.1	33.1	30.4	77.1	33.1	30.4	77.1

Note: The table presents results based on estimating Equation 1 using as outcomes either Patents (Columns 1-3), Trademarks (Columns 4-6) and Design rights (Columns 7-9) applications. Panel A uses the same definition as baseline, higher-educated individuals (with at least college education). Panel B uses higher-skilled occupations, while Panel C uses a subset of applications that are more likely to submit patent applications, as specified in Annex D. Panel D uses all migrants. The dependent variable is expressed as log changes in IP applications per worker, using either 5, 3, or 1-year differences. The independent variable is the increase in employment by type of migrant. The columns present different time intervals for both the dependent and independent variables. Columns 1, 4 and 7 represent five-year changes (2011-2016), Columns 2, 5 and 8 represent three-year changes (2011-2014 and 2014-2017), and Columns 3, 6, and 9 represent annual changes. No intervals overlap. All specifications control for baseline shares of higher-educated population, log of population and industry shares. All specifications are weighted by the number of employed natives in the region at the baseline year. Time fixed-effects are applied to account for time-varying shocks affecting the entire country, except for five-year changes due to collinearity. Standard errors, in parentheses, are clustered at the regional level in all specifications. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.
Source: OECD calculations based on MADIP and IPGOD (accessed July 2023).

The estimates indicate that the impact of migrant scientists on regional innovation is less than half of what was estimated for the US (Hunt and Gauthier-Loiselle, 2010[6]). The study for 1950-2000 shows that a one percentage point increase in the share of scientist migrants led to a substantial 52.4% rise in patents over a 10-year period.[21] However, the analysis might underestimate the positive impact as migrants in Australia could be patenting abroad, particularly in the US. This area remains a potential focus for future research.

The impact of migration on innovation depends on existing regional innovation activity but not on density or income

This subsection explores uneven effects across regions by focusing on several regional characteristics in the baseline year. The estimations in Table 4 expand the main estimation equation by interacting the variable of interest (i.e., the regional change in higher-educated migrants) with an indicator equal to one if the regions have below-the-median values for the considered regional characteristic in the initial year. The interaction term is additionally instrumented by the analogous interaction on the instrument, following previous literature (Özgüzel and Edo, 2023[30]).

The analysis focuses on three different regional characteristics to explore the uneven effects of migration on regional innovation. First, density, defined as the number of workers over the regional built-up area, can play a role in innovation as proximity among innovators can facilitate collaboration and foster innovation.[22] Secondly, income levels, measured using average regional labour income, may influence the degree of innovation in different regions, as places with higher income levels may disproportionately attract migrants who are better innovators. Lastly, regions already highly innovative may have innovation hubs that potentially use the migrants' skills more efficiently and, hence, generate larger effects. Alternatively, it is also possible that such places benefit relatively less from new migrant inflows as increasing innovation becomes more difficult at higher levels of innovation compared to less innovative places.

The impact of migrants on regional innovation is independent of regional income and density yet varies depending on the pre-existing innovation levels. Table 4 outlines the results. The interaction parameter for regional density and income (Panels A and B) lacks significance, suggesting no differences exist across regions above and below the median value. In contrast, in Panel C, regions below the median in existing innovation levels show significant interaction coefficients for patents (5- and 1-year changes) and design

[21] The study finds similar estimates whether they look at 10-year or 50-year windows, which suggests that the estimated effects are stable across time. This makes it possible to compare these magnitudes estimated for the US with those estimated in this study for Australia.

[22] Retrieved from the Global Human Settlement Layer (GHSL): Global Human Settlement - GHSL Homepage - European Commission (europa.eu). The baseline year is 2020.

rights (5 and 1-year changes), indicating that the effect is stronger in less innovative regions. This contrasts with the belief that innovation mainly occurs in already innovative places (Crescenzi, Dyevre and Neffke, 2022[31]). However, even though less innovative regions might catch up in patenting activity, the overall innovation gap across regions does not necessarily decrease.

Table 4. Uneven effects across regions with different characteristics

	Patents			Trademarks			Design rights		
	5 years	3 years	1 year	5 years	3 years	1 year	5 years	3 years	1 year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Density (built-up)									
∆ HE Migrants	4.773**	6.215**	6.110*	-1.318	-3.113	-0.157	2.466	9.098	13.857*
	(1.875)	(2.937)	(3.431)	(0.917)	(2.061)	(2.277)	(4.657)	(8.167)	(8.359)
∆ HE Migrants x Below Median Density	-0.061	0.615	4.501	-0.657	5.043	-0.744	-7.877	-1.697	6.09
	(2.572)	(3.582)	(4.773)	(2.608)	(4.113)	(4.174)	(9.534)	(10.293)	(9.998)
Panel B: Income									
∆ HE Migrants	4.335**	5.887**	5.784*	-1.56	-1.648	-0.333	-0.502	7.987	12.387
	(1.802)	(2.842)	(3.454)	(1.214)	(2.699)	(2.371)	(5.703)	(6.681)	(8.325)
∆ HE Migrants x Below Median Income	1.56	1.491	3.426	0.364	-0.845	0.387	4.631	2.229	8.922*
	(1.423)	(1.615)	(2.442)	(0.714)	(1.144)	(1.347)	(3.91)	(5.236)	(4.779)
Panel C: Innovation									
∆ HE Migrants	5.731***	7.790**	8.000**	-1.408	-1.649	-0.2	1.61	9.502	16.357**
	(1.804)	(3.364)	(3.359)	(1.379)	(2.839)	(2.638)	(5.437)	(7.237)	(8.188)
∆ HE Migrants x Below Median Innovation	2.764*	2.821	4.932**	0.517	1.579	0.254	13.368***	6.917	16.461***
	(1.43)	(1.923)	(2.143)	(1.486)	(2.355)	(2.631)	(4.766)	(5.554)	(4.608)

Note: The table presents results based on estimating an extended version of Equation 1 using as outcomes either Patents (Columns 1-3), Trademarks (Columns 4-6) and Design rights (Columns 7-9) applications. The specification includes an interaction term with an indicator of whether each region has below the median innovation, density or income. Panel A provides the results for different levels of innovation. Panel B provides the results for different levels of built-up density, and Panel C investigates the interaction with differing levels of income. The independent variables of interest are the increase in employment due to higher-educated migrants, where the measure of high education is tertiary educated (at least with a college degree), and its interaction with an indicator variable that takes value one if the region has below the median innovation, density or income, depending on the panel. The columns present different time intervals for both the dependent and independent variables. Columns 1, 4 and 7 represent five-year changes (2011-2016), Columns 2, 5 and 8 represent three-year changes (2011-2014 and 2014-2017), and Columns 3, 6, and 9 represent annual changes. No intervals overlap. All specifications control for baseline shares of higher-educated population, log of population and industry shares. All specifications are weighted by the number of employed natives in the region at the baseline year. Time fixed-effects are applied to account for time-varying shocks affecting the entire country, except for five-year changes due to collinearity. Standard errors, in parentheses, are clustered at the regional level in all specifications. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.
Source: OECD calculations based on MADIP and IPGOD (accessed July 2023).

6 Concluding remarks

Migrants play a crucial role in regional development by contributing significantly to income, international trade, and labour markets. This paper provides the first causal evidence of migrants' impact on regional innovation in the country, using comprehensive administrative data covering all Australian residents and IP rights applications. This study expands the understanding of migrants' influence on innovation, focusing on the case of an OECD country where migrants are higher-educated. Additionally, it presents new insights by examining how migrants affect regional innovation while considering regional characteristics.

The paper uncovers four main findings. First, it shows that a one percentage point increase in regional employment due to the arrival of higher-educated migrants leads to a 4.8% increase in patent applications per worker in the medium run (five years). Second, while migrants of all skill and education levels positively impact patenting, the effect is most pronounced for migrants in scientific occupations. Third, regions with lower pre-existing levels of patenting benefit relatively more from higher shares of migrants than those with higher patenting levels. Other regional factors like income or population density do not influence the effect. Finally, the analysis finds no effect on trademarks or design rights applications.

The analysis presented in this paper is an important step towards understanding the impact of migration on innovation, but more research is needed for a deeper understanding of this relationship and underlying mechanism. By contributing to innovation in patenting, migrants contribute to the development and growth of businesses, impacting overall welfare. Investigating how innovation driven by migrants contributes to productivity growth would be crucial for understanding their role in economic development. Moreover, investigating the underlying mechanism of the impact of migrants on innovation would be an essential complement to this research. Firm-level analysis using matched employer-employee data would allow to unfold and understand how migrants contribute to the increased patenting activity. Further, it would be valuable to investigate why the positive effects of migration on patents do not translate to trademarks or design rights. This divergence is possibly due to a lack of migrant entrepreneurship in sectors that use trademarks or design rights more intensively (Brown et al., 2020[29]). Lastly, examining the impact of employer-sponsored visas on firm-level innovation would provide valuable insights for policymakers.

References

Bartik, T. (1991), “Who Benefits from State and Local Economic Development Policies”, *W.E. Upjohn Institute*. [26]

Bernstein, S. et al. (2022), “The contribution of high-skilled immigrants to innovation in the United States”, *Working Paper*. [7]

Blit, J., M. Skuterud and J. Zhang (2020), “Can skilled immigration raise innovation? Evidence from Canadian Cities”, *Journal of Economic Geography*, Vol. 20/4, pp. 879-901. [23]

Borusyak, K., P. Hull and X. Jaravel (2022), “Quasi-Experimental Shift-Share Research Designs”, *The Review of Economic Studies*, Vol. 89, pp. 181-213, https://doi.org/10.1093/restud/rdab030. [27]

Bosetti, V., C. Cattaneo and E. Verdolini (2015), “Migration of skilled workers and innovation: A European Perspective”, *Journal of International Economics*, Vol. 96/2, pp. 311-322, https://doi.org/10.1016/J.JINTECO.2015.04.002. [13]

Bratti, M. and C. Conti (2018), “The effect of immigration on innovation in Italy”, *Regional Studies*, Vol. 52/7, pp. 934-947. [14]

Butts, K. (2023), “Difference-in-Difference Estimation with Spatial Spillovers”, *Working Paper*. [33]

Card, D. and G. Peri (2016), “Immigration Economics by George J. Borjas: A Review Essay”, *Journal of Economic Literature*, Vol. 54/4, pp. 1333-49, https://doi.org/10.1257/jel.20151248. [38]

Carlino, G. and W. Kerr (2015), *Agglomeration and Innovation*, Elsevier B.V., https://doi.org/10.1016/B978-0-444-59517-1.00006-4. [24]

Clarke, A., A. Ferrer and M. Skuterud (2019), “A Comparative Analysis of the Labor Market Performance of University-Educated Immigrants in Australia, Canada, and the United States: Does Policy Matter?”, *Journal of Labour Economics*, Vol. 37/S2, pp. S443-S490. [17]

Crescenzi, R., A. Dyevre and F. Neffke (2022), “Innovation Catalysts: How Multinationals Reshape the Global Geography of Innovation”, *Economic Geography*, Vol. 98/3, pp. 199-227, https://doi.org/10.1080/00130095.2022.2026766. [31]

Crown, D., A. Faggian and J. Corcoran (2020), “Foreign-Born graduates and innovation: Evidence from an Australian skilled visa program”, *Research Policy*, Vol. 49/9, pp. 1-17. [36]

Crown, D., A. Faggian and J. Corcoran (2020), “Foreign-Born graduates and innovation: Evidence from an Australian skilled visa program”, *Research Policy*, Vol. 49/9, p. 103945, https://doi.org/10.1016/J.RESPOL.2020.103945. [18]

Dotzel, K. and T. Wojan (2022), “An Occupational Approach to Analyzing Regional Invention”, *National Center for Science and Engineering Statistics (NCSES)*, https://ncses.nsf.gov/pubs/ncses22202/assets/ncses22202.pdf. [34]

Dustmann, C., U. Schonberg and J. Stuhler (2016), “The Impact of Immigration: Why Do Studies Reach Such Different Results?”, *Journal of Economic Perspectives*, Vol. 30/4, pp. 31-56. [37]

Gagliardi, L. (2015), “Does skilled migration foster innovative performance? Evidence from British local areas”, *Papers in Regional Science*, Vol. 94/4, pp. 773-794. [11]

Ganguli, I., S. Kahn and M. MacGarvie (eds.) (2020), *Immigrant Entrepreneurs and Innovation in the US High-Tech Sector*, University of Chicago Press. [29]

Goldsmith-Pinkham, P., I. Sorkin and H. Swift (2020), “Bartik Instruments: What, When, Why, and How”, *American Economic Review*, Vol. 110/8, pp. 2586-2624, https://doi.org/10.1257/aer.20181047. [28]

Hunt, J. (2011), “Which Immigrants are Most Innovative and Entrepreneurial? Distinctions by Entry Visa”, *Journal of Labor Economics*, Vol. 29/3, pp. 417-457. [5]

Hunt, J. and M. Gauthier-Loiselle (2010), “How much does immigration boost innovation”, *American Economic Journal: Macroeconomics*, Vol. 2/2, pp. 31-56, https://doi.org/10.1257/mac.2.2.31. [40]

Hunt, J. and M. Gauthier-Loiselle (2010), “How Much Does Immigration Boost Innovation?”, *American Economic Journal: Macroeconomics*, Vol. 2/2, pp. 31-56. [6]

Imbert, C. et al. (2022), “Migrants and Firms: Evidence from China”, *American Economic Review*, Vol. 112/6, pp. 1885-1914. [16]

Jaeger, D., J. Ruist and J. Stuhler (2018), *Shift-share instruments and the impact of immigration*, National Bureau of Economic Research. [25]

Jensen, P. (2014), “Understanding the impact of migration on innovation”, *The Australian Economic Review*, Vol. 47/2, pp. 240-250. [19]

Jensen, P. and E. Webster (2009), “Another look at the relationship between innivation proxies”, *Australian Economic Papers*, Vol. 48, pp. 252-269. [21]

Kerr, S. et al. (2016), “Global Talent Flows”, *Journal of Economic Perspectives*, Vol. 30/4, pp. 83-106. [10]

Mayda, A., G. Orefice and G. Santoni (2022), “Skilled Immigration, Task Allocation and the Innovation of Firms”, *IZA Discussion Paper Series*, Vol. 15693. [12]

Nguyen, T. and C. Parsons (2018), “The Labour Market Impact of a High Skilled Migration Wave:”, *Working Paper*. [39]

OECD (2023), *International Migration Outlook 2023*, OECD Publishing, Paris, https://doi.org/10.1787/b0f40584-en. [22]

OECD (2023), "Migration and regional productivity: Evidence from individual wages in Australia", *OECD Regional Development Papers*, No. 60, OECD Publishing, Paris, https://doi.org/10.1787/7bc64c78-en. [2]

OECD (2023), "Regional productivity, local labour markets, and migration in Australia", *OECD Regional Development Papers*, No. 39, OECD Publishing, Paris, https://doi.org/10.1787/3cc8f669-en. [1]

OECD (2023), "The impact of migration on regional labour markets in Australia", *OECD Regional Development Papers*, No. 64, OECD Publishing, Paris, https://doi.org/10.1787/d72110b5-en. [3]

OECD (2022), *International Migration Outlook 2022*, OECD Publishing, Paris, https://doi.org/10.1787/30fe16d2-en. [4]

OECD (2022), *The Contribution of Migration to Regional Development*, OECD Regional Development Studies, OECD Publishing, Paris, https://doi.org/10.1787/57046df4-en. [15]

OECD (2022), "Unlocking Rural Innovation", *OECD Rural Studies*, Vol. OECD Publishing, Paris, https://doi.org/10.1787/9044a961-en. [35]

OECD/Eurostat (2018), *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition*, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris/Eurostat, Luxembourg, https://doi.org/10.1787/9789264304604-en. [41]

Ozgen, C. (2021), "The economics of diversity: Innovation, productivity, and the labour market", *Journal of Economic Surveys*, Vol. 35/4, pp. 1168-1216. [8]

Özgüzel, C. and A. Edo (2023), "Immigration and employment dynamics in European regions", *OECD Regional Development Papers*, No. 52, OECD Publishing, Paris, https://doi.org/10.1787/a20155a2-en. [30]

Perez-Silva, R., M. Partridge and W. Foster (2019), "Are foreign-born researchers more innovative? Self-selection and the production of knowledge among PhD recipients in the USA", *Journal of Geographical Systems*, Vol. 21/4, pp. 557-594. [9]

Proctor, M. (2023), *https://www.matthewproctor.com/australian_postcodes*. [32]

Sarwar, A. (2022), "Australia's productivity challenge is an innovation challenge", *Australian Industry Group (AI group)*. [20]

Annex A. Data sources

Multi-Agency Data Integration Project (MADIP)

The Multi-Agency Data Integration Project (MADIP) dataset by the Australian Bureau of Statistics (ABS) is an individual-level panel dataset that provides longitudinal information for more than 27 million individual records between 2011 and 2020. MADIP combines administrative information from different departments, such as the Australian Taxation Office (ATO), the Department of Education, the Department of Health and Aged Care, the Department of Social Services, Services Australia, and the Department of Home Affairs.The availability of the dataset is subject to the agreement of the data custodians of the individual agencies and depends on the individual research question. In addition to administrative data, the MADIP includes one of the quinquennial Australian Census of Housing and Population. Besides the MADIP core data, this analysis relies on tax data by ATO, migration data by the Department of Home Affairs, and the Census 2018. The following subsections describe the individual components of the dataset.

MADIP core data

The MADIP core dataset is at the centre of every analysis using MADIP data. It contains demographic information like date of birth, gender, and date of death as well as location information on all residents in Australia. Moreover, the dataset includes a *spine* ID integral to merging the individual datasets from different agencies. By default, the dataset covers every Australian resident recorded in either *Social Security and Related Information*, *Personal Income Tax* data, or *Medicare Benefits Schedule* data between 2006 and 2020, resulting in a total of 27.1 million individual records. However, not every recorded person is listed in every individual dataset. For instance, income tax data is not available if the person has never reported taxes (e.g., children).

The geographical information is available at different granularity levels, including SA4, SA3, and SA2. Given the overwhelming coverage of the Australian population, the data is expected to be representative at every geographical level. Location information is distinguished by residential and mail address. For migrants, the business address is also reported. In the analysis, the individual location information is based on the residential address or mail address, depending on data availability.

MADIP is expected to cover the vast majority of Australian citizens and residents due to the combination of medicare, social benefits, and income tax records. According to the ABS, the following groups are potentially underrepresented: i) recently arrived migrants without Medicare, ii) non-earning partners and family members of working visa holders, iii) non-earning foreign students, iv) military personnel, v) prisoners, vi) recently born individuals, not yet included in the Medicare Benefits Schedule.

Census of Population and Housing 2016

The Australian Census of Population and Housing is conducted every five years and includes, among others, information on educational attainment, employment and work, family, and personal characteristics. This paper uses data from the Census wave of 2016, the latest available for research, linked to the MADIP universe. Due to Australian data confidentiality rules, only one Census wave at a time can be used in the MADIP environment. The Census data refers to the data collected on the 9th of August, 2016.

In Australia, participating in the census is mandatory for Australian residents, with very few exceptions. The ABS linked 20.7 million records of the Census 2016 to the MADIP data, which corresponds to 88% of all collected Census records in 2016. According to the ABS, the following groups are not within the scope of the Census: i) Australians overseas, ii) residents for less than six months, iii) visitors, iv) diplomatic personnel and their families. The paper retrieves information on age, occupation, industry, and country of birth from the Census.

Australian Taxation Office

The Australian Taxation Office provides administrative information on all employed individuals in Australia based on official tax returns. The dataset covers around 16.7 million individual records, including everyone with a tax return in Australia in at least one year from 2010/2011 – 2017/2018. The Australian financial tax year ranges from July until June of the following year. However, in order to combine the data with other datasets, the tax records are assumed to follow the calendar year (January-December rather than July-June).[23] Data spans from wages, total income, and insurance payments to job sector information. The variables of interest to this analysis are age, individual wage/salary, the main salary or wage occupation code, and industry. Employed individuals with an income below the threshold imposed by the ATO and, hence, without a tax record are not considered in the data. This also includes most migrants on a working holiday maker (WHM) visa.

Department of Home Affairs

The *Department of Home Affairs* provides administrative data on the native and migrant populations. The data includes every individual (native- or foreign-born) who crossed the border of Australia between 1990 and 2020. The dataset is used to retrieve information on the country of birth, date (month and year) of birth, and gender. Visa information is not available for all migrants. Moreover, due to changes in the visa status after arriving in Australia, the visa information might not be reliable for all migrants.

IP Government Open Data

The IP Government Open Data (IPGOD) is a data set provided by IP Australia, containing information on over 100 years of information on IP rights applications. The data are updated yearly, and this project uses the most recent available data, IPGOD 2022. The data set allows researchers to investigate the state of IP rights since applications were first filed to IP Australia. It includes rich microdata on applicants[24], key dates and events, the classification of IP rights and the history of IP transfers and exchanges.

The data cover four types of IP rights administered by IP Australia: trademarks, patents, design rights and plant breeder's rights. The latter are not included in this paper due to their limited scope.

Importantly, the data include information on applicant's location, namely postcode. These are used to assign applications to different regions, using an updated cross-walk of postcodes to SA4 regions (Proctor, 2023[32]).

[23] This means that the tax return for the financial year 2011/12 is treated as the tax return for the year 2012. The age retrieved from the ATO, is adjusted accordingly.

[24] For some applicants, the Australian Business number (ABN) is provided.

Historic Census

As discussed in Section 5, the identification strategy of the paper requires the use of a historical instrument based on the settlement patterns of migrants in the past. The information on the past settlement patterns is obtained from census data from 1981, 1986, 1991, and 2001. The historical data has been adjusted to 2016 borders by the ABS and provides data on the total employed population of the respective year disaggregated by country of birth, the industry of employment, and highest post-school qualification. The country of birth consists of 60 national groupings. Table 5 lists the national groupings and their share of the total migrant population in 1981.

Table 5. Migrant decomposition in terms of country of origin

Share of employed migrants among the employed migrant population, 1981

Country groupings	Share of total foreign-born employed population
Albania, Bulgaria & Romania	0.32%
Argentina & Uruguay	0.51%
Austria	0.87%
Bangladesh	0.03%
Belgium	0.15%
Brazil	0.05%
Cambodia, Laos & Myanmar	0.42%
Canada	0.46%
Chile	0.40%
China	0.99%
Colombia, Ecuador & Peru	0.13%
Cyprus	0.82%
Denmark, Finland, Norway & Sweden	0.87%
Egypt	1.20%
Fiji	0.27%
Former Czechoslovakia	0.63%
Former USSR	1.67%
Former Yugoslavia	5.63%
France	0.42%
Germany	4.52%
Greece	6.00%
Hong Kong & Macau	0.41%
Hungary	1.11%
India	1.44%
Indonesia & Timor-Leste	0.54%
Iran	0.11%
Iraq	0.11%
Ireland	1.62%
Israel	0.21%
Italy	10.97%
Japan	0.27%
Kenya, Malawi, Zimbabwe, Tanzania, Uganda & Zambia	0.27%
Korea	0.13%
Lebanon	1.48%
Malaysia & Brunei	0.76%
Malta	2.32%

Mauritius	0.33%
Mexico	0.01%
Netherlands	3.92%
New Zealand	4.63%
Other Africa	0.37%
Other Middle East	0.05%
Pakistan	0.08%
Papua New Guinea	0.25%
Philippines	0.51%
Poland	2.18%
Portugal	0.40%
Singapore	0.30%
South Africa & Namibia	0.73%
Spain	0.54%
Sri Lanka	0.56%
Switzerland	0.26%
Syria	0.11%
Taiwan	0.03%
Thailand	0.09%
Türkiye	0.66%
United Kingdom	33.58%
United States of America	0.96%
Vietnam	0.82%
All other countries	0.54%

Note: The table presents the share of the foreign-born employed population in 1981. The countries of origin are aggregated into 60 national groupings. Grouping was conducted by the ABS and refers to the international borders of 1981.
Source: Australian Census of Population 1981 from the Australian Bureau of Statistics (ABS).

Annex B. Construction and validity of the instrumental variable

This section details the construction of the instrumental variable and then discusses the validity tests to provide evidence on the main identifying assumption, i.e., that the instrumental variable affects regional innovation only through its impact on the increase of higher-educated migrants. The analysis draws from data from the 2001 and 2006 Australian Population Census.

Construction of the instrumental variable

The construction of the instrument follows several steps. First, migrants are grouped into 60 origin groups, as reported in Table 5. Second, the settlement patterns of the pre-existing migrant population by country of origin across Australian regions is calculated using data from the 1981 Census.[25] The shares are calculated as follows:

$$Share_{n,r}^{1981} = \frac{Migrants_{n,r}^{1981}}{\sum Migrants_{n,r}^{1981}}$$

The numerator $Migrants_{n,r}^{1981}$ is the number of employed migrants in 1981 in each of the 60 national groupings *n* in region *r*. The denominator $\sum Migrants_{n,r}^{1981}$ is the total migrant population from national grouping *n* in 1981 across Australia. These shares are used to estimate the total number of higher-educated migrants in a given region *r* in year *t*, $Migrants_{rt}$. This is done by predicting the total number of higher-educated migrants from origin country *n* in year *t* across Australia, $Migrants_{nt}$, using the shares of individuals by nationality in 1981:

$$\widehat{Migrants}_{rt}^{HE} = \sum_{n=1}^{60} Share_{n,r}^{1981} * Migrants_{nt}^{HE}$$

By assigning migrants from each country across different regions, an estimate of the expected number of migrants in each region based on settlement patterns from 1981 is obtained. Baseline shares of total employed people rather than higher-educated employed people are used in order to emphasize the role of non-economic factors and ethnic networks in determining location choices.

Similar to the migrant population, the settlement decision of natives may not be random as natives might also be attracted to places that are more dynamic and innovative. Furthermore, natives potentially react to the arrival of migrants by moving out of regions where migrants disproportionately locate (OECD, 2023[3]). Therefore, native population numbers, used in the denominator of the higher-educated migrant share, may

[25] This is the earliest year in which local data can be matched to the borders used in the analysis. Additionally, a forty years lag is useful to claim that shares are unrelated to contemporary economic factors.

also suffer from endogeneity problems. To address this concern, the current regional native population $Natives_{rt}$ is also predicted based on the settlement patterns in 1981:

$$\widehat{Natives}_{r,t} = \frac{Natives_r^{1981}}{\sum Natives_r^{1981}} * Natives_{rt}$$

Finally, the predicted numbers of migrants and natives are used to predict the inflow of migrants, which is used to instrument variables for the migrant inflow relative to the total population:

$$\Delta\widehat{Migrants}_{r,t}^{HE} = \frac{\widehat{Migrants}_{r,t}^{HE} - \widehat{Migrants}_{r,t-1}^{HE}}{\widehat{Migrants}_{r,t-1} + \widehat{Natives}_{r,t-1}}$$

Validity of the instrumental variable

Validity test 1: Past settlement patterns are not correlated with past regional characteristics

To provide evidence on the exogeneity of the shares, previous research suggests investigating whether initial shares of migrants in 1981 are correlated with regional characteristics, which can, in turn, be correlated with current innovation (Goldsmith-Pinkham, Sorkin and Swift, 2020[28]). In this case, the initial shares of the top-origin nationalities that drive most of the variation during the 2011-2018 period, i.e., India, China, Philippines, and Korea[26], should not be associated with regional characteristics in 1981.

Most of the regional variables are not associated with origin country shares in 1981. Table 6 provides the results of regressions of top origin-specific shares on a set of regional characteristics reflecting regional labour market characteristics and industry composition in 1981. These include shares of higher-educated workers, the distribution of workers across sectors, and the logarithm of wages and employment. Columns 1 to 4 show that out of 20 coefficients, only 2 are statistically significant. Additionally, Column 5 shows that when these four top nationalities are grouped together, none of the regional characteristics is correlated with the aggregate origin shares. In consequence, the instrument is not correlated with regional characteristics in 1981. Columns 6 to 8 assess the association of regional variables in 1981 with the predicted increase in the number of higher-educated migrants, i.e., the instrument. Neither the share of higher-educated individuals nor the sectoral shares or wage or employment levels are correlated with the instrument. Taken together, these results provide further evidence that the instrument is affecting current innovation only through its effect on migration flows.

Table 6. Explanatory variables in 1981

	India	China	Philippines	Korea	Top 4	Δ Predicted HE Migrant		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share higher-educated	0.038	-0.116*	-0.127**	-0.120	-0.205	0.057	0.060	-0.102
	(0.067)	(0.063)	(0.056)	(0.144)	(0.135)	(0.062)	(0.075)	(0.158)
Share primary sector	0.023	0.010	0.027	0.047	0.059		-0.015	-0.072
	(0.028)	(0.026)	(0.023)	(0.060)	(0.056)		(0.047)	(0.056)

[26] These are the four nationalities that contribute the most to the increase in migration during the 2011-2018 period.

Share secondary sector	0.042	-0.021	0.028	0.043	0.050		0.032	0.040
	(0.035)	(0.033)	(0.029)	(0.075)	(0.071)		(0.060)	(0.060)
Wages (logarithm)	0.100	-0.084	-0.066	-0.015	-0.050			0.020
	(0.064)	(0.061)	(0.054)	(0.139)	(0.130)			(0.087)
Employment (logarithm)	-0.025	0.053	0.048	0.009	0.076			0.189
	(0.037)	(0.035)	(0.032)	(0.081)	(0.076)			(0.135)
R^2	0.511	0.686	0.721	0.463	0.753	0.628	0.633	0.666

Note: The table presents correlations of origin shares in 1981 for the main origin countries (India, China, Philippines, Korea, and their sum) and the instrument, with respect to regional characteristics in 1981. Each column represents a dependent variable. The independent variables are regional characteristics, which include the share of higher-educated individuals, the share of workers in the primary or secondary sector, with the tertiary sector omitted to avoid multicollinearity and the logarithms of wage and employment levels in 1981. All specification control for shares of higher-educated population, log of population and industry shares in 2011. All specifications are weighted by the number of employed natives in the region at the baseline year. Robust standard errors are provided in parentheses. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.
Source: OECD calculations based on MADIP and Australian Census of Population 1981 (accessed July 2023).

Validity test 2: The effect is not driven by adjustment to previous inflows of higher-educated migrants

The arrival of migrants may continue affecting regional outcomes beyond their immediate effects in the short run. If inflows of higher-educated migrants prior to the study period have long-lasting effects, these can be conflated with the effects of contemporaneous increases in higher-educated migrants. To overcome this problem, accounting for past migration flows - the so-called lags - allows to separately identify the effect of past migration episodes (Jaeger, Ruist and Stuhler, 2018[25]).

To investigate the presence of such a bias, the following model is estimated, which adds the lagged increase in higher-educated migrants in Equation 1:

$$\Delta\, log\, Y_{r,t+5} = \beta_0 + \beta_1 \Delta Migrants_{rt}^{HE} + \beta_2 \Delta Migrants_{rt-5}^{HE} + \beta_3 X_{rt} + \alpha_t + \Delta\varepsilon_{rt} \quad (2)$$

The equation is estimated for the period 2011 to 2016 using a 5-year change. $\Delta Migrants_{rt-1}^{HE}$ represents the increase in higher-educated migrant population from 2001 to 2006, which corresponds to the previous period.

Accounting for the lagged migration inflows does not change the results. Table 7 provides the estimates of current and past increases in higher-educated migration on different measures of innovation. The estimates for the increase in higher-educated migration are similar to the baseline specification, and the difference is not statistically significant. In contrast, the effect of past migration inflows on current innovation is not statistically different from zero. Taken together, this points shows that the main estimates are driven by current migration inflows rather than long-term effects of past migration inflows. In consequence, the main estimates are not biased by adjustment to previous changes in higher-educated migrant population.

Table 7. Controlling for past migration flows

	Patents	Trademarks	Design Rights
	(1)	(2)	(3)
Δ HE Migrants	4.077*	-1.825	-0.177
	(2.144)	(1.533)	(6.460)
Δ HE Migrants, past	0.476	1.581	2.132
	(2.591)	(1.713)	(7.023)

Note: The table presents results based on estimating Equation 2 using as outcomes either Patents (Column 1), Trademarks (Column 2) and Design rights (Column 3) applications. The dependent variables are expressed as log changes in IP applications per worker, using 5-year differences. The independent variable is the increase in employment due to higher-educated migrants, where the measure of high education is tertiary education (at least with a college degree), which is instrumented. Additionally, the increase in higher-educated migrants from 2001 to 2006 is included as an additional explanatory variable. All specification control for baseline shares of higher-educated population, log of population and industry shares. All specifications are weighted by the number of employed natives in the region at the baseline year. Robust standard errors are provided in parentheses. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.
Source: OECD calculations based on MADIP, IPGOD and the Australian Census of Population of 1981, 2001 and 2006 (accessed July 2023).

Validity test 3: Pre-existing growth in innovation is not associated with the current migrant flows

Another potential concern is the presence of trends on pre-existing levels of innovation. The effect of higher-educated migrants on innovation could be driven by regions experiencing higher growth in innovation even before the arrival of migrants. In consequence, if the instrument is correlated with these pre-existing levels, then the exogeneity condition is violated. To test whether the instrument is related to previous trends in innovation, the following reduced-form regression is estimated:

$$\Delta \log Y_{r,t-1} = \beta_0 + \beta_1 \Delta \widehat{Migrants}^{HE}_{r,t} + \beta_2 X_{rt} + \alpha_t + \Delta\varepsilon_{rt} \quad (3)$$

Where the explanatory variable is the predicted change in higher-educated migration, $\Delta \widehat{Migrants}^{S}_{r,t}$, which is the instrumental variable, and the outcome represents the change in innovation in the 2001-2006 period.

The instrument is not correlated with previous changes in IP rights applications. Table 8 shows that the instrument is not associated with previous applications of patents, trademarks, or design rights. None of the coefficients is significant, indicating that the instrument is not associated with trends in the outcomes.

Table 8. Impact of the instrumental variable on past IP rights applications

	Patents, past			Trademarks, past			Design rights, past		
	5 years	3 years	1 year	5 years	3 years	1 year	5 years	3 years	1 year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
∆ Predicted HE Migrant	0.354	2.384	3.386	-0.045	0.148	-0.919	0.230	-5.612	-7.164
	(1.651)	(2.041)	(2.886)	(1.044)	(1.694)	(1.734)	(2.920)	(5.528)	(6.846)
N	88	176	616	88	176	616	88	25.357	616

Note: The table presents results based on estimating Equation 3 using as outcomes either Patents (Column 1), Trademarks (Column 2) and Design rights (Column 3) applications. The dependent variables are expressed as log changes in IP applications per worker, using 5-year differences, but in the period 2001-2006. The independent variable is the predicted increase in employment due to higher-educated migrants, as calculated by the instrumental variable, where the measure of high education is tertiary education (at least with a college degree). The columns present different time intervals for both the dependent and independent variables. Columns 1, 4 and 7 represent five-year changes (2011-2016), Columns 2, 5 and 8 represent three-year changes (2011-2014 and 2014-2017), and Columns 3, 6, and 9 represent annual changes. No intervals overlap. All specification control for baseline shares of higher-educated population, log of population and industry shares. All specifications are weighted by the number of employed natives in the region at the baseline year. Standard errors, in parentheses, are clustered at the regional level in all specifications. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.
Source: OECD calculations based on MADIP and IPGOD (accessed July 2023).

Annex C. Robustness checks

Conducting various robustness checks allows to assess the sensitivity of the results and validates the empirical strategy. This section provides the robustness checks, where each panel includes a different econometric specification. All regressions are estimated using 2SLS. For comparison, Panel A provides the main results from the 2SLS estimates in Panel B of Table 2. Overall, the magnitude and statistical significance of the coefficients are consistent with the baseline results, proving the robustness of the results.

Considering an alternative measure of cumulative innovation does not alter the results. The first robustness test employs cumulative innovation activity, acknowledging that the impact of migrants on innovation might not materialise immediately. While some previous research suggests a time frame of half a year to two years for innovation to develop (Hunt and Gauthier-Loiselle, 2010[6]), recent studies indicate that migrants may reach their peak in innovative activity four years after their arrival (Blit, Skuterud and Zhang, 2020[23]). Since the specific data required to calculate this time frame for Australia is unavailable, the analysis implements a robustness check using cumulative innovation after the increase in migration.27 Panel B of Table 9 uses cumulative IP rights applications during the three to five years after the increase in migration as the dependent variable. The coefficients remain robust, indicating that the influence of migration on innovation does not dissipate immediately after migrants' arrival.

Considering contemporaneous changes in the migrant and patenting activity allows to uncover whether the arrival of higher-educated migrants has an immediate effect on innovation. To demonstrate leading outcomes by one year does not drive the results, Panel C of Table 9 presents results using contemporaneous changes in IP rights applications. The fact that one-year changes do not exhibit a significant impact on innovation further validates the chosen strategy, as any existing impact within a maximum of one year should be relatively moderate, if present at all. Moreover, the coefficients are smaller compared to the baseline strategy, which is consistent with the notion that the impact of migration on innovation takes time to materialise.

To ensure that outliers do not drive the effects, the sample is winsorised by dropping regions in the top 5% and bottom 5% in terms of patenting activity. Panel D of Table 9 presents these results. The estimated effect supports the baseline estimates, as it is even stronger, indicating that outliers drive the significance and magnitude of the results.

Dropping specific regions from the sample does not affect the findings. Panels E and F of Table 9 drop either the Outback or the largest two metropolitan areas (Sydney and Melbourne), respectively. Panel E shows that excluding the Outback from the sample leaves results unchanged, thus showing that sparsely populated regions do not drive the results. Panel F excludes regions which are part of the two largest metropolitan (Sydney and Melbourne), home to 40% of the population and a large number of migrant populations. Yet, the results remain roughly unchanged, although the levels of significance become lower due to larger standard errors.

[27] Cumulative innovation is defined as the total number of applications three or five years after the migration inflow. For instance, for the five-year period 2011-2016, the outcome is cumulative IP rights applications from 2016 to 2021.

Aggregating regions composing the capital regions of the Australian States and Territories yield comparable but slightly less significant estimates. Native workers may either move to other regions as a reaction to migration, creating a problem of spatial interdependence. Such mobility might lead to spill-overs across different regions, potentially biasing the results in Australian metropolitan areas that are disaggregated into multiple regions. To test whether spatial interdependence drives the results, the analysis uses the OECD's TL3 classification. This classification merges regions of metropolitan areas in the capital regions of Australian States and Territories into a single geographical unit, reducing the number of regions from 88 to 50. The remaining regions remain unchanged. Panel G of Table 9 shows that the results are robust to this choice, although the point estimate from the regression using one-year differences becomes noisier. The point estimates are generally slightly smaller but always within the standard errors of the baseline estimates. The positive bias of coefficients in regressions using variation across regions is likely due to increases in innovation in metropolitan regions being driven by increases in higher-educated migrants in neighbouring regions (Butts, 2023[33]).

Alternative definitions of innovation confirm that the total count definition does not drive the effect. A definition used in the literature is the fractional count of patents. Panel H of Table 9 investigates the sensitivity to using fractional rather than the total count of IP rights assigned to each region. The magnitudes remain similar to the total count definition, thus showing that the potential under-representation of less dense regions due to the fractional count does not affect the results.

Excluding the increase in higher-educated natives from the regression does change the estimated coefficients of the impact of higher-educated migrants. A correlation between increases in higher-educated migrants and higher-educated natives can potentially lead to biased results due to high multicollinearity. However, the correlation between these two variables is 0.73 for one-year changes, 0.71 for three-year changes and 0.68 for five-year changes. Although high, these variables are not perfectly correlated, and Panel I of Table 3 shows that dropping the increase in higher-educated natives from the estimation does not significantly alter the results.

Table 9. Robustness of the main results

	Patents			Trademarks			Design rights		
	5 years	3 years	1 year	5 years	3 years	1 year	5 years	3 years	1 year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Baseline									
∆ HE Migrants	4.754***	6.351**	6.564**	-1.459	-1.919	-0.245	0.754	8.725	14.621*
	(1.726)	(2.832)	(3.264)	(1.225)	(2.671)	(2.403)	(5.727)	(7.235)	(8.241)
F-stat	27.1	25.3	43.1	27.1	25.3	43.1	27.1	25.3	43.1
Panel B: Cumulative innovation									
∆ HE Migrants	3.063**	4.284***		-1.066	-1.057		4.202*	5.592	
	(1.137)	(1.949)		(0.651)	(1.166)		(2.420)	(4.562)	
F-stat	29.9	18.3		29.9	18.3		29.9	18.3	
Panel C: Contemporaneous changes									
∆ HE Migrants	3.794*	5.412***	2.934	-0.157	-1.946	0.434	1.915	-6.413	-0.425
	(2.045)	(1.988)	(2.469)	(1.130)	(1.393)	(1.515)	(4.982)	(6.914)	(7.531)
F-stat	27.1	25.3	43.1	27.1	25.3	43.1	27.1	25.3	43.1
Panel D: Dropping top and bottom five percentiles of innovation									
∆ HE Migrants	6.243***	8.751***	9.025***	-1.951	-4.790**	-1.62	2.348	13.486	10.817
	(1.597)	(2.778)	(3.219)	(1.207)	(1.976)	(2.281)	(6.585)	(8.919)	(9.412)
F-stat	19.1	18.6	34	19.1	18.6	34	19.1	18.6	34
Panel E: Omitting Outback									
∆ HE Migrants	4.582**	6.885**	6.872**	-1.597	0.343	0.104	2.509	10.293	15.207

	(1.789)	(2.956)	(3.354)	(1.395)	(3.358)	(2.920)	(6.267)	(7.909)	(8.817)
F-stat	23.1	21.8	37.4	23.1	21.8	37.4	23.1	21.8	37.4
Panel F: Omitting Sydney and Melbourne									
Δ HE Migrants	5.026**	8.111**	7.374*	-1.597	0.343	0.104	-4.215	1.668	3.423
	(2.178)	(3.653)	(4.182)	(1.395)	(3.358)	(2.920)	(3.371)	(7.013)	(7.637)
F-stat	26.7	21.8	28.5	26.7	21.8	28.5	26.7	21.8	28.5
Panel G: TL3									
Δ HE Migrants	3.890**	5.735**	5.587	-1.482	0.199	-0.347	-4.215	1.668	3.423
	(1.657)	(2.374)	(4.446)	(1.300)	(3.104)	(3.388)	(3.371)	(7.013)	(7.637)
F-stat	111.3	121.6	75.7	111.3	121.6	75.7	111.3	121.6	75.7
Panel H: Fractional count									
Δ HE Migrants	4.229**	7.701***	7.367**	-0.818	-2.082	0.416	3.365	6.645	15.834*
	(1.667)	(2.940)	(3.318)	(1.150)	(2.326)	(2.230)	(6.074)	(6.920)	(8.750)
F-stat	27.1	25.3	43.1	27.1	25.3	43.1	27.1	25.3	43.1
Panel I: Excluding Δ HE Natives									
Δ HE Migrants	4.788**	6.329**	5.480*	-1.550	-1.940	-0.259	0.631	8.895	13.115*
	(1.676)	(2.827)	(2.835)	(1.256)	(2.655)	(1.946)	(5.752)	(7.365)	(7.639)
F-stat	26.8	27.1	48.4	26.8	27.1	48.4	26.8	27.1	48.4

Note: The table presents results based on estimating Equation 1 using as outcomes either Patents (Columns 1-3), Trademarks (Columns 4-6) and Design rights (Columns 7-9) applications. Panel A provides baseline results, while Panels B to I perform different robustness checks, as explained in the text. The independent variable is the increase in employment due to higher-educated migrants, where the measure of high education is tertiary education (at least with a college degree). The columns present different time intervals for both the dependent and independent variables. Columns 1, 4 and 7 represent five-year changes (2011-2016), Columns 2, 5 and 8 represent three-year changes (2011-2014 and 2014-2017), and Columns 3, 6, and 9 represent annual changes. No intervals overlap. All specification control for baseline shares of higher-educated population, log of population and industry shares. All specifications are weighted by the number of employed natives in the region at the baseline year. Time fixed-effects are applied to account for time-varying shocks affecting the entire country, except for five-year changes due to collinearity. Standard errors, in parentheses, are clustered at the regional level in all specifications. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.
Source: OECD calculations based on MADIP and IPGOD (accessed July 2023).

Annex D. Definitions of higher-skilled and higher-educated migrants

Different definitions of skills and education are used: higher-educated (at least a college degree), higher-skilled occupations and scientists. To assign whether an individual has tertiary education or not, data from the 2016 Census is used, using the Level of Highest Educational Attainment (HEAP) variable. Individuals with Postgraduate Degree Level, Graduate Diploma and Graduate Certificate Level, and Bachelors Degree Level, which represent categories 1 to 3 in the data, are the ones classified as tertiary educated.

The classification of individuals to higher-skilled occupations follows previous research and uses ANZSCO occupation classification[28]. Categories 1 (Managers) and 2 (Professionals) of the ANZSCO one-digit classification are used to identify higher-skilled individuals. Similarly, to identify scientists, occupations that are more likely to innovate are used (Dotzel and Wojan, 2022[34]), following previous literature (OECD, 2022[35]). The following two-digit occupations are used: 13 (Specialist Managers), 23 (Design, Engineering, Science and Transport Professionals), 26 (ICT Professionals), and 31 (Engineering, ICT and Science Technicians).

[28] More information on the Occupation and Skill Classification are provided by the ABS (1220.0 - ANZSCO -- Australian and New Zealand Standard Classification of Occupations, 2013, Version 1.3 (abs.gov.au)).

Annex E. Additional results

Table 10. Additional results for uneven effects of migrants with different education levels

	Patents			Trademarks			Design rights		
	5 years	3 years	1 year	5 years	3 years	1 year	5 years	3 years	1 year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: All migrants except scientists									
∆ Migrants, no scientist	1.776**	2.518**	2.653*	-0.438	-0.848	0.090	0.997	2.653	5.191
	(0.713)	(1.109)	(1.653)	(0.459)	(1.019)	(0.973)	(2.271)	(2.990)	(3.890)
F-stat	30.5	28.8	79.4	30.5	28.8	79.4	30.5	28.8	79.4
Panel B: Tertiary-educated migrants, except scientists									
∆ HE Migrants, no scientists	7.707***	10.349**	8.100*	-2.381	-2.882	0.119	1.937	17.830	17.899
	(2.843)	(1.408)	(4.493)	(2.191)	(4.712)	(4.712)	(9.575)	(12.869)	(11.881)
F-stat	21.4	21.4	49.2	21.4	21.4	49.2	21.4	21.4	49.2

Note: The table presents results based on estimating Equation 1 using as outcomes either Patents (Columns 1-3), Trademarks (Columns 4-6) and Design rights (Columns 7-9) applications. Panel A uses all migrants but omits scientists. Panel B uses higher-educated migrants but omits scientists. The dependent variable is expressed as log changes in IP applications per worker, using either 5, 3, or 1-year differences. The independent variable is the increase in employment by each type of migrant. The columns present different time intervals for both the dependent and independent variables. Columns 1, 4 and 7 represent five-year changes (2011-2016), Columns 2, 5 and 8 represent three-year changes (2011-2014 and 2014-2017), and Columns 3, 6, and 9 represent annual changes. No intervals overlap. All specification control for baseline shares of higher-educated population, log of population and industry shares. All specifications are weighted by the number of employed natives in the region at the baseline year. Time fixed-effects are applied to account for time-varying shocks affecting the entire country, except for five-year changes due to collinearity. Standard errors, in parentheses, are clustered at the regional level in all specifications. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.
Source: OECD calculations based on MADIP and IPGOD (accessed July 2023).