Highly skilled migration, science and innovation : three essays

Edoardo Ferrucci

▶ To cite this version:

Edoardo Ferrucci. Highly skilled migration, science and innovation : three essays. Economics and Finance. Université de Bordeaux; Università di Torino (Turin, Italie), 2017. English. NNT : 2017BORD0854 . tel-03116719

HAL Id: tel-03116719
https://theses.hal.science/tel-03116719

Submitted on 20 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE EN COTUTELLE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ET DE L'UNIVERSITÉ DE TURIN

Ecole Doctorale Entreprise,économie, societé

Scuola di Dottorato in Scienze Umane e Sociali

SPÉCIALITÉ : ECONOMIE

Par Edoardo FERRUCCI

HIGHLY SKILLED MIGRATION, SCIENCE AND INNOVATION:
THREE ESSAYS

Sous la direction de Francesco LISSONI
et de Fabio MONTOBBIO

Soutenue le 15 Decembre 2017

Membres du jury :

M. SCELLATO, Giuseppe	Professeur, Università di Torino	Président - Rapporteur
M. NATHAN, Max	MCF, University College London	Rapporteur
M. LISSONI, Francesco	Professeur, Universitè de Bordeaux	Co-directeur de Thèse
M. MONTOBBIO, Fabio	Professeur, UCSC Milan	Co-directeur de Thèse

Titre :

Migration hautement qualifiées, science et innovation: trois essais

Résumé :

Comment les travailleurs des domaines des STIM, à savoir les chercheurs universitaires et les inventeurs, affectent-ils les activités et les résultats de la recherche et de l'innovation? Dans la présente thèse, j'essaie de répondre à la demande croissante d'une recherche centrée sur l'Europe. Elle s'articule autour de trois chapitres empiriques s'écartant de la plupart des´etudes précédentes qui évaluent les effets de l'immigration sur le marché du travail, en ce sens qu'elle essaie d'apporter des preuves à l'impact des travailleurs qualifiés étrangers sur l'innovation en Europe.

Les deux premiers chapitres sont centrés sur les activités de brevets européens. Ils se réfèrent à la littératureéconomique sur la qualité des brevets et / ou les brevets en tant qu'indicateurs de l'innovation. Le troisième chapitre est davantage axé sur la recherche sociologique, l'identité sociale et le comportement de groupe.

Dans le premier chapitre J'explore le lien entre performance et connexions ethniques d'une catégorie professionnelle, les inventeurs, dont l'engagement dans les activités d'innovation demeure uneévidence. Je me base sur une vaste littérature quiétudie l'effet de la diversité ethnique et culturelle sur l'innovation, ainsi que sur l'auto-sélection basée sur les compétences venant de la littérature des migrations.

Les deux premiers chapitres sont centrés sur les activités de brevets européens. Ils se réfèrent à la littératureéconomique sur la qualité des brevets et / ou les brevets en tant qu'indicateurs de l'innovation. Le troisième chapitre est davantage axé sur la recherche sociologique, l'identité sociale et le comportement de groupe.

Dans le premier chapitre J'explore le lien entre performance et connexions ethniques d'une catégorie professionnelle, les inventeurs, dont l'engagement dans les activités d'innovation demeure uneévidence. Je me base sur une vaste littérature quiétudie l'effet de la diversité ethnique et culturelle sur l'innovation, ainsi que sur l'auto-sélection basée sur les compétences venant de la littérature des migrations.

Le deuxième chapitre est le plus pertinent d'un point de vue politique. Il explore la contribution à l'innovation dans les pays de destination des inventeurs ayant quitté l'ex- URSS pour l'Europe et Isral, après l'effondrement de l'Union soviétique en 1991. Ce choc politique a eu des répercussions importantes en termes de flux migratoires de travailleurs hautement qualifiés vers les pays occidentaux. L'Europe, Isral et lesétats-Unis, et ainsi constitue unévènement naturel dans les migrations.

Le dernier chapitre cible les institutions académiques européennes, en particulier lesécoles de commerce publiques et privées, y compris les départements de gestion des grandes universités.

J’examine des publications scientifiques ayant plus d’un auteur, des revues de management. J’analyse ainsi leur impact scientifique et leur visibilité sur la base de la diversité culturelle des coauteurs. Dans la société moderne, lesécoles de commerce comptent parmi les organisations les plus internationalisées et les plus actives en termes d’activités dans le domaine des connaissances. La littérature contemporaine de management a toujours mis l’accent sur l’importance de la diversité de la main d’uvre dans l’amélioration ou la limitation de l’efficacité des groupes. Ces deux derniers points font de ce domaine un sujet approprié età la fois potentiellement attractif pour ces memes chercheurs qui y contribuent.

Mots clés : Innovation; Diffusion des connaissances; Migration internationale

---

Title :

Highly Skilled Migration, Science And Innovation: Three Essays

Abstract :

How STEM workers, namely academic researchers and inventors, affect research and innovation activities and outputs?
The present dissertation tries to meet such demand for a European-based research. Its three papers , which is made of three empirical chapters diverge from most of previous studies that evaluate the labour market effects of immigration in that it seeks to provide evidence on the impact of foreign skilled workers on innovation in Europe. The first two chapters center on European patenting activities They relate to and draw from the economic literature on patent quality and/or patents as innovation indicators. The third chapter is more compelled to the sociological research on social identity and group behavior.

In the first chapter I explore the link between performance and ethnic connections of a professional category, inventors, whose engagement in innovation activities is self-evident. I build upon a vast literature that investigates the effect of ethnic and cultural variety on innovation, as well as on the migration literature on skill-based self selection.

The second chapter is the most relevant from a policy perspective. It explores the contribution to innovation in destination countries by inventors who left former USSR countries for Europe and Israel, after the collapse of the Soviet Union in 1991. This was a political shock that had both important repercussions in terms of highly skilled migratory flows towards Western Europe, Israel, and the United States, and serves well as a natural experiment in migration.

The last chapter targets European academic institutions, in particular both public and private business schools, including management departments of large universities.
I examine multi-authored scientific publications on peer-reviewed management journals, and explain their scientific impact and visibility with the cultural diversity of the co- authors. Business schools are one of the most internationalized and knowledge-intensive organizations in modern society, and contemporary management literature has widely investigate the importance of workforce diversity in enhancing or limiting group effectiveness, two circumstances that make the chosen field of study both appropriate and extremely likely to attract the attention of the same scholars that contribute to it.

Keywords : Innovation; Knowledge diffusion; International migration

---

Unité de recherche : GREThA

Groupe de Recherche en Économie Théorique et Appliquée

UMR CNRS 5113

Acknowledgments

Over the past four years I have received support and inspiration from a great number of individuals.

First and foremost I wish to thank my supervisor, Professor Francesco Lissoni. He has been more than a mentor, more than a colleague. His guidance has made this a thoughtful and rewarding journey. Ever since my first year in Turin he has backed me up, and during the most difficult times when writing this thesis, he gave me the moral support and the freedom I needed to move on. Thank you Francesco, I will never forget it.

I would like to thank the members of my thesis committee Max Nathan, Giuseppe Scellato and Fabio Montobbio for their support and for the time they spent evaluating this work.

The colleagues that I met both in France and in Italy have contributed immensely to my personal and professional growth. They have been a source of friendships as well as good advice and collaboration. In particular I would like to thank Claudia, Gianluca, Ernest, Valerio and Gabriele: I wish you guys all the best, and I hope we will be friends for a very long time.

This thesis would not have been possible without the help and support of the University of Turin and the University of Bordeaux. A very special gratitude goes out to all the GREThA members, in particular to Cristophe Carrincazeaux who funded the research through the PROXIMO project.

I want to thank Martina for everything she gave me during my good and bad times; we grew a lot in these years and I am really proud of us.

Finally, thank you to my parents Manrico and Lucia. They have always left me the freedom to choose what to do with my life; they have cherished with me every great moment and supported me whenever I needed it. Grazie.

Contents

Introduction **11**

1 Foreign inventors in the US and EU15: diversity and productivity **15**

1.1 Introduction 15
1.2 Highly skilled migrants 17
1.3 Data and descriptives 21
1.3.1 Data 21
1.3.2 An overview of foreign inventors in the US and EU15 23
1.4 Diversity test: nationality and inventor teams' composition 24
1.5 Econometric analysis: migration, diversity and patent quality 29
1.5.1 Methodology 29
1.5.2 Main results 32
1.5.3 Robustness checks and endogeneity issues 37
1.6 Discussion and conclusions 41
Appendix A 43
Appendix B 61

2 Migration and local innovation: post-USSR inventors in Europe and Israel **67**

2.1 Introduction 67
2.2 Migration and innovation literature 69
2.3 Historical context 72
2.3.1 The breakdown of USSR 72
2.3.2 Science and technology in the Soviet Union 73
2.3.3 The Intellectual Property Rights system 74
2.4 Data and sample 76
2.5 Descriptive evidence 79
2.6 Empirical strategy 84
2.6.1 Model 1: post-1991 evidence 85

2.6.2 Model 2: pre-1992 evidence 86
2.7 Results 88
2.8 Concluding remarks 95

3 Internationalization of Business Research Teams: How Cultural Diversity affects Research Performance? 97

3.1 Introduction 97
3.2 Literature review 99
3.2.1 Diversity in organizations and its performance implications 99
3.2.2 Business knowledge production landscape 102
3.3 Hypotheses development 103
3.3.1 Cultural diversity and team research performance 103
3.3.2 Geographic dispersion of teams 105
3.4 Research Design 106
3.4.1 Sample 106
3.4.2 Econometric framework and variables 107
3.5 Empirical Results 113
3.6 Discussion and conclusions 121

List of Figures

1.1 Graphical diversity test. 28
A1 US Main Immigrant Inventors Groups Patenting Activity 44
A2 US Main Immigrant Inventors Groups 44
A3 US Inventor Team Size by Nationality Composition 47
A4 US Inventors by Applications *ISI-OST-INPI* Technological Class 48
A5 EU15 Main Immigrant Inventors Groups Patenting Activity 50
A6 EU15 Main Immigrant Inventors Groups 50
A7 EU15 Inventor Team Size by Nationality Composition 54
A8 EU15 Inventors by Applications *ISI-OST-INPI* Technological Class 54
A9 EU15 Resident Foreign Nationality Inventors (Intra-European Migrants Included) 55
A10 EU15 Main Immigrant Inventors Groups (Intra-European Migrants Included) 55
A11 EU15 Main Immigrant Inventors Groups Patenting Activity 57
A12 EU15 Main Immigrant Inventors Groups 57
A13 US Main Immigrant Inventors Groups Patenting Activity 59
A14 US Main Immigrant Inventors Groups 59

2.1 EPO patents by USSR applicants. 76
2.2 1991-1995 patent production by country. 81
2.3 1996-2000 patent production by country. 81
2.4 IPC35 classification 82
2.5 Yearly arrivals of ex-USSR inventors by age of the technological class. 83
2.6 Average patent production. 88
2.7 Average active inventors. 88
2.8 OLS estimates, annual coefficients. 95

3.1 Research quality, interaction. Marginal effects. 116
3.2 Average Team Cultural diversity. 117

List of Tables

1.1 US sample, comparison of diversity indexes. 26

1.2 Distribution comparison of diversity indexes 27

1.3 Summary statistics by sample, EU15 33

1.4 Summary statistics by sample, US 33

1.5 OLS estimation, inventor team diversity and patent quality (Eq. 3). 34

1.6 OLS estimation, inventor team weighted diversity and patent quality (Eq. 3). . 35

1.7 OLS estimation, foreign inventorship and patent quality (Eq. 4). 36

1.8 OLS estimation for top cited patents. Dependent variable is log(1+cit3). 38

1.9 OLS estimation for top cited patents. Dependent variable is log(1+cit3). 38

1.10 OLS estimation of US patents by main technological class. Dependent variable is log(1+cit3). 39

1.11 Baseline model (Eq. 3), OLS estimation of EU15 patents by main technological class. Dependent variable is log(1+cit3). 39

1.12 2SLS estimation. We instrument **diversity** with **foreign share (t-1)** and **country diversity (trend)**. US patents signed before 2000 are excluded due to missing data. 40

1.13 2SLS estimation. We instrument **weighted diversity** with **foreign share (t-1)** and **country diversity (trend)**. US patents signed before 2000 are excluded due to missing data. 41

A1 US Inventors by origin 43

A2 US Main Immigrant Inventors Groups 45

A3 US Immigrant Inventors by Main CSAs 45

A4 US Native Inventors by Main CSAs 46

A5 US Inventors Team Size by Nationality Composition* 47

A6 EU15 Inventors by origin 49

A7 EU15 Main Immigrant Inventors Groups 51

A8 EU15 Immigrant Inventors by Main Metropolitan Regions 51

A9 EU15 Native Inventors by Main Metropolitan Regions 52

A10 EU15 Inventors Team Size by Nationality Composition* 52

A11 EU15 Main Immigrant Inventors Groups (Intra-European Migrants Included) 53

A12 EU15 Immigrant Inventors by Main Metropolitan Areas 58
A13 US Immigrant Inventors by Main CSAs 60
B1 Negative Binomial estimation, inventor team diversity and patent quality 61
B2 Negative Binomial estimation, inventor team weighted diversity and patent quality 62
B3 Negative Binomial estimation, foreign inventorship and patent quality. 63
B4 Negative Binomial estimation on top cited patents. Dependent variable is cit3. . 64
B5 Cross-correlation table, EU15 65
B6 Cross-correlation table, US 66

2.1 IPC classification of fast fission reactors. 78
2.2 Number of ex-USSR inventors by country of origin and area of destination. 79
2.3 Temporal distribution of technological classes' birth. 84
2.4 Model 1. 90
2.5 Model 1, controlling for patent quality. 90
2.6 Model 1, controlling for non-exUSSR foreign activity. 91
2.7 Model 1, Controlling for previous inventor activity. 91
2.8 Model 2, USSR-based PCT technologies. Whole sample. 92
2.9 Model 2, USSR-based PCT technologies. Matched sample. 92
2.10 Model 2, USSR-based PCT technologies Matched sample. Controlling for previous inventor activity. 94
2.11 Model 2, USSR-based EPO technologies Matched sample. IPC Sections. . 94

3.1 Cultural groups. 108
3.2 Summary statistics. 111
3.3 Correlation matrix 112
3.4 Research quality, ordinal logit model. 114
3.5 Research quality: marginal effects of Model (3). 114
3.6 Research quality, testing for moderating effect. Ordinal logit. 115
3.7 Research impact. 118
3.8 Research impact. 119
3.9 Research impact, testing for moderating effect. 120

Summary in French - Résumé en Français

Les trois essais que comportet cette dissertation contribuent á la littérature émergente sur les migrations internationales des travailleurs hautement qualifiés et l'innovation (Breschi et al. 2016), Kerr et al. 2016). Autrefois réservée au domaine de l'histoire économique, l'étude des liens entre les deux phénomènes s'est longtemps limitée á l'analyse des diffusions (Lissoni 2017), plus particulièrement celles des vagues migratoires de groupes minoritaires de travailleurs qualifiés ou d'entrepreneurs originaires de pays plus ou moins développés (Cipolla 1972; Hornung 2014; Moser et al. 2014). Durant es années 1960 et 1970, des économistes du développement en mettant un accent sur le brain-drain, ont donné une dimension de problème politique et de question théorique á ce phénomène. L'argument majeur avancé par ces économistes était que ces flux migratoires représentaient une perte de richesse pour les pays d'origine tels que l'Inde et d'autres anciennes colonies; perte qui devait être compensée par la suite par des transferts en provenance des pays de destination. (Bhagwati and Hamada 1974; McCulloch and Yellen 1977).

La littrature rcente est beaucoup plus hétérogène, á la fois d'un point de vue disciplinaire et en termes de questions de recherche. Premièrement, elle est motivée par la reprise des masses migratoires (dont les chiffres ont atteint et même dépassé ceux d'avant la crise financière) et les proportions croissantes de sa portion de travailleurs hautement qualifiés (Arslan et al. 2015, Hatton 2014). Les niveaux d'éducation des migrants venant des pays développés et en voie de développement se sont considérablement améliorés au cours de la dernière décennie. Les étrangers hautement éduqués (c'est á dire ceux ayant une formation de niveau tertiaire, souvent dispensée dans les collèges et / ou les universités) représentent aujourd'hui un chiffre d'environ 30 millions dans la zone de l'OCDE, avec près de 40% de la croissance des populations étrangères (OECD 2016).

Deuxièmement, le nouveau champs de littérature met en évidence un nouveau sujet qui s'ajoute á la liste de ceux ayant été abordés par l'ancien champs: il s'agit de l'impact des migrants hautement qualifiés sur le potentiel d'innovation des pays de destination et des lieux ou secteurs d'activité économique spécifiques. Etant donné qu'une grande majorité des nouveaux migrants hautement qualifiés vienne des pays les moins avancés pour les pays les plus avancés sur le plan scientifique ou technologique, ou se compose

de jeunes poursuivant leurs études (pour un doctorat ou un post-doctorat ou d'autres emplois), l'attention n'est plus ou pas seulement á la diffusion de l'innovation, ce qui rend ce nouveau champs de littérature différent de l'ancienne tradition historique.

Les géographes économiques, par exemple, examinent l'impact des migrations sur les taux d'innovation des régions et des villes (Bellini et al. 2013; Kemeny and Cooke 2015; Niebuhr 2010; Ozgen et al. 2013). Les économistes du travail se concentrent sur la place qu'occupent les scientifiques et les ingénieurs étrangers dans les économies des états-Unis et d'autres pays développés, notamment en relation avec les éventuels effets de déplacement aux dépens des chercheurs et ingénieurs locaux (Hunt and Gauthier-Loiselle 2010, Hunt 2015) ou des équilibres géographiques préexistants. Ceci représente donc une rupture traditionnelle de l'attention initialement portée aux migrants peu qualifiés et de leur potentiel á provoquer une baisse des salaires et et une augmentation du chmage des travailleurs locaux (Borjas 1987; Borjas 2003; Altonji and Card 1991; Ottaviano and Peri 2012). Cependant, cette rupture qui demeure principalement centrée sur les états-Unis, s'accorde sur une principal point qui est l'impact significatif sur la croissance des taux d'innovation des pays de destination venant des étudiants et travailleurs étrangers spécialisés dans les domaines des STIM; tout cela étant d á une sélection positive basée sur le niveau de qualification et á des effets de composition (STIM signifie sciences, technologie, ingénierie mathématiques, qui sont des domaines dans lesquels on retrouve le plus d'étudiants internationaux en comparaison aux étudiants américains locaux).

Néanmoins, ces axes de recherche trouvent écho auprès des décideurs politiques européens qui sont confrontés á une pénurie croissante de compétences scientifiques et technologiques, en raison de leur volonté de transformer leurs propres économies en économies de connaissances, ce qui nuit aux facteurs démographiques. En outre, l'ouverture des frontières nationales á des individus hautement qualifiés rencontre de moins en moins de résistance auprès d'une opinion publique qui á la base se serait opposée aux migrations, ceci en raison de la baisse d'appréhension sur l'intégration ou la compétition pour la richesse.

C'est le cas par exemple de l'Europe, qui avec une population vieillissante et moins importante, pourrait faire face dans un avenir proche á un manque de disponibilité de main-d'uvre qualifiée suffisante (Gavigan et al. 1999, Mahroum 2001). Il y a eu plusieurs initiatives afin d'attirer plus de travailleurs qualifiés étrangers en Europe. Par exemple, la Commission européenne a tenté á plusieurs reprises d'assouplir les exigences légales de la carte bleue (un permis de travail européen) en termes d'exigences salariales puisqu'á ce jour, elle n'est accordée qu'aux ressortissants de pays tiers avec un revenu valant une fois et demi le revenu moyen de l'état de l'UE o ils prévoient de s'installer.

Dans la présente thèse, j'essaie de répondre á la demande croissante d'une recherche centrée sur l'Europe. Elle s'articule autour de trois chapitres empiriques s'écartant de la plupart des études précédentes qui évaluent les effets de l'immigration sur le marché du travail, en ce sens qu'elle essaie d'apporter des preuves á l'impact des travailleurs qualifiés étrangers sur l'innovation en Europe. Comment les travailleurs des domaines des STIM, á savoir les chercheurs universitaires et les inventeurs, affectent-ils les activités et les résultats de la recherche et de l'innovation?

Les deux premiers chapitres sont centrés sur les activités de brevets européens. Ils se réfèrent á la littérature économique sur la qualité des brevets et / ou les brevets en tant qu'indicateurs de l'innovation. Le troisième chapitre est davantage axé sur la recherche sociologique, l'identité sociale et le comportement de groupe.

Dans le premier chapitre[1] J'explore le lien entre performance et connexions ethniques d'une catégorie professionnelle, les inventeurs, dont l'engagement dans les activités d'innovation demeure une évidence. Je me base sur une vaste littérature qui étudie l'effet de la diversité ethnique et culturelle sur l'innovation, ainsi que sur l'auto-sélection basée sur les compétences venant de la littérature des migrations.

Je mène mon exploration pour le cas des états-Unis et l'Europe en tant que pays de destination. L Par Europe, j'entends l'UE15, á savoir l'Union européenne á ses frontières de 1995, qui englobe les plus grandes économies du continent, ainsi que les pays européens ayant le plus de demandes de brevets. J'utilise des données originales, á savoir les demandes de brevet déposées auprès de l'Office européen des brevets (OEB) selon la procédure du Traité de coopération en matière de brevets (PCT), par des équipes d'inventeurs résidant aux états-Unis et dans l'UE15. Ma première remarque porte sur la tendance disproportionnée des inventeurs migrants á intégrer des équipes composées de migrants venant d'un même pays d'origine, limitant ainsi la diversité observée. Deuxièmement, mes résultats mettent en exergue un fait important: la capacité de la présence de migrants dans une équipe d'inventeurs á accroitre la qualité des brevets est plus importante lorsqu' elle augmente également la variété de l'équipe.

Le deuxième chapitre[2] est le plus pertinent d'un point de vue politique. Il explore la contribution á l'innovation dans les pays de destination des inventeurs ayant quitté l'ex-URSS pour l'Europe et Isral, après l'effondrement de l'Union soviétique en 1991. Ce choc politique a eu des répercussions importantes en termes de flux migratoires de travailleurs hautement qualifiés vers les pays occidentaux. L'Europe, Isral et les états-Unis, et ainsi constitue un évènement naturel dans les migrations.

Je teste l'hypothèse selon laquelle les inventeurs de l'ex-URSS auraient généré des externalités de connaissances positives dans les pays de destination, en relation avec des résultats similaires dans la littérature au sujet de la recherche scientifique (notamment en mathématiques). Sur la base d'un ensemble de données similaires á celles du chapitre 1, je retrace l'activité des inventeurs actifs en Europe occidentale et en Isral avant et après l'effondrement de l'Union soviétique, et teste l'hypothèse d'une activité inventive dans

[1]Basé sur un article co-écrit avec Francesco Lissoni.
[2]Basé sur un article écrit par moi seul.

les classes technologiques les plus touchées par le choc migratoire accrue par rapport á l'activité dans les autres classes.

Ceci en:

- identifiant les classes technologiques dans lesquelles les scientifiques et ingénieurs soviétiques pourraient potentiellement contribuer.
- créant un groupe apparié de technologies de contrle, sur la base d'un ensemble de caractéristiques mesurées avant 1992.

Les preuves étayent l'idée d'un important flux de connaissances transmises aux inventeurs européens et israéliens par ceux l'ex-URSS, et les résultats proviennent principalement de domaines scientifiques tels que la chimie et la physique. En s'appuyant sur des preuves historiques des schémas migratoires Est-Ouest, vérifiant l'hétérogénéité des effets á l'intérieur des cohortes de migrants, je trouve peu ou pas du tout de preuves soutenant l'hypothèse d'une auto-sélection basée sur les compétences.

Le dernier chapitre[3] cible les institutions académiques européennes, en particulier les écoles de commerce publiques et privées, y compris les départements de gestion des grandes universités.

J'examine des publications scientifiques ayant plus d'un auteur, des revues de management. J'analyse ainsi leur impact scientifique et leur visibilité sur la base de la diversité culturelle des coauteurs. Dans la société moderne, les écoles de commerce comptent parmi les organisations les plus internationalisées et les plus actives en termes d'activités dans le domaine des connaissances. La littérature contemporaine de management a toujours mis l'accent sur l'importance de la diversité de la main-d'uvre dans l'amélioration ou la limitation de l'efficacité des groupes. Ces deux derniers points font de ce domaine un sujet approprié et á la fois potentiellement attractif pour ces mêmes chercheurs qui y contribuent.

Les résultats de mon analyse soutiennent la vision de la théorie du traitement de l'information o la diversité culturelle apporte de multiples perspectives á la résolution de problèmes, á la créativité et á l'innovation en équipe: en termes de qualité de recherche, mes résultats confirment un lien positif avec la diversité culturelle, quoique un niveau d'hétérogénéité élevé semble être bénéfique seulement pour des publications publiées dans les meilleures revues. L'effet s'accentue lorsque les équipes sont plus dispersées dans l'espace: des valeurs communes et une origine culturelle partagée pourraient pallier le manque de proximité géographique entre chercheurs. L'impact de l'équipe sur la recherche est mesuré par le nombre de citations reues par son journal depuis la date de publication. L'effet de la diversité culturelle n'est positif que dans des équipes géographiquement dispersées.

[3]Basé sur un article co-rédigé avec David Yoon et Mustapha Belkhouja.

Introduction

The three essays that compose this dissertation contributes to the emerging literature on international highly skilled migration and innovation (Breschi et al. 2016, Kerr et al. 2016). Once the preserve of economic history, the investigation of links between the two phenomena has for long been confined to diffusion studies (Lissoni 2017), in particular to migratory waves of skilled or entrepreneurial minority groups from more to less advanced countries (Cipolla 1972; Hornung 2014; Moser et al. 2014). In the 1960s and 1970s, developement economists raised *brain-drain* to the status of political issue and theoretical topic, by positing the existence a welfare loss for the origin countries such as India and other former colonies, to be compensated by financial transfers from the destination countries (Bhagwati and Hamada 1974; McCulloch and Yellen 1977).

Recent literature is much more heterogeneous, both from a disciplinary viewpoint and in terms of research questions. First, it is motivated by the reprise of mass migration (whose figures reached and surpassed the pre-financial crisis ones) and the increasing weight of its highly skilled share (Arslan et al. 2015, Hatton 2014). The education attainments of migrants from both advanced and developing countries have greatly improved over the past decade. Highly-educated foreigners (i.e. the ones who own a tertiary-level education, often delivered at colleges and/or universities) nowadays amount to 30 millions in the OECD area, and they account for almost the 40% of the increase in the foreign born population (OECD 2016). Second, the new literature adds a new topic to the list of those treated by the old one: it investigates the impact of highly skilled migrants on the innovation potential of destination countries and specific locations or sectors of economic activity therein. Since much of the new highly skilled migrants move from less to more scientifically and technologically advanced countries, or includes young individuals who still need to complete their education (with a PhD or a post-doc or other formative early career jobs), the focus of attention is no more or not only diffusion of innovation, which makes this new wave of study different from the old historical tradition.

Economic geographers, for example, examine the impact of migration on the innovation rates of regions and cities (Bellini et al. 2013; Kemeny and Cooke 2015; Niebuhr 2010; Ozgen et al. 2013). Labour economists focus on how well foreign scientists and engineers fit in the economies of the US and other advanced countries, also in relation to possible displacement effects at the expenses of local knowledge workers (Hunt and Gauthier-

Loiselle 2010, Hunt 2015) or pre-existing geographical equilibria. This is a traditional departure from the original attention on low-skilled migrants and their potential to depress wages and increase unemployment of native workers (Borjas 1987; Borjas 2003; Altonji and Card 1991; Ottaviano and Peri 2012). However, such departure is still by and large US-centered, and converge in finding that both international STEM students and workers have contributed significantly to sustain the innovation rates in that country, both due to skill-based positive selection and composition effects (STEM stands for Science, Technology, Engineering & Mathematics, four subjects in which graduate international students enroll disproportionally, compared to US natives).

Still, these lines of research resonate well with policy-makers in Europe, too, who face increasing skill shortages in science and technology, due to their wish to transform their own economies into knowledge-driven ones, which hurts against demographic factors. Besides, opening the country borders to highly educated individuals meet less resistance from an otherwise migration-opposed public opinion, due to lower fears about integration or competition for welfare.

That is the case for Europe, whose population is rapidly shifting towards an older and smaller profile, the availability of sufficient skilled workforce may be compromised in the near future (Gavigan et al. 1999,Mahroum 2001). Proposals to attract more qualified foreigners to work in Europe. For example, the European Commission has made several attempts to relax the legal requirements of the Blue Card (a European work permit) in terms of salary requirement since as of today it is granted only to non-EU nationals who who will earn one and a half times the average wage in the EU country where they plan to move.

The present dissertation tries to meet such demand for a European-based research. Its three papers , which is made of three empirical chapters diverge from most of previous studies that evaluate the labour market effects of immigration in that it seeks to provide evidence on the impact of foreign skilled workers on innovation in Europe. How STEM workers, namely academic researchers and inventors, affect research and innovation activities and outputs?

The first two chapters center on European patenting activities They relate to and draw from the economic literature on patent quality and/or patents as innovation indicators. The third chapter is more compelled to the sociological research on social identity and group behavior.

In the first chapter[4] I explore the link between performance and ethnic connections of a professional category, inventors, whose engagement in innovation activities is self-evident. I build upon a vast literature that investigates the effect of ethnic and cultural variety on innovation, as well as on the migration literature on skill-based self selection.

[4]Based on a paper co-authored with Francesco Lissoni.

I conduct my exploration for both the US and Europe as destination countries. I define Europe as EU15, namely the European Union at its 1995 boundaries, which encompasses the largest economies of the Continent, as well as its most patent-intensive countries. I make use of original data, namely the patent applications filed at the European Patent Office (EPO) according to the Patent Cooperation Treaty procedure (PCT), by US- and EU15-resident teams of inventors. I first find that migrant inventors tend disproportionately to enter teams composed of fellow migrants from the same country, thus limiting observed variety. Second, I find that the presence of migrants in an inventor team increases patent quality the most when it also increases the team variety.

The second chapter[5] is the most relevant from a policy perspective. It explores the contribution to innovation in destination countries by inventors who left former USSR countries for Europe and Israel, after the collapse of the Soviet Union in 1991. This was a political shock that had both important repercussions in terms of highly skilled migratory flows towards Western Europe, Israel, and the United States, and serves well as a natural experiment in migration.

I test the hypothesis of former USSR inventors generating positive knowledge externality in destination countries, in line with similar findings in the literature concerning scientific research (most notably in mathematics). Based on a similar dataset as of chapter 1, I track the activity of inventors active in Western Europe and Israel before and after the collapse of the Soviet Union, and test the hypothesis that the inventive activity in the technological classes most affected by the migratory shock increased relatively to the activity in other classes.

I do so by:

- identifying the technological classes to whose advancement Soviet scientists and engineers could potentially contribute.
- creating a matched group of control technologies, on the basis of a set of characteristics measured prior to 1992.

Evidence supports the idea of a significant knowledge flow transmitted to European and Israelian inventors from ex-USSR ones, and results are mainly driven from scientific fields such as chemistry and physics. Building on historical evidences of East-to-West migratory patterns, check for heterogeneity of effects among cohorts of migrants, find little to no evidence that support the idea of skill-based self-selection.

The last chapter[6] targets European academic institutions, in particular both public and private business schools, including management departments of large universities.

I examine multi-authored scientific publications on peer-reviewed management jour-

[5]Based on a single authored paper.
[6]Based on a paper co-authored with David Yoon and Mustapha Belkhouja.

nals, and explain their scientific impact and visibility with the cultural diversity of the co-authors. Business schools are one of the most internationalized and knowledge-intensive organizations in modern society, and contemporary management literature has widely investigate the importance of workforce diversity in enhancing or limiting group effectiveness, two circumstances that make the chosen field of study both appropriate and extremely likely to attract the attention of the same scholars that contribute to it.

Current results support the view of information-processing theory where cultural diversity brings along multiple perspectives to problem-solving, creativity, and innovation in teams: in terms of research quality I find a positive association with team cultural diversity, although high levels of heterogeneity seem to be beneficial only in order to get published on the very top journals. The effect strengthened in teams that are more dispersed in space: common values and a shared cultural background may overcome the lack of geographical proximity among researchers. The research impact of the team is measured as the number of forward citations received by its paper since the publication date. The effect of cultural diversity is positive only in geographically dispersed teams.

Overall, this thesis confirms the existence of a link between migration and innovation: there is evidence on a circumstance which has been theorized in literature, that is the positive association between diversity and innovative performance (although I observe such dynamic at the level of the team, while most of previous research employ a geographical approach).

The third chapter underlines that ethnic diversity is beneficial only for teams of researchers who are dispersed in space; this hints to the central role of migration flows in contributing to the formation of multinational research teams. The presence of a self-selection process is less evident: while such mechanism could at least partially explain the results of the first chapter, it is almost completely absent from the scenario provided by the second chapter.

As for future plans of research, chapters 2 and 3 both require further developments. Chapter 2 will include the geographical heterogeneity in patenting/innnovativeness as additional explanatory factors, together with the variations in the diffusion of specific technological sectors across European regions. In chapter 3 the connection between geographical dispersion and migration will be explored more thoroughly, possibly including the whole history of institutional affiliations of the authors in the database.

Chapter 1

Foreign inventors in the US and EU15: diversity and productivity

1.1 Introduction

We live in a new era of international mass-migration, one that involves an unprecedented, rising number of highly skilled individuals (Arslan et al. 2015; Chaloff and Lemaitre 2009; Docquier et al. 2006; Hatton 2014; Özden et al. 2011). These include professionals such as scientists and engineers (S&E), who are directly involved in research and innovation activities (Franzoni et al. 2012; Freeman 2010). Policy-makers and scholars in top destination countries, enquire on the net benefits generated by these inflows (Borjas and Doran 2012a; Chellaraj et al. 2006; Kerr 2010). Those in countries where inflows are limited, and possibly offset by substantial outflows, investigate on how to increase them. All of them toy with the idea of selective, highly-skilled oriented immigration policies (Hatton 2014; Boeri 2012; Cerna and Czaika 2016).

For migrant S&E to enhance the innovation potential of destination countries, some skill-based, positive self-selection processes ought to be in place. Positively self-selected migrants are more likely both to have a career in the chosen field of study or expertise (Hunt 2011), and to deliver above-average performances (Moser et al. 2014; No and Walsh 2010; Stephan and Levin 2001).

Career-building in science and technology, however, also require social connections, both for finding a grant or a job and for joining or setting up a research team (little or no research can be conducted in isolation; see: Jones 2009; Wuchty et al. 2007). In this respect, ethnic connections may play an important role. They may be at the origin of the decision to move, as with chain migration; and they may compensate for the lack of social capital at destination (Beine et al. 2011; Kerr and Mandorff 2015). In this sense, they may help any highly skilled migrant to develop his/her full innovation potential. At the same time, however, over-reliance on ethnic connections may be indicative of some weakness at the individual level, as when a scientist or engineer is not brilliant enough to make it alone in the profession; or excessive social control, as when the ethnic group, or a reference person therein, limits its members' outside interactions. Notice that professional segregation along ethnic lines may limit the exploitation of the cultural variety that migrants bring along, which several studies have found to increase innovativeness at the city or regional level (Bellini et al. 2013; Ottaviano and Peri 2006; Ozgen et al. 2011).

In this paper, we explore the link between performance and ethnic connections of a professional category, inventors, which both comprises many scientists or engineers and whose engagement in innovation activities is self-evident.

In particular, we try to answer the following research questions : do migrant inventors activate or rely upon community-mediated mechanisms of collaboration and exchange? And if so, do such mechanisms simply compensate for limited access to other assets (thus resulting in undistinctive or possibly inferior performances)? Or do they pave the way for superior performances?

We conduct our exploration on both the US and Europe as destination countries. While the US have been the object of many inquiries (albeit none like ours), Europe is a much less studied case, possibly because its highly skilled immigration rates are lower than the US ones and mostly intra-continental (Dumont et al. 2010; Özden et al. 2011). Yet, highly skilled migration is hot topic in Europe, and its increase has been pursued, with mixed success, by several policies (Cerna and Chou 2014). In what follows, we define Europe as EU15, namely the European Union at its 1995 boundaries, which encompasses the largest economies of the Continent, as well as its most patent-intensive countries (with the main exception of Switzerland). We make use of original data, which consist of all patent applications filed at the European Patent Office (EPO), by US- and EU15-resident teams of inventors (at least two inventors per patent, with the same residence). Information on the inventors include their nationality, as derived from the WIPO-PCT database (Miguelez et al. 2013). We regress the quality of patents (as measured by forward citations) on the inventor team composition by nationality and find that less homogeneous teams produce higher quality patents. While lacking a clear causal interpretation, we consider our results to be an important starting point for future research on the emerging topic of migration and innovation.

The paper proceed as follows. In section 2 we review the rather heterogeneous literature on migration and innovation, with special emphasis on highly skilled migrants, their self-selection, and social connectedness. On this basis, we formulate some testable propositions. In section 3 we present our data and variables of interest. In section 4 we conduct our econometric exercise and comment its results. Section 5 concludes.

1.2 Highly skilled migrants

In this section we review, in both a selective and critical fashion, the existing literature on the innovation impact of highly skilled migration onto destination countries, with a special focus on S&E. We organize the review around three themes: self-selection, diversity, and ethnic ties. Most research concerns the US, due both to its historical importance as a country of immigration, and to the extraordinary attraction it exerts over foreign S&E. Fewer studies have concerned Europe, whose countries generally attract fewer highly skilled migrants than the US, and mostly from within Europe. The concept of self-selection is a key one in the migration literature, and it refers primarily to whether an individual's decision to migrate is positively or negatively related to her skill (self-selection, Belot and Hatton 2012). Extreme cases of positive self-selection show up in in studies such as Stephan and Levin 2001 and No and Walsh 2010. Stephan and Levin pioneer study examined highly productive or distinguished S&E (including inventors), active in the 1980s and 1990s in the US, and found that the foreign-born were over-represented (compared to foreign-born shares in the overall US S&E labour force). Quite interestingly, Stephan and Levin also detected a cohort effect, with the foreign-born entered in the US before 1945 being particularly productive. This is coherent with historical accounts of the specific importance, for the development of US science, of Jewish and other scientists fleeing Germany in the wake of WWI (Moser et al. 2014). At the same time, it reminds us that the largest benefits are to be expected when immigration occurs from a more advanced country (as Germany was, with respect to the US, in several scientific fields), possibly as a result of an exogenous shock[1].

Still, similar evidence has been found for more recent cohorts of migrants to the US. No and Walsh 2010 survey 1900 US-based inventors of triadic patents, asking them to self-evaluate their inventions' technological impact and economic value. Both measures are found to be higher for inventions by foreign-born inventors, after controlling for several confounding factors.

Chellaraj et al. 2006 make use of a production function approach to estimate the impact of both foreign-born highly skilled workers and international students on innovation in the US. The elasticity of patents to the presence of skilled immigrants is found to be positive and significant, and even more so the elasticity with respect to foreign graduate students. This difference is explained by an important composition effect: while foreign-educated, highly skilled immigrants comprise many professions, foreign graduate students in the US are concentrated in science and engineering and therefore have a much more

[1] The importance of migration shocks of this type has bee recognized for a long time by historians of technical change in modern Europe. For a classic discussion see Cipolla 1972.

direct impact on innovation[2]. The composition effect plays a role also in explaining the evidence produced by Hunt and Gauthier-Loiselle 2010 and Hunt 2015. However, Hunt 2015 also shows that, when looking at engineers by degree, rather than occupation, migrants overall are lower achievers than locals, since engineering graduates from the least developed and non-Anglophone countries face difficulties in getting an engineering job or in reaching managerial positions, being impeded by lack of language skills or social capital. On the contrary, engineering graduates from richer countries both find proper jobs and perform better than natives. This finding points to an important, albeit overlooked determinant of foreign S&E contribution to innovation, namely social capital, on which we will come back below. Wadhwa et al. 2007 investigate engineering and technology companies established in the U.S. between 1995 and 2005, and find that around 25% of them were founded or co-founded by at least one foreign-born entrepreneur. The percentage increases remarkably in high-tech clusters such as the Silicon Valley (52%) or New York City (44%). These foreign entrepreneurs are mostly found to hold doctoral degrees in S&E, and to be better educated than control groups of natives. However, it is not clear whether this is a sign of successful integration in the US economy, or the result of a classic mechanism, by which migrants go for self-employment face to the difficulties they met in corporate careers (Kerr and Mandorff 2015).

Based on ethnic analysis of names and surnames of inventors listed on EPO data (patent applications at the European Patent Office), Breschi et al. 2017 test whether foreign inventors in the US and in several EU countries have outstanding performances. The latter are defined either as a higher-than-average probability of being ranked among the top 5% most prolific inventors, or a higher-than-average h-index, as computed on the forward citations to one's own patents, and controlling for cohort effects (supposed migration year).

Breschi et al. 2015 find similar results for the US and Europe as a whole, in line with the hypothesis of foreign inventors being over-represented among the most prolific and the most cited.

A second way migrant S&E can contribute to innovation in the destination country is by generating net positive externalities, as long as they compensate crowding out of native peers . The main source of externality investigated by the literature has been cognitive diversity and cultural variety. Most studies in the field come from economic geography, and build upon Jane Jacobs'view of the importance of variety for innovation, and of cities

[2] In a related paper, Stuen et al. 2012 examine the impact of foreign-born (by origin country) vs. native students on the scientific publications (number and citations received) by 2300 US university departments. Foreign-born and local students are found to impact similarly on the publication activity (and quality) of their departments, which goes in the direction of suggesting their substitutability.

as key providers of such variety, including ethnic one[3].

These studies generally follow an innovation production function approach and relate some productivity or innovation output measure for cities or regions to either the share of migrant population or a measure of population diversity. Some emphasis is placed on the research for causal links, with the choice of several instruments for explaining the location decision of migrants. A joint reading of this literature reveals how complex the notion of diversity can be, both at the methodological and conceptual level (we come back on this point in section 3). The most popular measure of population diversity is the fractionalization index (the reciprocal of an Herfindhal index, based on ethnic or nationality shares of the population). We find an application of this measure in Ottaviano and Peri 2006 and Bellini et al. 2013, who examine respectively US cities and NUTS3 European regions, with no specific reference to highly skilled migration, but to migration in general. They find a positive and causal link running from diversity to productivity, through a system of equations with income and prices as the dependent variables. Ozgen et al. 2011 investigate the patent output per capita in Nuts 2 regions in Europe (based on a two-period pooled cross-section) as a function of the diversity of nationalities in each region. In the absence of direct information on the skill level of migrants, the authors proxy it by distinguishing between migrants from advanced and less advanced macro-regions of origin. They find that more diversity leads to more innovation, and so it does any increase in the average skill level of migrants. In addition, the authors provides evidence of a strong correlation between the fractionalization index (computed including natives) and the share of foreign residents. Nijkamp and Poot (2015) expand upon this theme and provide a thorough methodological discussion of diversity measures. Niebuhr 2010 finds a positive relationship between the patenting rate of 95 German regions over two years (1995 and 1997) and the diversity of R&D staff of local companies[4]. The results holds for several alternative diversity indexes. Bosetti et al. (2015) obtain similar results with a study at the country level, for 20 European countries, which exploit information on the education level and nationality of workers, as contained in the EU Labour Force Survey.

Nathan 2014a departs from this literature by focusing on the individual performance of inventors, based on a dataset of UK-resident inventors, whose ethnic origin is identified by means of linguistic analysis of names and surnames. A positive association is found between individual patent productivity and the ethnic diversity of inventors in the same

[3] Two further sets of studies focus respectively on either ethnic variety and macro-economic performance, and the related problems of building consensual political institutions (for a survey: Alesina and La Ferrara 2005), or on ethnic variety of employees, managers or teams in organizations, and the associated trade-off of cognitive variety and communications costs (for a representative study: Earley and Mosakowski 2000; for a survey: Kirkman et al. 2006). For a firm-level analysis on the diversity-innovation relationship see also Parrotta et al. 2014

[4]Nijkamp and Poot 2015 extensively discuss diversity indexes, stating that there is no consensus in literature upon a "best" solution. Any alternative index suffers of its own drawbacks, and it is suited only for specific purposes. Plenty of room still exists for each scholar to propose an ad hoc index, suited to his/her research objectives.

travel-to-work area, as measured with a fractionalization index. In addition, by means of a two-step approach, the author estimates whether ethnic inventors add further to the UK innovation potential due to self-selection, and find moderately positive results for "star" inventors.

One limitation of the diversity-innovation literature is the absence of modeling and information on the interactions between highly skilled natives and migrants, and between migrants from the same or different groups. Studies that rely on labour force or migration official statistics do not have the means to examine to what extent workers in a given area, however mixed, interact across ethnic groups, or on the contrary operate in a segregation regime. As for Nathan 2014a, he does not examine the composition of inventor teams, nor other information on social interactions one can get from patent data.

Theory-wise, social connections between inventors with a migrant status are of particular interest. A vast literature from both migration and entrepreneurship studies exists, which points out how such connections are leveraged to overcome limited access to financial capital and selective professions, or other forms of exclusion at destination (Kerr and Mandorff 2015; Waldinger 1986). Sensitive information on prices and business contacts is exchanged among members of the same community, which originate distinctive specializations. The latter have been historically visible in low-end service and manufacturing activities, such as retail, catering, or textiles. As for highly-skilled, innovation-related migrants, two research questions can be put forward and will lead our analysis:

Q1 Do migrant S&E, like lower skilled migrants, also activate or rely upon community-mediated mechanisms of collaboration and exchange?

Q2 If so, do such mechanisms simply compensate for limited access to other assets (thus resulting in undistinguished or possibly inferior performances)? Or do they pave the way for superior performances?

Some evidence concerning these questions is provided by Freeman and Huang 2015, who examine the ethnic composition of US-based scientific teams, as reported on their publications from the 1980s onward. Authors are classified according to nine, very broad ethnic groups, as defined by Kerr 2008 for inventors. For each paper, the authors calculate an homophily index, which is best described as the reciprocal of the fractionalization index used in diversity studies, and include natives as one of groups. A negative correlation is found between homophily and: (i) the productivity of first and last authors on the paper; (ii) two citation-based quality measures of the papers. Although little is offered in order of causality analysis, these results suggest that over-reliance on co-ethnic co-authors is associated to weaker scientific status and may lead to weaker scientific performances. Almeida et al. 2014, for a sample of Indian semiconductor engineers, similarly find a negative association between reliance of ethnic networks (as measured by patent citations to co-ethnic inventors) and inventive performance. In contrast, Borjas et al. 2017 find that ethnic Chinese supervisors of Chinese PhD students at US universities increase their productivity (albeit at the cost of displacing their colleagues who do not have access to the same network). More indirect evidence on the role of ethnic ties among migrant S&E is provided by the analysis of patent citations conducted, among others, by Agrawal et al. 2008 and Breschi et al. 2017. The latter, in particular, find that US-resident inventors of foreign origin (where the origin is deduced by name analysis) have a higher-than-expected propensity to cite patents by same-origin inventors. The result, however, hold only for certain ethnic groups (most notable, East Asians and Russians, as opposed to Western Europeans) and it does not tell us anything on whether the same-origin preference witness of a preferential access to superior knowledge, or it reflects some form of segregation.

Last, we feel the necessity to point out that team diversity constitutes a significant excerpt of management literature: often the focus is either on firms as a whole, or on the composition of Top Management Teams (TMT). Most of the literature builds on the original idea developed in Hambrick and Mason 1984, who asserted that TMT characteristics can be used to unravel corporate strategies.

Management scholars have tackled the consequences of several definitions of team heterogeneity. The seminal contributions of Hambrick et al. 1996 provides clear evidence on the positive effects of educational and functional background diversity on their propensity to react; in a similar fashion Bantel and Jackson 1989 analyze the top management composition in a sample of 199 banks, finding again that heterogeneity in terms of age, tenure and functional background is associated to higher degrees of innovation. The nexus between diversity and innovative behavior. Bassett-Jones 2005 claim the existence of a dual mechanism: if on one hand diversity can foster creativity and innovation, and be in turn associated with higher performances, at the same time it can induce misunderstandings, low morale and loss of competitiveness.

Most recent research, in line with contemporary political and sociological debates, has focused on specific demographic traits such as nationality, ethnicity, and gender-related diversity (Nielsen and Nielsen 2013, Baixauli-Soler et al. 2015, Richard et al. 2013, Nielsen and Nielsen 2011 and Carter et al. 2010)

1.3 Data and descriptives

1.3.1 Data

One of the challenges faced in exploring migratory patterns of highly-skilled workers is the lack of reliable micro-information. Census-based data sources, such as National Labour Force surveys[5] or DIOC-OECD[6] provide only aggregate information, and not a very fine

[5]Eurostat EU LFS

[6]OECD-DIOC database

one in terms of educational/occupational levels. Social security data are the most detailed microdata, but they mainly come on a national basis.

As far as studies on migration and innovation are concerned, academic publications and/or patent records provide a solid alternative. The latter contains detailed information on scientists and technologists such as names, country of residence, and field of expertise. Based on country of residence, name analysis can be applied to assign a scientist or inventor a "local" or "migrant" status, as described by Lakha et al. 2011, Kerr 2008 and Breschi et al. 2017.

Name analysis methods, however,still suffer from some limitations, mostly related to the obstacle of discerning recent migrants from second-generation ones and ethnic minorities, and the difficulty to distinguish between locals and migrants that share the same language (as with British in the United States, or Spanish in South America).

Concerning patents, an alternative to name analysis exists, one that comes however at the price of losing potential observations. It consists in exploiting the information on inventors' nationality to be found in patent applications filed according to the Patent Cooperation Treaty (PCT) procedure[7]. The PCT procedure (as opposed to a national one) allows applicants to benefit from an harmonized prior-art search procedure and a longer delay for international extension (that is, it buys time before filing for protection in several national offices).

Up to 2011, a combination of legal requirements from both the World Intellectual Property Organization (WIPO, which administers the PCT) and the US Patents & Trademarks Office (USPTO), made it compulsory for inventors filing in the US to declare their nationality (Miguelez et al. 2013).

For information on the contents of patent applications in our samples we make use of the CRIOS-Patstat database Coffano and Tarasconi 2014, which constitutes the primary source for the construction of most of the variables described in Section 5. This choice came quite naturally, as it provides stable data on European patents that are also strongly consistent and harmonized across years; moreover it constitutes a strong reference for many previous EPO-centered empirical works. Another advantage in using the CRIOS-Patstat database is that its inventor records are disambiguated, according to the procedure described in Pezzoni et al. 2014.

In order to exploit PCT information, we first retain all the EPO patents with a corresponding PCT application, and with priority date comprised 1990 and 2010. We justify this choice as follows. PCT membership has grown over time; this could obviously result in a bias toward inventors coming from or residing in one of the early member countries. In particular, before 1990 the PCT member states were just a few but they dramatically increase after the introduction of TRIPs (the Trade Related IP Agreements, Maskus

[7]The Patent Cooperation Treaty was originally signed, and became active in 1978 for 18 states. Nowadays its adherents count 148 countries.

2000). We then retain only the patent applications whose entire team of inventors was resident, at the priority date, either in the United States or in a country of the European Union (EU15)[8]. We also drop all the patents for which some of the inventors were left with unassigned nationality[9].

In the reminder of the paper we will always treat the two samples separately. The main reason to do so concerns the different levels and composition of migratory flows in the two areas. As discussed in Section 2, the US have attracted many more highly skilled migrants than the EU, especially from East Asia. At the same time, highly skilled migration in the EU15 has been more limited and dominated by intra-regional flows.

Finally we draw on the following, secondary data sources.

First, for producing the weighted diversity index in Section 4, we made use of the CEEPI database on pairwise country relationships, which provides, among others, a new measure of linguistic proximity for 90 country pairs, based on the lexical similarity scores between 40-item word lists of different languages[10]. Furthermore, for the purpose of instrumental variable estimation in our regression exercises, we will exploit population data from two different sources. For the EU15 sample we use the OECD database on working age population[11], that will be integrated with the International Migration Database[12]. For US states, the same information is derived from the IPUMS databas[13].

1.3.2 An overview of foreign inventors in the US and EU15

Our US sample consists of 165,085 applications and 347,284 inventors, 17.76% of which are migrants (see Table A1 in Appendix A). Table A2 in Appendix A provides a breakdown of this figure by major migrant groups. Chinese and Indian inventors stand out with respectively 22.7% and 18.4% of all migrant inventors to the US, followed by Canadian, British and German nationals with respectively 9.6%, 8.7% and 4.8%. South Korea, France, Rus-

[8]The country of residence of inventors is derived from the latter's addresses as reported on the patent documents. Most often, patent applicants indicate the inventor's home address, which is then a rather accurate source of information for our purposes. However, it may occur that applicants report instead the address of their headquarters or of a controlled entity in charge of IP, which may even be located in a different country than the inventors'. This would create a type I error in our data, with local inventors possibly mistaken for migrant ones, due to the substitution of their home address with that of the US or EU applicant's headquarters. For these reasons, we examined in some detail the inventors' addresses and corrected or dropped problematic records when necessary, as described in the Appendix.

[9]Moreover during the matching phase of inventors to EPO applications, we find that in some cases we have totally or partially missing record, that is for some patents we can identify the nationality of just a subset of the inventors listed in the document. We choose to drop those observations. Additionally, we observe that some inventors were assigned to multiple nationalities across different applications. We decided to perform a simple name analysis on these inventors via the IBM-GNR database, taking the nationality more coherent with the ethnic origin of the inventor's name.

[10]http://www.cepii.fr/cepii/en/bdd_modele/bdd.asp. For an extensive overview of methodologies, see Melitz and Toubal 2014

[11]https://data.oecd.org/pop/population.htm

[12]https://stats.oecd.org/Index.aspx?DataSetCode=MIG

[13]https://usa.ipums.org/usa/

sia and Japan account for in between 2% and 3% each; Australia, Netherlands, Israel and Italy for about 1% each. Figure A1 shows that while the presence of both Chinese and Indian inventors became noticeable in the mid-1990s, it boomed since 2000. Four major innovation hubs in the US (namely Silicon Valley, New York, San Diego and Boston) account for 50% of migrant inventors (see Table A3 in Appendix A). Migrant inventors mostly appear on mixed teams, that is inventor teams comprising both natives and migrants, which in turn tend to be rather large teams, with 4 inventors or more (Table A5 in Appendix A). As for technological classes, migrant inventors are relatively more numerous in Chemistry (including Pharma and Biotech) and Electrical engineering, as opposed to Instruments, Mechanical engineering and all the other patent classes. As for the EU15 sample, this includes 271,538 and 472,687 inventors. Notice that, if not otherwise stated, we define as migrant inventor to EU15 only the inventors whose nationality is not one among the EU15 ones (for example, we consider as migrant an Indian inventor residing in the Netherlands; but not a German one). This reduces considerably the percentage of migrant inventors, to approximately 3% (Table **??** in Appendix A). When extending the definition of migrants to all non-nationals of the country of residence, the percentage rises to 8%. The main non-EU15 group of inventors is that of US nationals, followed by Chinese ones (13% and 12% respectively, out of all non-EU15 nationals). Indians and Russians follow at around 7%. Other groups accounting for more than 2% include both intra-European (extra-EU15) migrants (such as the Polish, Swiss and Romanians) as well as extra-Europeans (such as Australians and Canadians). The dominant hubs for foreign presence are Eindhoven (which host the headquarters and main R&D labs of Phillips), London, and Paris. However, the foreign presence is much less concentrated than in the US, as it takes 10 top hubs (instead of 4 as in the US) to account 50% of foreign inventors (table **??**; but see Appendix A for technical remarks on geo-localization in Eindhoven). The distribution of foreign inventors by size of the team and by technology appear similar to those for the US. When including intra-EU15 nationals among the migrants (that is, when counting Germans in the Netherlands as migrants), it is the latter who account for the largest country groups in the sample, with Germans accounting for around 15% of all migrants followed by the French (9%), the Dutch and the British (both at around 8%). US and Chinese nationals come only after the Italians, who stand at 5.5%.

1.4 Diversity test: nationality and inventor teams' composition

We answer Q1, as put forward in Section 2, by testing whether the observed diversity in patent teams may originate from inventors'homophily, defined here as the tendency of inventors of the same nationality to collaborate more among themselves than with local inventors or migrants from different countries. Specifically, we test the null hypothesis of no homophily by which the observed diversity distribution would simply result from random associations between individuals, irrespective of nationality. In order to do so, we first produce two measures of team diversity for our data. For both measures, we then compare the observed distribution of diversity with a random generated distribution (benchmark distribution). The test is performed separately for the US and for the top four EU15 countries in terms of foreign inventors activity (i.e. Germany, the Netherlands, United Kingdom and France).

We calculate diversity by means of two alternative indexes. First, for each application, we compute a standard fractionalization index (as employed by several studies surveyed in Section 2). For patent j we have:

$$diversity_j = 1 - \sum_{m=1}^{M} s_m^2 \tag{1.1}$$

where s_m is the share of inventors of nationality m listed in application j. Albeit easy to handle, the index does not exploit any other data than s[14].

In particular, *diversity* returns the same value whatever the specific countries of origin of the inventors, regardless of whether inventors are locals or migrants, and of whether they move from countries with high or low migration rates to the US or EU15. Furthermore it ignores that some ethnic combinations may be more likely than others to appear due to sociological factors, such as when the association is made easier by linguistic, religious or cultural proximity. To put it simply, think of a US patent with two inventors, a local and a Korean one. Compare this to the case of another US patent with two inventors, one from Korea and one from Thailand. In both cases diversity will take the value 0.5, but while in the former case the team is a rather natural combination of a local inventor and an inventor from a large community of high-skilled migrants, in the second one the combination is a much rarer one, and we could possibly think that such social connection is partly induced by the cultural proximity shared by Korea and Thailand.

To mitigate at least to some extent these distortive effects we experiment with the following measure:

$$weighted\ diversity_j = 1 - \sum_{m=1}^{M} \sum_{k=1}^{K} s_m s_k \gamma_{mk} \tag{1.2}$$

where γ_{mk} is the m-k element of a cultural distance matrix Γ, and with m and k being two generic countries[15]. We build Γ by making use of data on linguistic proximity

[14] Nijkamp and Poot 2015 discuss extensively a number of diversity indexes, stating that there is no consensus in the literature on the best one. Each index suffers of its own drawbacks, and it is suited only for specific purposes. Plenty of room still exists for proposing ad hoc indexes, suited to different research objectives.

[15] The formulation in (2) was originally proposed by Desmet et al. 2009

between country pairs provided by CEEPI, which was originally conceived to examine the role of countries' linguistic barriers within the framework of bilateral trade flows. Since (2) requires a distance measure, we rearranged the original linguistic proximity scores and standardized them so as to get $\gamma_{mk} \in [0, 1]$ for each country-pair. We also assigned the value of 0 to γ_{mm}. The weighted index is slightly more stable than the original fractionalization index: while the latter is strongly correlated with the percentage of foreign inventors listed on the patent ($\sigma_{US} = 0.83$, $\sigma_{EU15} = 0.802$), for the weighted index the correlation drops to 0.1 in both samples. In order to show the different behavior of diversity and weighted diversity, Table 1.1 presents the values we obtain by applying the two indexes to three patent applications from the US sample. In all patents two inventors out of three are from the United States, while the third is either Indian, Chinese or British. We can see that while diversity is always at 0.44 (which is the result of 1 - $(0.6^2 + 0.3^2)$), *weighted diversity* varies depending on the linguistic distances between countries. Patents signed by American and British inventors shows almost 0 diversity while the one in which there's a Chinese inventor has a much higher value. The last patent, with two Americans and one Indian inventor exhibits an intermediate value.

Table 1.1: US sample, comparison of diversity indexes.

Application ID	diversity	weighted diversity	Inventors' nationality
16916575	0.44	0.000047	[US, US, GB]
16969477	0.44	0.004630	[US, US, CN]
16971459	0.44	0.000201	[US, US, IN]

Notes: The table reports the diversity and weighted diversity index for three different patents (Application ID), with the same number of inventors but different nationalities.

Diversity is heavily concentrated towards a few distinct values that represent precise team configurations. While 0 indicates patents with no foreign inventors, the value of 0.5 is associated with applications in which two nationalities are equally represented. The values of the weighted diversity are more dispersed, but they also are concentrated close to 0. Apart from the high skewness of the two distributions, a codifiability issue applies to both our indexes. These values are in fact hard to interpret, especially within an econometric framework. Yet, they lend themselves to our simple test of homophily, based on benchmarking.

Benchmark distributions are generated in several steps: we first create five samples of inventors, one for each country we consider (Germany, the Netherlands, United Kingdom, France, and the United States). Second, we break down the distribution of nationalities in each sample, per priority year of the patent. Third, for each year we randomly sample (one at a time and without replacement) the nationalities from the correspondent distribution, and we assign them sequentially to the inventors in our database. The sampling takes place so that each element of the distribution of nationalities has equal probability of being picked (simple random sampling). We iterate the procedure 500 times, and each

Table 1.2: Distribution comparison of diversity indexes

Country	Mean difference	p-value
Great Britain	-32.577	0.0
Netherlands	-85.322	0.0
Germany	-18.301	0.0
France	-22.636	0.0
United States	-85.710	0.0

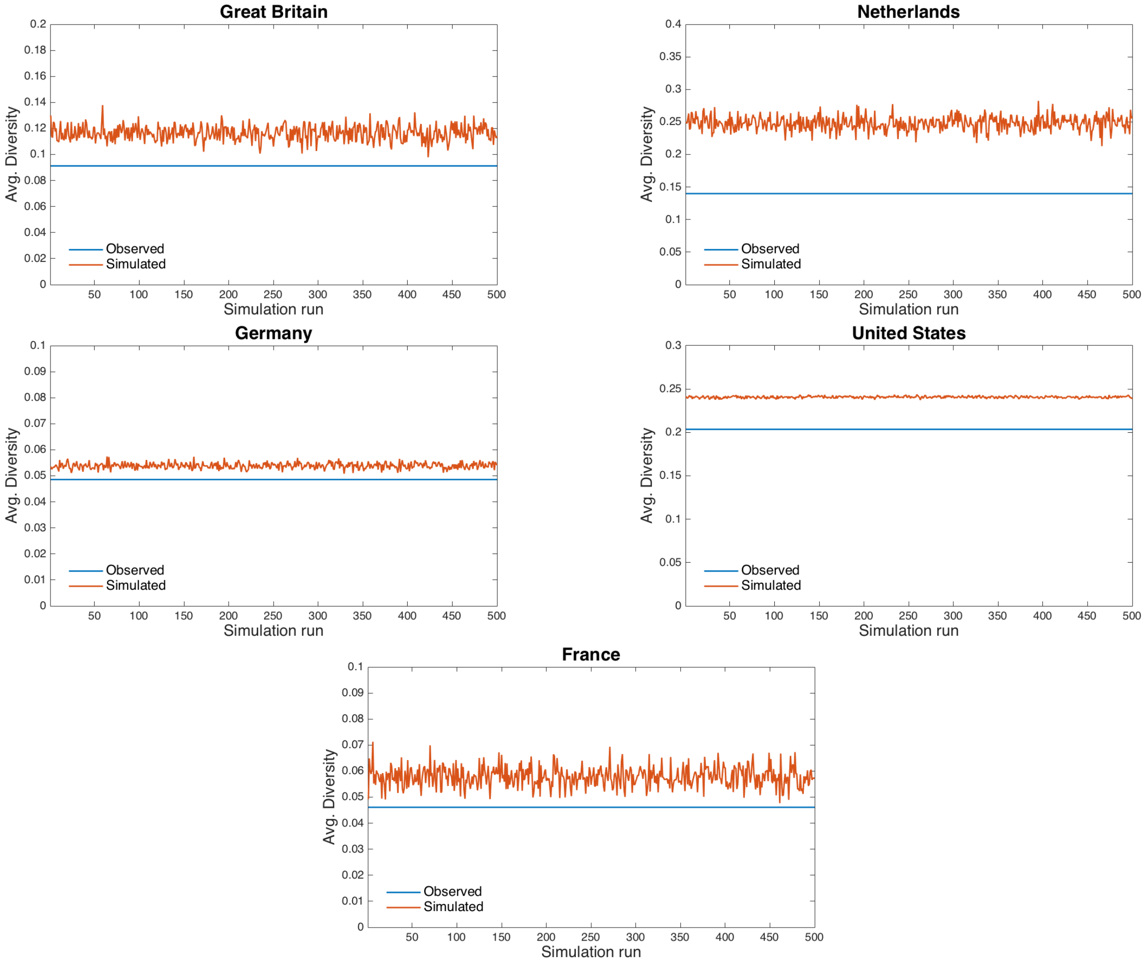
Notes: for five country sample, we report the difference between means of the distribution of the diversity index over all patents in the sample. The p-value refers to the null hypothesis of the difference being equal to zero and it is computed according to the procedure in Section 4.

time we calculate a *diversity* index, which is simply the mean of the randomly-generated *diversity* indexes for all patents. We similarly create benchmark distributions for *weighted diversity*.

We finally proceed to test the null hypothesis of random distribution of both diversity and weighted diversity, first through a graphical comparison of the observed and benchmark distributions of the two indexes. The results for diversity are shown in Figure 1.1 for each country and for each simulation run, the average values of the benchmark distributions are always higher than the observed ones. That is, the observed diversity is lower than what one would expect if inventors associated in teams irrespective of their nationality. This goes against the null hypothesis and in favor of the presence of homophily. We observe some variability across country, though. While the benchmark-observed difference is extremely small for Germany, France and Great Britain, it is much higher for the United States and the Netherlands, where we know the incidence of migration on total inventorship to be higher.

Coming to formal means testing, we need first to overcome some pitfalls originating from the non-normality of the distribution, together with unequal variances between observed and simulated data. In order to do so we exploit a bootstrap procedure to obtain confidence intervals, as suggested by Davison and Hinkley 1997. The process consists in iteratively re-sampling small portions of the two distributions and compare each time the relative t statistics. In Table 1.2 we report the average difference between observed and simulated t-statistics, computed after 10k replications, and their corresponding p-values, for the *diversity* index. As expected, the mean difference is negative and highly significant: the observed mean diversity is lower than the randomly generated one. The intensities of such gaps are in line with the ones of the graphic plots: the difference is quite higher for the United States and for the Netherlands.

Figure 1.1: Graphical diversity test.



Notes: The simulations compare average values of the benchmark distributions to the observed ones. Red lines correspond to simulation outcomes. Blue lines refer to observed diversity values.

1.5 Econometric analysis: migration, diversity and patent quality

1.5.1 Methodology

We answer Q2, as put forward in Section 2, by regressing of the observed quality of patents on alternative measures of diversity of nationalities in the team of inventors. Our data are a pooled cross-section with patents as observations. We test the robustness of our results to endogeneity issues by means of IV analysis.

Our baseline model is as follows:

$$E(citations_i) = \beta_0 + \beta_1 diversity_i + \gamma C_i + \delta f_i + \varepsilon_i \tag{1.3}$$

where β_1 captures the marginal effect of the diversity (or, alternatively, weighted diversity) in the inventor team of patent i on the patent's quality; and f is a vector of year and technology fixed effects. (C refers to a set of controls upon which we come back below). We proxy the quality of patents with the count of forward citations at the DOCDB family level, excluding self-citations within the same family. While there exists some mixed evidence about the use of patent citations as a measure of knowledge diffusion (Alcacer and Gittelman 2006), numerous contributions in the patent literature stress the capability of such measure to predict the technological importance and/or economic value of a patent (Jaffe and De Rassenfosse 2016, Trajtenberg 1990).

We calculate citations over a three- and a five-year window after its priority date (cit_3years, cit_5years), as well as for the whole time span up to 2010, irrespective of the priority date (cit_all). The different specifications of the dependent variable serve the purpose of dealing with right-truncation issues by experimenting with different truncation points (Lanjouw and Schankerman 2004).

The standard model to deal with count data is the Poisson regression or, in case of overdispersion of the dependent variable, Negative Binomial regression (Scott Long 1997)[16]. However, we prefer relying on OLS estimation, after transforming the citation count into the continuous variable log(1+citations).

This is a widespread approach, which allows for direct interpretation of coefficients as marginal effects (semi-elasticities), particularly in the case of two-stage regressions with instruments (as examples, see Harhoff et al. 1999 and Lissoni et al. 2011, among others).

[16] First, one of the basic assumptions of Poisson regressions is that events take place independently of time. On the contrary we expect some kind of time dependence in patent citations: the probability of a patent receiving an additional citation may differ depending on having already been cited in previous periods. The second assumption is that the mean of the predicted variable equals its variance. In the case of forward patent citations, however, the variance is always very much higher than the mean. This holds also for the patents in our sample.

In any case, we also run Poisson and Negative Binomial regressions for the baseline model, whose results are reported in Appendix D and do not differ much from OLS ones.

Commenting further on Eq. (3) it is important to highlight one important limitations of both *diversity* and *weighted diversity* when it comes to quantifying the contribution to innovation directly ascribed to migratory events, as they do not isolate diversity coming from a mix of natives and migrants, and diversity coming teams entirely composed of migrants. Consider for example two US-based patents, signed by three inventors each: the first team is made of three American inventors, while the second patent is developed by three Chinese. For both patents *diversity* (and *weighted diversity*) will be equal to zero, but any alternative index meant to capture the migrant status of inventors would clearly signal a difference between them. In particular, with reference to research question Q2, we are interested in testing whether the zero-diversity, all-Chinese patent is of lower quality than the zero-diversity, all-natives patents, as a reflection of a more limited, quality-constraining choice of team partners for migrant inventors, due to lack of social capital. This clearly calls for an alternative measure to both diversity and weighted diversity, one that may enable us to isolate the contribution to innovation made by foreign inventors, irrespective of team diversity.

Hence, we propose a second specification of the model, where *diversity* is replaced by a set of dummies, as follows:

$$\begin{aligned} E(citations_i) = \beta_0 &+ \beta_1 foreign\ same + \beta_2 foreign\ mix + \\ &+ \beta_3 locals\ foreign\ same + \beta_4 locals\ foreign\ mix + \\ &+ \gamma C_i + \delta f_i + \varepsilon_i \end{aligned} \tag{1.4}$$

where the baseline case for the dummy set is that of a patent signed exclusively by local inventors, and:

- *foreign same* equals one if the inventor team is entirely made of foreigners, all of whom with the same nationality
- *foreign mix* equals one for teams entirely made of foreign inventors, with at least two of them having different nationalities
- *locals foreign same* equals one if the team includes both local and foreign inventors, with all the foreign ones having the same nationality
- *locals foreign mix* equals one if the team includes both local and foreign inventors, with at least two foreign inventors having different nationalities

The categories of patents described by our dummies are mutually exclusive, meaning that each patent cannot be assigned to more than one of them. Each dummy captures a different mix of diversity and the migrant status of the inventors in the patent's team. The coefficients β_1 through β_4 will provide an estimate of the response of patent quality to various team structures, taking as reference the patents signed solely by local inventors.

Coming finally to the set of controls in C, which is the same in both Eq. 3 and 4, this includes several patent characteristics that could induce variations in quality, as measured by forward citations. Some of such controls, if omitted, could induce biased estimates for the coefficient of diversity, being possibly correlated to the latter:

- *Team size*: the number of inventors who signed the patent. We know from the literature that patent citations are positively correlated with the number of inventors involved as collaboration can boost both the quality of inventive out- put together with the speed at which it is achieved (Simonton 1988, *guimera2005team*)

- *Team experience*: a measure of inventors' performance. A deeper knowl- edge of the field and a complete understanding of the dynamics involved in patenting activities are likely to increase the quality of inventions Schettino et al. 2013. We calculate the number of patents signed by each inventor on the focal patent, and take the maximum value. Alternatively we use the average level of previous patenting activity and we obtain very similar results.

- *Family size*: the total number of patents that form the focal patent's DOCDB family. We know that such measure depends chiefly on the number of patent offices in which the applicant seeks for protection, with increasing costs in terms of both filing fees, legal charges, and documentation expenses (e.g. translations). A large patent family both signals the quality of the protected invention (which must be worth the extension costs) and increases the probability of citations, ceteris paribus, due to the multiplication of related patent documents.

- *Previous collaborations* is a dummy which equals one if at least two of the inventors listed on the application have been previously engaged in patenting activities together. This may both indicate the accumulation of social capital and signal the existence of a successful partnership, both expected to be positively associated to patent quality.

- *University* is a dummy equal to one if the invention involves an academic research center. In terms of quality, patents owned by European universities receive on average fewer citations, especially in the first few years after filing, while the opposite holds true for the US (Sterzi 2013, Bacchiocchi and Montobbio 2009).

- *Triadic*: triadic patents are defined as filed at the United States Patent and Trademark Office (USPTO), the European Patent Office (EPO), and the Japanese Patent Office (JPO). Recent literature suggest that, with a few exceptions, these applications show a higher citations impact with respect to ordinary patents (Sternitzke 2009).

- *NUTS3 regions*: a dummy that equals one whenever the inventors reside in different NUTS3 regions. This dummy, which is a proxy for geographical distance, signals the absence of frequent face-to-face interactions between inventors.

Finally, in order to make our results at least partially robust to various sources of bias, namely reverse causality and measurement errors, we instrument our focal regressors with the share of foreign nationals over the total population, as well as with the level of diversity among foreign residents. Intuitively, both the foreign presence and the diversity levels in teams of inventors should be correlated with the ones of the environment in which they operate. More in detail, we make use of the following instruments[17]:

- *foreign share*: the percentage of foreign born individuals over the total population.

- *country diversity*: within the foreign population, we compute the diversity index described above.

The instruments are employed in a classical 2SLS framework, and both of them enter the regressions with a one-year lag with respect to the patent priority date t.

1.5.2 Main results

The results for estimations of Eq. 3 are presented in Table 1.5. Columns (1) to (3) reports our results for the US sample, respectively for specification with dependent variables log(1+cit_3years), log(1+cit_5years) an log(1+cit_all). Columns (4) to (6) refer to the EU15 sample, for the same dependent variables.

We find that the coefficients of diversity are always positive and significant, regardless of the dependent variable of choice. A decimal increase in diversity is associated with a 1.7% increase in three-years forward citations for the US. The percentage increase ranges from 2.2% to 2.5% for the EU15 sample.

The coefficients of control variables are all significant, and in line with previous contributions. Ownership by a academic institution/university results in an opposite effect on forward citations: negative for the EU15 sample (as in Sterzi 2013)and slightly positive for the US one.

Results are similar for weighted diversity (Table 1.6): a decimal increase in the weighted index is associated to a percentage change in 3-years forward citations in between 1.7% and 1.8% for the US sample, and between 2.9% and 3.3% for the EU15 one.

[17]As anticipated, population data for European countries come from the OECD repository, while for the US we refer to the IPUMS database.

Table 1.3: Summary statistics by sample, EU15

Variable	Mean	Std. Dev.	Min.	Max.	N
diversity	0.066	0.164	0	0.833	246295
weighted diversity	0.034	0.102	0	0.765	246014
log(1+cit)	1.201	1.141	0	6.748	246295
log(1+cit5) years	0.947	0.904	0	5.864	246295
log(1+cit3) years	0.783	0.759	0	5.591	246295
team size	3.144	1.434	2	10	246295
university	0.038	0.192	0	1	246294
team experience	1.823	0.96	0.693	5.832	246295
previous collaborations	0.275	0.447	0	1	246295
triadic	0.271	0.444	0	1	246295
NUTS3 regions	1.935	1.048	0	10	246295
family size	6.505	4.497	1	54	246295

Table 1.4: Summary statistics by sample, US

Variable	Mean	Std. Dev.	Min.	Max.	N
diversity	0.189	0.245	0	0.864	163690
weighted diversity	0.126	0.191	0	0.815	163268
log(1+cit)	1.752	1.224	0	7.783	163690
log(1+cit5) years	1.485	1.038	0	6.529	163690
log(1+cit3) years	1.284	0.919	0	6.182	163690
team size	3.401	1.621	2	10	163690
university	0.076	0.265	0	1	163690
team experience	1.681	0.866	0.693	5.765	163690
previous collaborations	0.235	0.424	0	1	163690
triadic	0.35	0.477	0	1	163690
NUTS3 regions	1.962	1.005	0	9	163690
family size	6.678	4.711	1	53	163690

Table 1.5: OLS estimation, inventor team diversity and patent quality (Eq. 3).

	(1) US sample log(1+cit_3)	(2) US sample log(1+cit_5)	(3) US sample log(1+cit)	(4) EU sample log(1+cit_3)	(5) EU sample log(1+cit_5)	(6) EU sample log(1+cit)
diversity	0.170***	0.180***	0.175***	0.219***	0.242***	0.252***
	(0.00910)	(0.00996)	(0.0112)	(0.00943)	(0.0108)	(0.0128)
team experience	0.00305***	0.00248***	0.00148***	0.00175***	0.00169***	0.00126***
	(0.000217)	(0.000225)	(0.000239)	(0.000102)	(0.000115)	(0.000138)
family size	0.0258***	0.0294***	0.0319***	0.0249***	0.0297***	0.0356***
	(0.000530)	(0.000582)	(0.000655)	(0.000417)	(0.000480)	(0.000583)
patent scope	0.0313***	0.0328***	0.0258***	0.0462***	0.0530***	0.0608***
	(0.00268)	(0.00298)	(0.00345)	(0.00225)	(0.00263)	(0.00329)
university	0.00705	0.0263**	0.0341***	-0.124***	-0.124***	-0.118***
	(0.00787)	(0.00878)	(0.00997)	(0.00691)	(0.00818)	(0.00991)
triadic	0.277***	0.320***	0.374***	0.235***	0.314***	0.436***
	(0.00534)	(0.00597)	(0.00684)	(0.00409)	(0.00489)	(0.00619)
prev. collaborations	0.0214***	0.0192**	0.0121	0.0356***	0.0381***	0.0398***
	(0.00549)	(0.00600)	(0.00671)	(0.00366)	(0.00426)	(0.00518)
team size	0.0519***	0.0579***	0.0622***	0.0324***	0.0375***	0.0413***
	(0.00158)	(0.00172)	(0.00191)	(0.00131)	(0.00151)	(0.00181)
NUTS3 regions	-0.0150***	-0.0162***	-0.0162***	-0.00919***	-0.00933***	-0.00323
	(0.00242)	(0.00265)	(0.00296)	(0.00171)	(0.00198)	(0.00241)
Constant	0.519***	0.745***	1.680***	0.187***	0.288***	0.744***
	(0.0320)	(0.0380)	(0.0582)	(0.0134)	(0.0169)	(0.0258)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech. dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	163572	163572	163572	246056	246056	246056
R^2	0.180	0.218	0.275	0.139	0.163	0.199
F	562.2	798.6	1097.6	590.1	828.2	1147.0

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.6: OLS estimation, inventor team weighted diversity and patent quality (Eq. 3).

	(1) US sample log(1+cit_3)	(2) US sample log(1+cit_5)	(3) US sample log(1+cit)	(4) EU sample log(1+cit_3)	(5) EU sample log(1+cit_5)	(6) EU sample log(1+cit)
weighted diversity	0.173***	0.181***	0.171***	0.290***	0.317***	0.338***
	(0.0117)	(0.0127)	(0.0142)	(0.0150)	(0.0172)	(0.0203)
team experience	0.00306***	0.00249***	0.00150***	0.00177***	0.00171***	0.00128***
	(0.000217)	(0.000226)	(0.000239)	(0.000102)	(0.000115)	(0.000138)
family size	0.0259***	0.0295***	0.0320***	0.0250***	0.0299***	0.0357***
	(0.000532)	(0.000584)	(0.000657)	(0.000417)	(0.000481)	(0.000584)
patent scope	0.0321***	0.0336***	0.0266***	0.0466***	0.0533***	0.0612***
	(0.00268)	(0.00299)	(0.00346)	(0.00225)	(0.00264)	(0.00329)
university	0.00630	0.0256**	0.0336***	-0.125***	-0.126***	-0.119***
	(0.00790)	(0.00881)	(0.0100)	(0.00693)	(0.00821)	(0.00993)
triadic	0.277***	0.319***	0.374***	0.235***	0.314***	0.437***
	(0.00535)	(0.00598)	(0.00685)	(0.00410)	(0.00490)	(0.00619)
prev. collaborations	0.0221***	0.0200***	0.0129	0.0349***	0.0373***	0.0389***
	(0.00550)	(0.00600)	(0.00672)	(0.00367)	(0.00426)	(0.00519)
team size	0.0531***	0.0592***	0.0636***	0.0335***	0.0388***	0.0426***
	(0.00158)	(0.00172)	(0.00191)	(0.00131)	(0.00151)	(0.00181)
NUTS3 regions	-0.0156***	-0.0169***	-0.0170***	-0.0105***	-0.0109***	-0.00479*
	(0.00243)	(0.00265)	(0.00297)	(0.00171)	(0.00198)	(0.00241)
Constant	0.597***	0.859***	1.832***	0.172***	0.253***	0.691***
	(0.0302)	(0.0362)	(0.0540)	(0.0135)	(0.0169)	(0.0264)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech. dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	163151	163151	163151	245776	245776	245776
R^2	0.180	0.217	0.275	0.138	0.162	0.199
F	558.6	794.3	1092.4	587.1	824.6	1143.3

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Overall, the regression results for equation (3) suggest that diversity within a team of inventors is positively associated to patent quality, regardless of the specific diversity index we use.

We next focus on regression results for equation (4), as per Table 1.7. All the coefficients for the foreign-and-mix dummy set are positive and significant, which indicates a positive association of patent quality both to diversity and to the migrant status of inventor. However, the coefficients for *foreign mix* and *locals foreign mix* are generally higher than those for for *foreign same* and *locals foreign same* (the coefficients for *foreign same* also are generally less significant, and not significant in column 6). This suggests teams hosting foreign inventors from different countries, whether along with local inventors or not, are associated to patents of higher quality than those hosting same-national foreign inventors. For instance, a US patent produced by a team composed entirely of foreign same-nationals inventors country is associated to a 5.6% increase in three-year forward citations with respect to the baseline case (*all locals*). When the team is composed of foreign inventors from different countries (*foreign mix*), the percentage increase goes up to 14% . The results are similar for teams including also local inventors: in this case the observed citations increase ranges from 6% (for *locals foreign same*) to 12.5% (for *locals foreign mix*). This relationship holds across the three definitions of patent quality, and for both the US and the EU15 sample.

Table 1.7: OLS estimation, foreign inventorship and patent quality (Eq. 4).

	(1) US sample log(1+cit_3)	(2) US sample log(1+cit_5)	(3) US sample log(1+cit)	(4) EU sample log(1+cit_3)	(5) EU sample log(1+cit_5)	(6) EU sample log(1+cit)
foreign same	0.0568**	0.0605**	0.0525*	0.0424**	0.0421*	0.0203
	(0.0193)	(0.0209)	(0.0234)	(0.0155)	(0.0178)	(0.0207)
foreign mix	0.140***	0.139***	0.125***	0.0834***	0.0932***	0.102***
	(0.0129)	(0.0138)	(0.0152)	(0.0158)	(0.0181)	(0.0209)
locals foreign same	0.0606***	0.0671***	0.0679***	0.0859***	0.0973***	0.100***
	(0.00494)	(0.00544)	(0.00613)	(0.00472)	(0.00545)	(0.00649)
locals foreign mix	0.125***	0.129***	0.121***	0.179***	0.185***	0.186***
	(0.00821)	(0.00888)	(0.00979)	(0.0117)	(0.0131)	(0.0149)
Constant	0.528***	0.753***	1.688***	0.195***	0.296***	0.752***
	(0.0321)	(0.0381)	(0.0582)	(0.0134)	(0.0169)	(0.0258)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech. dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	163572	163572	163572	246056	246056	246056
R^2	0.180	0.218	0.275	0.139	0.163	0.199
F	538.3	764.3	1050.1	564.4	791.9	1096.6

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

These results both confirm and extend those obtained with diversity indexes. They support the importance of diversity, and show it to be quantitatively more relevant than the mere presence, in a team, of foreign inventors. In other words, migrant inventors increase patent quality possibly because of self-selection (they are on average more skilled than native inventors) but also and especially by adding variety to inventor teams. This result requires interpretation, as it may be determined by two different mechanisms, one having once more to do with self-selection and sorting, the other with the distinctive effect of diversity on team's creativity. According to the first mechanism, migrant inventors with higher skills would find more opportunities to join more innovative companies or projects, thus being associated both to higher quality patents and more diverse teams. According to the second mechanism, on the contrary, the causal link would go from diversity to patent quality, regardless of the migrant inventors' skills. While deciding which of the two mechanisms prevails requires treating endogeneity in our regressions (an exercise we attempt below), we can observe right now that some ground exists for causality to run at least in part from diversity to patent quality. If self-selection and sorting were the exclusive mechanism in place, we would not observe the estimated coefficients for locals foreign mix and locals foreign same to differ, being self-selection and sorting certainly in place for both cases, but diversity occurring only in the latter.

1.5.3 Robustness checks and endogeneity issues

We first investigate whether our results may be driven by the presence of foreign inventors in a small number of very high quality patents. This robustness check is dictated by the very skewed distribution of forward citations, with many patents receiving zero or very few citations and a long right tail of highly cited ones. Hence we verify whether our baseline results hold for different portions of our database, by running the model in Eq. 3 on an increasingly smaller subset of highly cited patents (from top 25% to top 1%, with the top quantiles having been computed separately for each year-technological class combination).

Table 1.8 refers to the specification with *diversity* as the explanatory variable of interest. Our baseline result cease to hold in both the US and the EU15 samples for the Top 1% cited patents, while the coefficients remain positive and significant for the Top 25% subsample, although lower in absolute value. We get analogous results with *weighted diversity* (Table 1.9).

Next, we run separate regressions for five large technological classes (with each patent application being assigned to one and only one class). In Table 1.10 we observe that, for patents in the US sample, the only estimated coefficient for *diversity* that change drastically is the one for Instruments, which is way below the one estimated in Section 4.2, although it remains positive and significant. We find no differences across classes for the EU15 patents: within this sample, results are much more in line with the baseline estimation.

Table 1.8: OLS estimation for top cited patents. Dependent variable is log(1+cit3).

	Top 1% US sample	Top 10% US sample	Top 25% US sample	Top 1% EU sample	Top 10% EU sample	Top 25% EU sample
diversity	-0.0217	0.0215	0.0566***	0.0485	0.0752***	0.0970***
	(0.0267)	(0.0110)	(0.00878)	(0.0631)	(0.0185)	(0.0124)
Constant	3.238***	2.276***	1.789***	3.284***	2.059***	1.659***
	(0.0571)	(0.0363)	(0.0272)	(0.147)	(0.0290)	(0.0194)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech. dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	3285	26930	57798	850	12082	36213
R^2	0.643	0.438	0.321	0.632	0.491	0.354
F	100.1	302.8	399.3	35.35	180.9	306.6

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.9: OLS estimation for top cited patents. Dependent variable is log(1+cit3).

	Top 1% US sample	Top 10% US sample	Top 25% US sample	Top 1% EU sample	Top 10% EU sample	Top 25% EU sample
weighted diversity	-0.0509	-0.00244	0.0436***	-0.0375	0.0866**	0.135***
	(0.0340)	(0.0140)	(0.0112)	(0.100)	(0.0281)	(0.0193)
Constant	3.142***	2.278***	1.825***	3.453***	2.060***	1.707***
	(0.0778)	(0.0363)	(0.0321)	(0.0846)	(0.0290)	(0.0207)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech. dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	3263	26845	57608	848	12063	36157
R^2	0.643	0.439	0.321	0.632	0.491	0.353
F	99.37	302.4	397.7	34.92	180.7	306.7

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.10: OLS estimation of US patents by main technological class. Dependent variable is log(1+cit3).

	(1) Mech. Eng.	(2) Chemistry	(3) Instruments	(4) Elect. Eng.	(5) Others
diversity	0.198***	0.148***	0.0427*	0.233***	0.162***
	(0.0253)	(0.0126)	(0.0179)	(0.0175)	(0.0439)
Constant	0.302***	0.383***	0.430***	0.169***	0.128**
	(0.0676)	(0.0378)	(0.0663)	(0.0329)	(0.0495)
Controls	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes
Tech. dummies	Yes	Yes	Yes	Yes	Yes
N	25113	75167	45902	63121	17032
R^2	0.171	0.187	0.183	0.137	0.112
F	79.73	279.7	168.5	147.9	32.05

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.11: Baseline model (Eq. 3), OLS estimation of EU15 patents by main technological class. Dependent variable is log(1+cit3).

	(1) Mech. Eng.	(2) Chemistry	(3) Instruments	(4) Elect. Eng.	(5) Others
diversity	0.143***	0.216***	0.164***	0.233***	0.162***
	(0.0204)	(0.0140)	(0.0201)	(0.0175)	(0.0439)
Constant	0.141***	0.175***	0.119***	0.169***	0.128**
	(0.0231)	(0.0201)	(0.0312)	(0.0329)	(0.0495)
Controls	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes
Tech. dummies	Yes	Yes	Yes	Yes	Yes
N	73573	98870	50666	63121	17032
R^2	0.120	0.146	0.145	0.137	0.112
F	158.5	261.8	132.7	147.9	32.05

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The last, but most important issue that we tackle is a possible reverse causality bias in our estimates. Specifically, more international teams could represent the response of firms to their needs for more ambitious projects, from which they obtain higher quality patents. We can think of the extreme case in which a team of inventors from all over the world is assembled in order to develop an outstanding, breakthrough idea: in such a case, we would have a clear endogeneity problem. We choose to proceed by instrumenting the focal explanatory variable (either diversity of weighted diversity) and re-estimate the model of Eq. 3 through 2SLS. The exogenous instruments that we include are *foreign share* and *country diversity*, as described above. This comes at the cost of losing observations, due to limited time coverage of both the OECD and the IPUMS data. In particular, the OECD International Migration Database, for some of the EU15 countries such as Portugal, Austria and Ireland has limited coverage[18]. IPUMS data, on the other hand, are available only from 2000. Hence, we lose about seven thousand patents for the EU15 sample and fifty thousand for the US, that is about a third of the entire sample. Results in Tables 1.12 and 1.13 suggest that endogeneity issues, if any, may have caused us to underestimate, rather than overestimate, the marginal effect of *diversity* on patent quality. Similar comparisons can be drawn for all the other coefficients. We do not want to push ourselves too far and accept this value as the true one, but we can conclude that the positive effect from the baseline estimation holds against a reverse causality issue.

Table 1.12: 2SLS estimation. We instrument **diversity** with **foreign share (t-1)** and **country diversity (trend)**. US patents signed before 2000 are excluded due to missing data.

	(1) US sample	(2) US sample	(3) US sample	(4) EU sample	(5) EU sample	(6) EU sample
diversity	1.239***	1.311***	1.441***	1.869***	1.887***	1.811***
	(0.0754)	(0.0821)	(0.0898)	(0.291)	(0.335)	(0.403)
Constant	0.105***	0.0499**	0.0787***	-0.0977***	-0.167***	-0.227***
	(0.0141)	(0.0153)	(0.0167)	(0.0196)	(0.0225)	(0.0272)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	116590	116590	116590	239922	239922	239922
Log-likelihood	-147425.3	-157366.8	-167800.5	-272403.2	-306300.7	-350979.0
F	438.8	548.5	671.6	509.6	642.8	863.3

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[18] The first set of datapoints date back to 1996 for Austria, 1992 for Portugal and 1994 for Ireland.

Table 1.13: 2SLS estimation. We instrument **weighted diversity** with **foreign share (t-1)** and **country diversity (trend)**. US patents signed before 2000 are excluded due to missing data.

	(1) US sample	(2) US sample	(3) US sample	(4) EU sample	(5) EU sample	(6) EU sample
weighted diversity	1.842***	1.946***	2.135***	5.513***	5.565***	4.736***
	(0.117)	(0.127)	(0.139)	(1.039)	(1.159)	(1.310)
Constant	0.107***	0.0526***	0.0821***	-0.189***	-0.259***	-0.293***
	(0.0144)	(0.0156)	(0.0171)	(0.0400)	(0.0446)	(0.0504)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	116290	116290	116290	239644	239644	239644
Log-likelihood	-149296.0	-159089.3	-169532.6	-308776.3	-334974.3	-364192.0
F	420.7	527.1	645.1	374.9	504.1	769.9

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

1.6 Discussion and conclusions

In this paper we have investigated the impact of migration on innovation in destination countries, as proxied by the impact of migrant inventors on patent quality in booth the US and Europe (EU15 countries). Theoretical insights from the literature suggests that migrants can impact positively on innovation in destination countries either directly, as when they are positively self-selected and more skilled or productive than natives, or indirectly, by co-inventing along with natives and/or migrants from different countries than theirs. Our evidence points in the latter direction.

We first find that, as suggested by the literature discussing migrants'social capital, migrant inventors tend disproportionately to enter teams composed of fellow migrants with the same nationality, thus producing observed average values of diversity indexes that are inferior to the benchmark values one would otherwise obtain. Second, we find that the presence of migrants in an inventor team most increases patent quality when it also increases the team's diversity. This result is suggested both by an alternative econometric specification, one in which we accoung separately for the diversity of the inventor teams and the presence of migrants therein, and by instrumenting diversity with EU country-specific or US state-specific variables. Our findings on the role played by migration on team diversity complement early findings on the similar role played by migration in increasing diversity in cities and regions (Ozgen et al. 2011; Bellini et al. 2013; Nathan 2014a). They also contribute to the long-standing debate in the organization literature on the trade-off between coordination costs and creativity associated to increasing cultural diversity in teams. At the same time, our results appear distinctive insofar they are based on extensive microdata, more closely associated to the inventive act and innovation in general than anything found in the literature. They are more strictly connected to the highly skilled migration literature, especially for what concern migrants' self-selection and social capital. As for policy conclusions, while our evidence does not lend itself to provide any specific indications, it goes in the direction of favoring open border for highly skilled migrants, as well as recommending their rapid integration into local labour markets, and to ease their association with natives and migrants from other countries, so to obtain more diversity in teams. Plans for future research include enriching our datasets with biographical information on inventors, especially with reference to their qualifications, date of entry into destination countries, and temporary vs permanent migrant status. Based on such information, we expect to be produce more specific evidence on mechanisms underlying the association of inventors in more or less diverse teams, as well as better controlling for personal characteristics associated to self-selection.

Appendix A - Migrant inventors data

To perform our study we constructed two different samples including patent applications listing inventors all resident in the United States (US sample), or in one of the EU15 countries (EU15 sample).

The US sample consists of 163,690 applications, providing information about 347,284 inventors. We define an inventor as *migrant* if his/her nationality (as reported in the WIPO database) is not that of the United States. The EU15 sample includes 246,295 patent applications, covering 472,687 inventors. In this case, we consider an inventor as *migrant* if his/her nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

To account for intra-European migration flows we created an additional sample including patents whose inventors reside within the same EU15 country, considering as *migrants* all the individuals with a different nationality than their country of residence.

US Sample

As indicated in Table A1, 17.76% of US resident inventors are foreign-born, an immigrant community of 61,670 individuals.

Table A1: US Inventors by origin

	Inventors	Percent
Natives	285,614	82.24
Immigrants	61,670	17.76
Total	347,284	100.00

Figure A1 portrays the evolution of foreign inventors patenting activity in the US. Chinese nationals' activity started growing around the mid 1990s, experiencing a surge at the beginning of the 2000s, along with Indian natives. Perhaps spurred by a boost in foreign inventors inflow, this pattern might also be related to an increase in the overall quest for patent protection in the United States. Other major migrant groups are represented by Canadians, British and German inventors, which to a lesser extent, reported a similar boost around the same period of time.

Table A2 lists all the major US-resident immigrant inventors groups (Figure A2 provides a graphic representation of it). Chinese and Indian migrants constitute over 40% of all US foreign-born inventors, followed by Canadian, British and German nationals who account for 9.60%, 8.74% and 4.79% of total immigrant inventors respectively.

Table A3 provides an overview of US-resident foreign-born inventors geographical localisation, grouping them according to their Combined Statistical Area (CSA) of resi-

Figure A1: US Main Immigrant Inventors Groups Patenting Activity

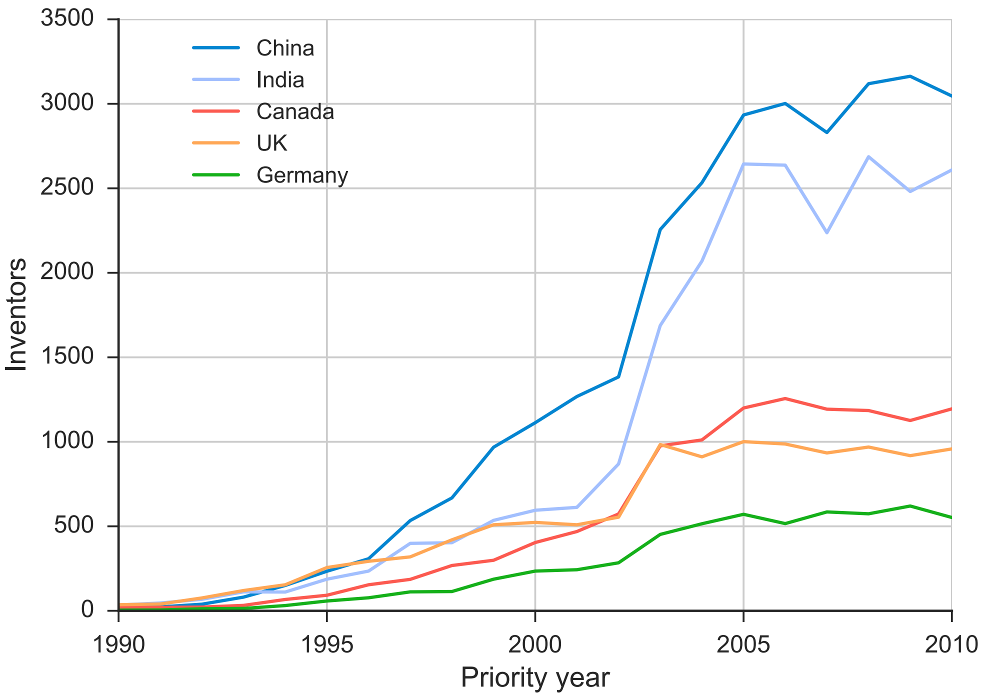


Figure A2: US Main Immigrant Inventors Groups

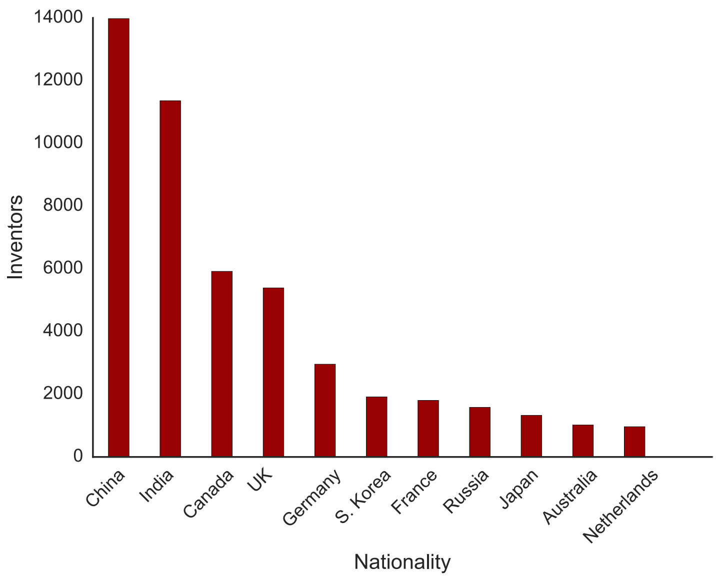


Table A2: US Main Immigrant Inventors Groups

Nationality	Inventors	Percent	Cumulative Percent	Percent on Resident Inventors
China	13,979	22.67	22.67	4.03
India	11,358	18.42	41.08	3.27
Canada	5,918	9.60	50.68	1.70
UK	5,391	8.74	59.42	1.55
Germany	2,957	4.79	64.22	0.85
South Korea	1,915	3.11	67.32	0.55
France	1,804	2.93	70.25	0.52
Russia	1,582	2.57	72.81	0.46
Japan	1,327	2.15	74.97	0.38
Australia	1018	1.65	76.62	0.29
Netherlands	964	1.56	78.18	0.28
Israel	835	1.35	79.53	0.24
Italy	806	1.31	80.84	0.23
Turkey	582	0.94	81.78	0.17
Spain	531	0.86	82.64	0.15
Other	10,703	17.36	100.00	3.08
Total	61,670	100.00	100.00	17.76

Table A3: US Immigrant Inventors by Main CSAs

Combined Statistical Area (CSA)	Inventors	Percent	Cumulative Percent
San Jose - San Francisco - Oakland (CA)	11,348	18.44	18.44
New York - Newark - Bridgeport (NY-NJ-CT-PA)	8,031	13.05	31.49
San Diego - Carlsbad - San Marcos (CA)	6,226	10.12	41.60
Boston - Worcester - Manchester (MA-NH)	5,507	8.95	50.55
Los Angeles - Long Beach - Riverside (CA)	2,859	4.65	55.20
Philadelphia - Camden - Vineland (PA-NJ-DE-MD)	2,205	3.58	58.78
Washington - Baltimore - N. Virginia (DC-MD-VA-WV)	1,950	3.17	61.95
Chicago - Naperville - Michigan City (IL-IN-WI)	1,827	2.97	64.92
Houston - Baytown - Huntsville (TX)	1,573	2.56	67.47
Detroit - Warren - Flint (MI)	1,315	2.14	69.61
Raleigh - Durham - Cary (NC)	1,222	1.99	71.59
Seattle - Tacoma - Olympia (WA)	1,162	1.89	73.48
Minneapolis - St. Paul - St. Cloud (MN-WI)	1,128	1.83	75.32
Dallas - Fort Worth (TX)	888	1.44	76.76
Atlanta - Sandy Springs - Gainesville (GA-AL)	801	1.30	78.06
Other	13,610	21.94	100.00
Total	61,652	100.00	100.00

*Given the definition of Combined Statistical Area (i.e. further agglomeration of adjacent Metropolitan and Micropolitan Statistical Areas) 18 observations were not assigned to any of them.

dence. San Jose-San Francisco-Oakland area attracts almost a fifth of all migrant inventors, confirming Silicon Valley and the whole Bay Area as a top destination for high skilled migrants. As expected, New York City's greater metropolitan area holds the second place in the list, accounting for 13.05% of total foreign-born inventors. Other relevant regions are the San Diego area (with its important biotech and electronic communities) and the region gravitating around Boston, being home to 10.12% and 8.95% of immigrant inventors respectively.

It is interesting to note how immigrant inventors tend to concentrate in the top four CSAs of our list, which together account for about half of the migrant population in our sample. This pattern becomes more noticeable if compared with US native inventors localization (Table A4). While not displaying important changes in terms of main CSAs of destination, US nationals appear less concentrated. Important regions like San Francisco's Bay Area and New York City metropolitan area still account for the biggest inventors' communities, but native inventors localization appears to be more evenly distributed.

Table A4: US Native Inventors by Main CSAs

Combined Statistical Area (CSA)	Inventors	Percent	Cumulative Percent
San Jose - San Francisco - Oakland (CA)	32,916	11.55	11.55
New York - Newark - Bridgeport (NY-NJ-CT-PA)	26,959	9.46	21.00
Boston - Worcester - Manchester (MA-NH)	22,547	7.91	28.91
Los Angeles - Long Beach - Riverside (CA)	13,162	4.62	33.53
Minneapolis - St. Paul - St. Cloud (MN-WI)	12,290	4.31	37.84
San Diego - Carlsbad - San Marcos (CA)	12,012	4.21	42.06
Philadelphia - Camden - Vineland (PA-NJ-DE-MD)	11,640	4.08	46.14
Chicago - Naperville - Michigan City (IL-IN-WI)	10,582	3.71	49.85
Washington - Baltimore - N. Virginia (DC-MD-VA-WV)	8,826	3.10	52.95
Detroit - Warren - Flint (MI)	7,225	2.53	55.48
Houston - Baytown - Huntsville (TX)	7,073	2.48	57.96
Seattle - Tacoma - Olympia (WA)	6,526	2.29	60.25
Raleigh - Durham - Cary (NC)	6,043	2.12	62.37
Denver - Aurora - Boulder (CO)	5,030	1.76	64.14
Cleveland - Akron - Elyria (OH)	4,656	1.63	65.77
Other	98,116	34.23	100.00
Total	285,603	100.00	100.00

*Given the definition of Combined Statistical Area (i.e. further agglomeration of adjacent Metropolitan and Micropolitan Statistical Areas) 11 observations were not assigned to any of them.

Table A5: US Inventors Team Size by Nationality Composition*

Full Native			Mixed**		
Team size	Applications	Percent	Team size	Applications	Percent
2	40,588	41.29	2	16,919	24.65
3	25,781	26.23	3	16,500	24.04
4	14,050	14.29	4	11,851	17.26
5	7,549	7.68	5	8,057	11.74
6	3,854	3.92	6	4,977	7.25
Other	6,472	6.58	Other	10,340	15.06
Total	98,294	100.00	Total	68,644	100.00

***Patent applications count; **Teams including at least one immigrant inventor**

Figure A3: US Inventor Team Size by Nationality Composition

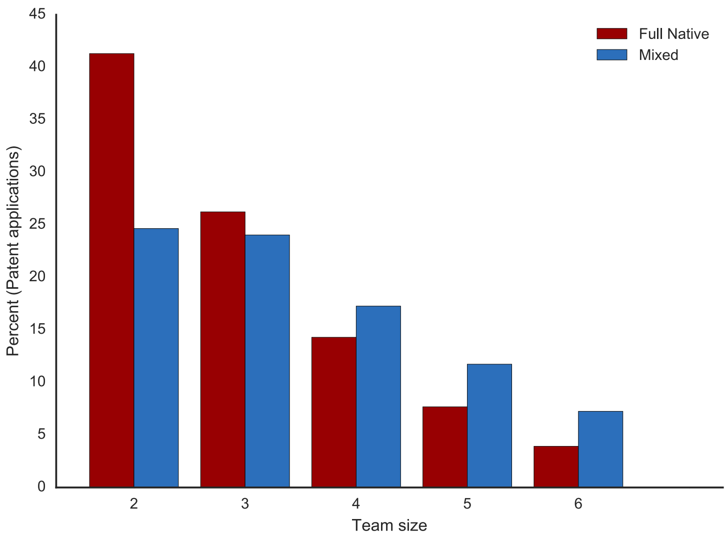
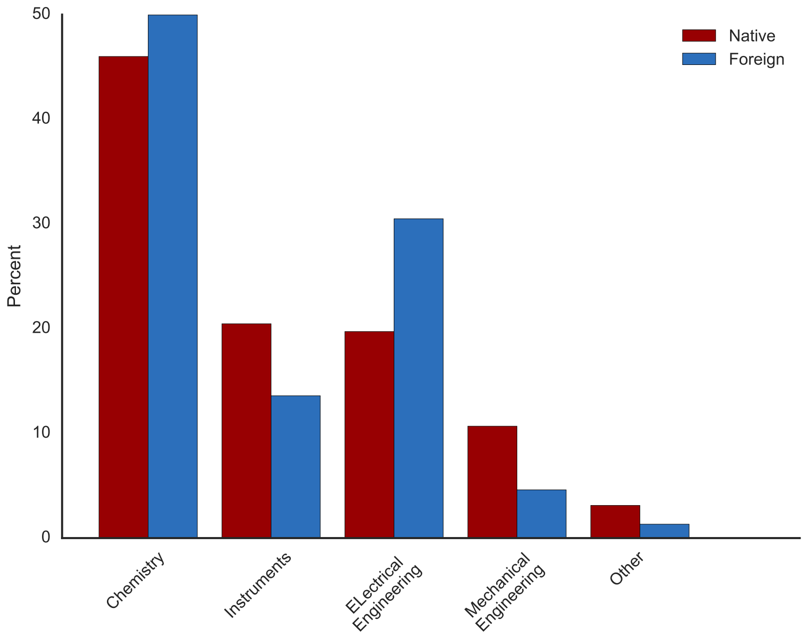


Table A5 and Figure A3 group patent applications according to inventors team size, distinguishing between teams constituted only by US nationals and those including at least one immigrant inventor. Mixed teams appear to be larger, as a greater fraction of applications reporting foreign nationality inventors lists groups larger than two individuals. On the other hand, approximately 40% of teams comprising only natives are represented by two inventors, with a smaller proportion of 4+ inventors teams.

Figure A4: US Inventors by Applications *ISI-OST-INPI* Technological Class



***According to the underlying invention, patent applications can be associated with multiple technological classes.**

Figure A4 gathers inventors by their application's *ISI-OST-INPI*[19] technological class, comparing native and foreign-born inventors. While the chemistry class accounts for the biggest part of both US nationals and immigrants patenting output, foreign nationality inventors appear more active in the electrical engineering domain. A possible explanation for this pattern could be related to the greater attraction exerted by ICT firms on foreign-born inventors, especially from a leading technological hub like Silicon Valley, famous for its international community.

[19]For a comprehensive description of *ISI-OST-INPI* technology classification see Schmoch, Ulrich, *Concept of a technology classification for country comparisons. Final Report to the World Intellectual Property Office (WIPO)*, Karslruhe: Fraunhofer ISI (2008).

EU15 Sample

As reported by Table A6, immigrant inventors share is much smaller in our EU15 sample than in the United States' one. Foreign nationality inventors residing in one of the EU15 countries are 14,112, constituting approximately 3% of the total inventors' community[20].

Table A6: EU15 Inventors by origin

	Inventors	Percent
Natives	458,575	97.01
Immigrants	14,112	2.99
Total	472,687	100.00

*Following our definition of *migrants*, the EU15 Native group includes all inventors with a EU15 nationality, regardless of their country of residence (e.g. a German inventor residing in the UK), while the EU15 Immigrant group includes all inventors whose nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

Figure A5 depicts the evolution of immigrant inventors patenting activity in the EU15 area. American and Chinese nationals constitute the more prolific foreign-nationality groups, followed by Indians, Russians and Australians. A progressive increase in patent applications request can be observed for most nationalities.

Table A7 and Figure A6 list the major foreign nationality groups resident in the region. US and Chinese nationals represent the biggest groups, accounting for 13.18% and 12.22% of total EU15 immigrant inventors. Indians and Russians constitute 6.92% and 6.77% respectively, followed by Australian and Polish inventors which represent 3.44% and 3.36% of the total.

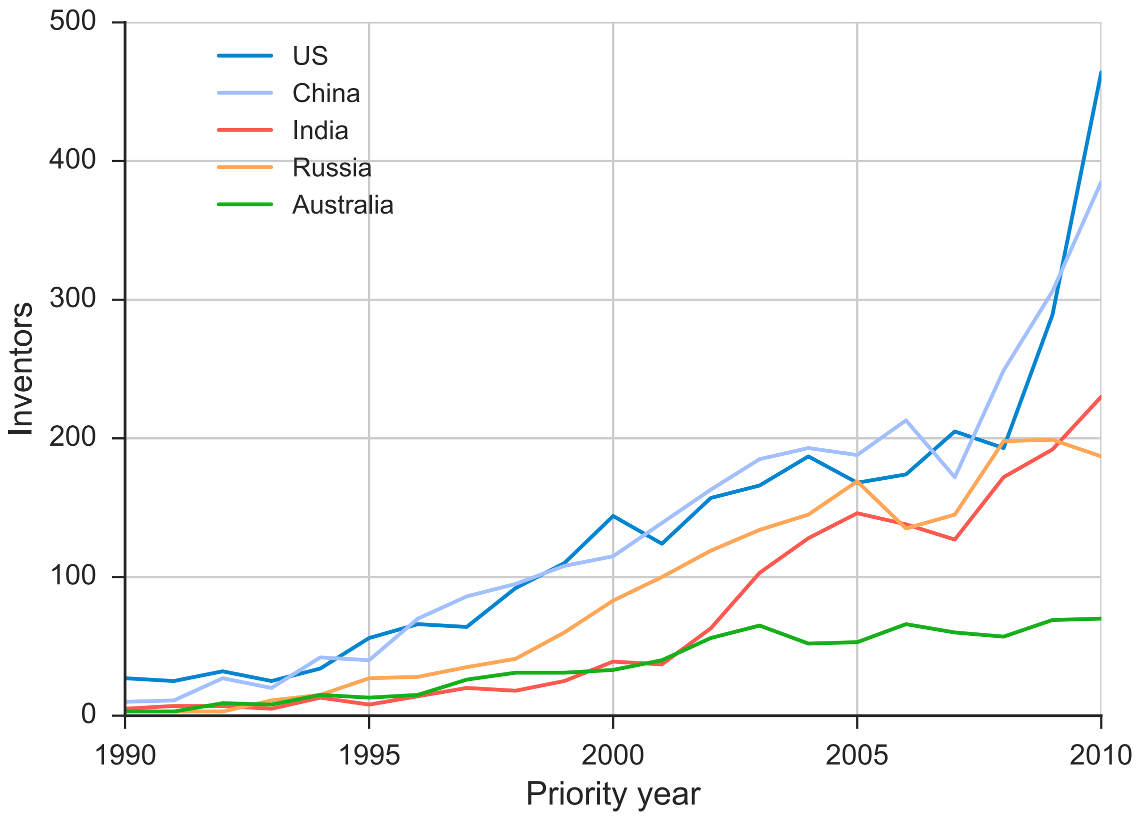
Table A8 groups EU15 resident foreign inventors according to their Metropolitan Region of residence. The Eindhoven area is reported to be the first one, given the presence of Philips research center, which attracts a significant number of foreign high skilled individuals. London and Paris constitute the second and third preferred destinations (9.22% and 7.22% respectively), followed by Munich and Cambridge regions. It is important to highlight the relevant presence of many German cities in the list.

Table A9 provides the geographical distribution of European natives only. Overall, EU15 nationals tend to be less concentrated in the main metropolitan regions. The presence and share of German areas slightly increases, along with the lower ranking of the Cambridge and especially Eindhoven regions.

Table A10 and Figure A7 group patent applications according to inventors team size. The distinction between teams constituted only by EU15 nationals and those including at least one immigrant inventor reveals how also in the European case mixed teams tend

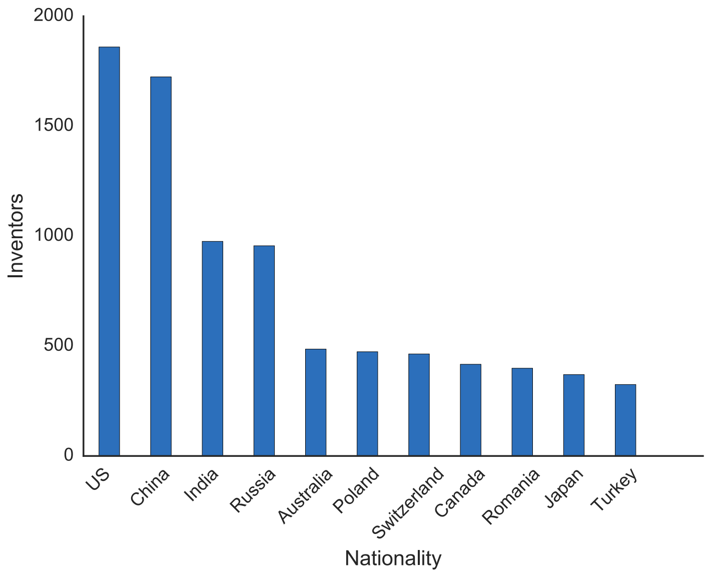
[20]As explained below, these small figures suffer from our definition of migrant inventor, given that much of EU15 high-skilled migration takes place within its boundaries

Figure A5: EU15 Main Immigrant Inventors Groups Patenting Activity



*Following our definition of *migrants*, the EU15 Native group includes all inventors with a EU15 nationality, regardless of their country of residence (e.g. a German inventor residing in the UK), while the EU15 Immigrant group includes all inventors whose nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

Figure A6: EU15 Main Immigrant Inventors Groups



*Following our definition of *migrants*, the EU15 Native group includes all inventors with a EU15 nationality, regardless of their country of residence (e.g. a German inventor residing in the UK), while the EU15 Immigrant group includes all inventors whose nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

Table A7: EU15 Main Immigrant Inventors Groups

Nationality	Inventors	Percent	Cumulative Percent	Percent on Resident Inventors
US	1,860	13.18	13.18	0.53
China	1,724	12.22	25.40	0.36
India	976	6.92	32.31	0.21
Russia	956	6.77	39.09	0.21
Australia	486	3.44	42.53	0.11
Poland	474	3.36	45.89	0.10
Switzerland	464	3.29	49.18	0.11
Canada	417	2.95	52.13	0.10
Romania	399	2.83	54.96	0.10
Japan	370	2.62	57.58	0.09
Turkey	325	2.30	59.89	0.09
Ukraine	256	1.81	61.70	0.07
Israel	248	1.76	63.46	0.06
Malaysia	244	1.73	65.19	0.05
Norway	222	1.57	66.76	0.05
Other	4,691	33.24	100.00	1.03
Total	14,112	100.00	100.00	3.22

*Following our definition of *migrants*, the EU15 Native group includes all inventors with a EU15 nationality, regardless of their country of residence (e.g. a German inventor residing in the UK), while the EU15 Immigrant group includes all inventors whose nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

Table A8: EU15 Immigrant Inventors by Main Metropolitan Regions

Metropolitan Region	Inventors	Percent	Cumulative Percent
Eindhoven	1,674	14.35	14.35
London	1,076	9.22	23.58
Paris	842	7.22	30.80
Munich	390	3.34	34.14
Cambridge	384	3.29	36.79
Stockholm	356	3.05	40.48
Helsinki	285	2.44	45.05
Stuttgart	248	2.13	46.48
Copenhagen	242	2.07	47.13
Berlin	237	2.03	49.16
Mannheim - Ludwigshafen	206	1.77	50.93
Brussels	204	1.75	52.67
Frankfurt	175	1.50	54.18
Achen	169	1.45	55.62
Vienna	148	1.27	56.89
Other	5,028	43.11	100.00
Total	11,664	100.00	100.00

*Given the definition of Metropolitan Regions (i.e. agglomerations of at least 250,000 inhabitants) 2,448 observations were not assigned to any of those.);**Following our definition of *migrants*, the EU15 Native group includes all inventors with a EU15 nationality, regardless of their country of residence (e.g. a German inventor residing in the UK), while the EU15 Immigrant group includes all inventors whose nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

Table A9: EU15 Native Inventors by Main Metropolitan Regions

Metropolitan Region	Inventors	Percent	Cumulative Percent
Paris	26,666	7.66	7.66
London	18,074	5.19	12.85
Stuttgart	13,343	3.83	16.68
Munich	12,810	3.68	20.36
Eindhoven	11,027	3.17	23.53
Berlin	8,167	2.35	25.88
Frankfurt	7,640	2.19	28.07
Copenhagen	7,522	2.16	30.23
Stockholm	7,077	2.03	32.27
Helsinki	7,016	2.02	34.28
Ruhr district	6,435	1.85	36.13
Mannheim - Ludwigshafen	5,738	1.65	37.78
Milan	5,544	1.59	39.37
Nuremberg	5,423	1.56	40.93
Cambridge	5,322	1.53	42.46
Other	200,327	57.52	100.00
Total	348.131	100.00	100.00

*Given the definition of Metropolitan Regions (i.e. agglomerations of at least 250,000 inhabitants) 110,444 observations were not assigned to any of those;**Following our definition of *migrants*, the EU15 Native group includes all inventors with a EU15 nationality, regardless of their country of residence (e.g. a German inventor residing in the UK), while the EU15 Immigrant group includes all inventors whose nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

Table A10: EU15 Inventors Team Size by Nationality Composition*

Full Native			Mixed**		
Team size	Applications	Percent	Team size	Applications	Percent
2	103,641	41.12	2	5,125	26.32
3	65,118	25.83	3	4,950	25.42
4	36,936	14.65	4	3,432	17.62
5	18,144	7.20	5	2,145	11.01
6	9,919	3.94	6	1,340	6.88
Other	18,305	7.26	Other	2,483	12.75
Total	252,063	100.00	Total	19,475	100.00

*Patent applications count; **Teams including at least one immigrant inventor;***Following our definition of *migrants*, the EU15 Native group includes all inventors with a EU15 nationality, regardless of their country of residence (e.g. a German inventor residing in the UK), while the EU15 Immigrant group includes all inventors whose nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

to be larger. As for the US, roughly 40% of teams comprising only EU15 natives are represented by two inventors, with a smaller proportion of 4+ inventors teams.

Figure A8 groups inventors by their application's *ISI-OST-INPI* technological class, comparing immigrant inventors with EU15 nationals. As for the US sample, the chemistry domain accounts for most patenting activity of both nationality cohorts, with immigrants more active in the electrical engineering class.

The definition of *migrant* inventors for the EU15 sample inevitably ignores intra-European migration flows. However, if we include EU15 nationals residing in a different country than their native one in our accounts, it's easy to notice how a great portion of EU15 inventors' mobility originates within the region[21].

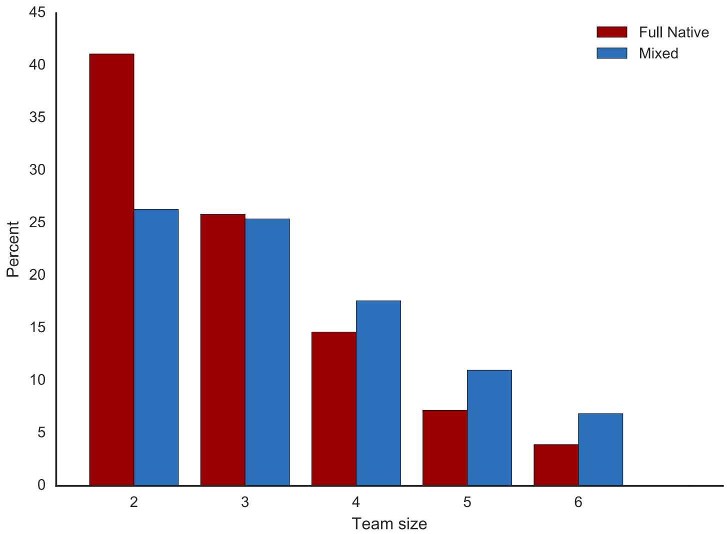
As reported by Figure A9, Table A11 and Figure A10, German natives represent the biggest group of migrants, accounting with French and Dutch inventors for roughly a third of EU15 inventors' migrant flows. The contribution to inventors' migration towards the EU15 of nationalities outside the region has a decisively smaller magnitude, with US and Chinese nationals constituting 4.89% and 4.53% of total immigrants, followed by Indians and Russians who represent 2.56% and 2.51% of foreign-nationality inventors.

Table A11: EU15 Main Immigrant Inventors Groups (Intra-European Migrants Included)

Nationality	Inventors	Percent	Cumulative Percent	Percent on Resident Inventors
Germany	5,616	14.75	14.75	1.19
France	3,543	9.31	24.06	0.79
Netherlands	3,044	8.00	32.05	0.66
UK	2,911	7.65	39.70	0.53
Italy	2,107	5.53	45.24	0.46
US	1,860	4.89	50.12	0.36
China	1,724	4.53	54.65	0.32
Spain	1,311	3.44	58.09	0.29
Austria	1,199	3.15	61.24	0.27
Belgium	1,193	3.13	64.38	0.26
India	976	2.56	66.94	0.21
Russia	956	2.51	69.45	0.21
Ireland	672	1.77	71.22	0.16
Greece	639	1.68	72.90	0.14
Sweden	610	1.60	74.50	0.13
Other	9.708	25.50	100.00	2.15
Total	38,069	100.00	100.00	8.13

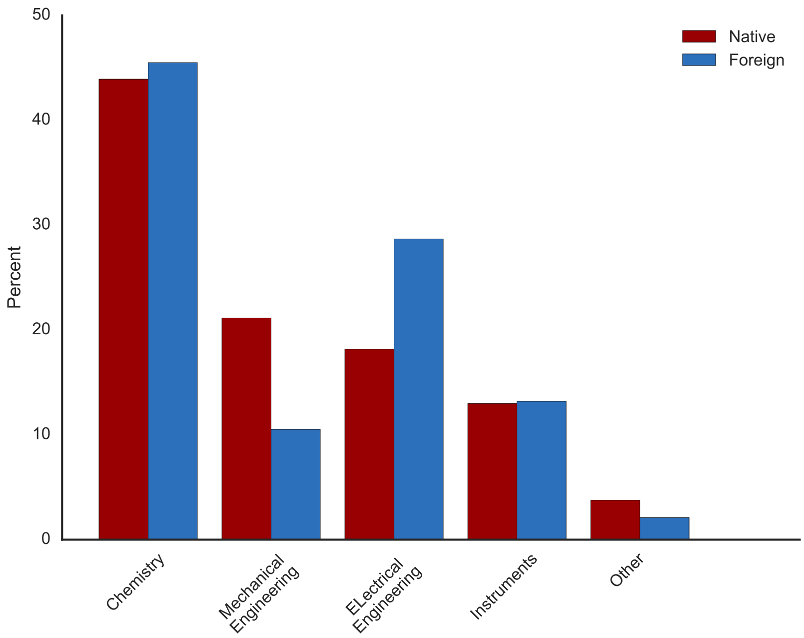
[21] To account for the magnitude of intra-European inventors mobility, we performed our regressions on an additional sample including patents whose inventors reside within the same EU15 country, considering as *migrants* all the individuals with a different nationality than their country of residence.

Figure A7: EU15 Inventor Team Size by Nationality Composition



*Following our definition of *migrants*, the EU15 Native group includes all inventors with a EU15 nationality, regardless of their country of residence (e.g. a German inventor residing in the UK), while the EU15 Immigrant group includes all inventors whose nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

Figure A8: EU15 Inventors by Applications *ISI-OST-INPI* Technological Class



*According to the underlying invention, patent applications can be associated with multiple technological classes;**Following our definition of *migrants*, the EU15 Native group includes all inventors with a EU15 nationality, regardless of their country of residence (e.g. a German inventor residing in the UK), while the EU15 Immigrant group includes all inventors whose nationality is not that of a EU15 country (e.g. an Indian inventor residing in the Netherlands).

Figure A9: EU15 Resident Foreign Nationality Inventors (Intra-European Migrants Included)

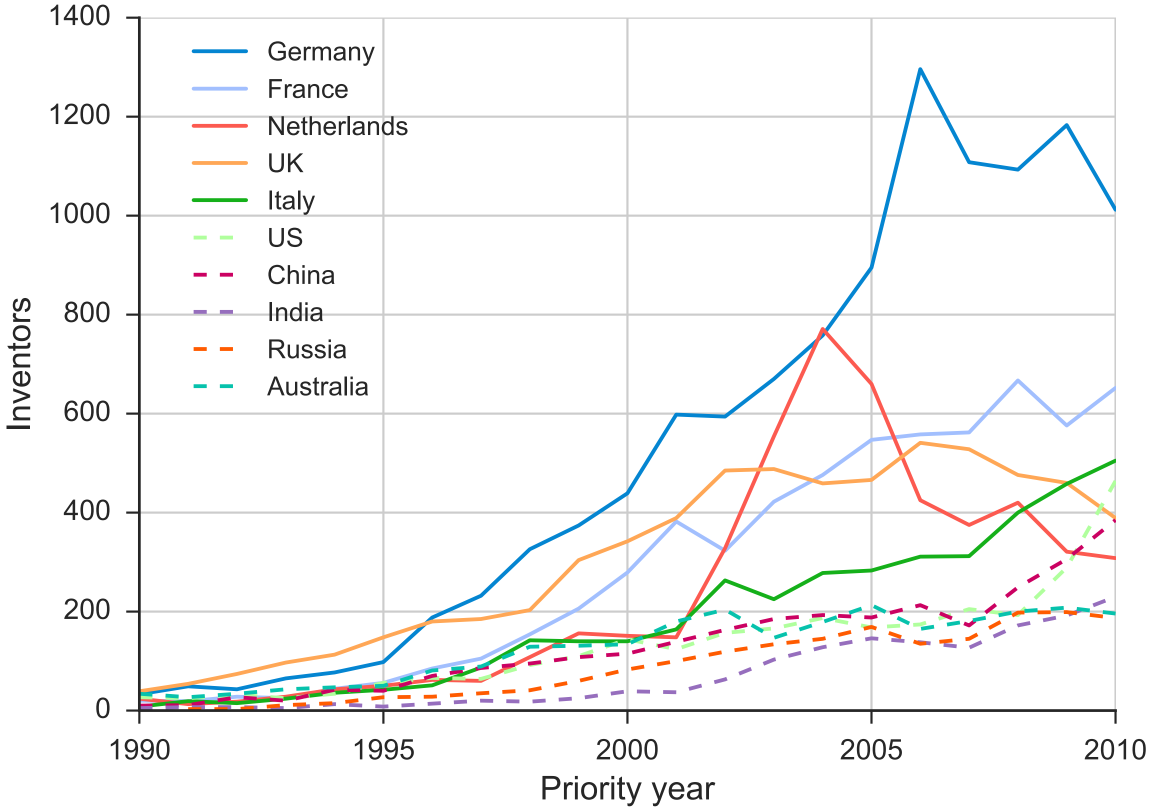
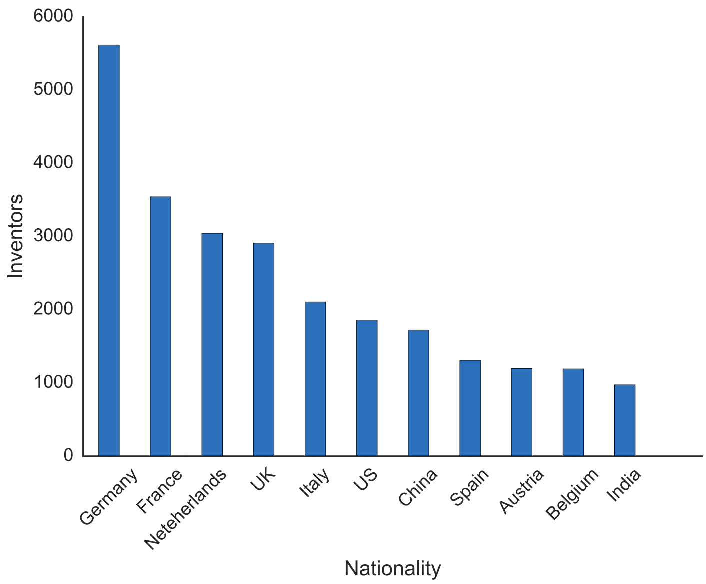


Figure A10: EU15 Main Immigrant Inventors Groups (Intra-European Migrants Included)



Inventors Address Technical Issues

A crucial information for the purposes of our study is represented by inventors' addresses. According to the location listed on patent applications, we are able to distinguish between inventors whose country of residence is the same as their nationality (which we treat as *natives*), and inventors whose nationality differs from their country of residence, which we define as *migrants*[22]. Given its critical role, we verified the reliability of this information.

The main issue we encountered pertains to the ambiguity of addresses which clearly do not represent the inventors' real residence. In most cases, the address of the company, a corporate research center or that of a subsidiary, substituted inventors' actual address, while for few observations we stumbled into the address of the law firm taking care of the patent application procedure. The problem arising from these addresses is that we would observe a different country of residence than the nationality even for inventors who didn't actually relocated in another country. While for research centers it could be the case for some inventors to report a corporate addresses and still be considered as migrants, inventors with the address of the company headquarters, a subsidiary or of a law firm, almost surely cannot be included in the group of immigrants. Companies' patenting policies in terms of information diffusion about their employees would inflate our count of migrants, mistakenly including individuals reporting a foreign address only because employed by a multinational firm. This section presents the main groups of ambiguous addresses we detected, and how our samples would change if we didn't consider those observations as reliable.

Most of inventors' critical addresses we found in the EU15 sample are related to Philips or its spin-off NXP. Of 4,802 observations, 3,469 addresses were associated with Philips research center in Eindhoven[23], 489 with Philips research center in Achen[24], 550 with a NXP british office[25] and 294 with a law firm taking care of Philips and NXP patent applications in France.

Figure A11 reports how EU15 main immigrant inventors groups' patenting activity would change if we didn't consider the inventors associated with those critical addresses as *migrants*. The main changes concern US, Chinese and Indian nationals groups who, especially after 2000, experience a progressive drop in their patenting activity with respect to the original sample. It is worth noting how for Americans and Chinese the surprising surge in patenting over the last two years considered by our sample would disappear, making their patenting patterns less unusual.

[22]As previously discussed, this definition slightly differs for the full EU15 sample where we consider a EU15-resident inventor as *immigrant* only if its nationality does not fall within the EU15 nationalities spectrum, thus not accounting for intra-european migration.

[23]Philips Research Eindhoven, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands.

[24]Philips GmbH Innovative Technologies Research Laboratories, Pauwelsstraße 17, 52074 Aachen, Germany.

[25]NXP Semiconductors, Redcentral, 60 High Street, Redhill, Surrey, RH1 1SH, United Kingdom.

Figure A12 clearly shows how EU15 immigrant groups would reduce their size if we were to remove the inventors associated with a critical address from our *migrants'* cohort.

An important change in our description of data can be also seen in the main metropolitan areas attracting foreign nationality inventors A12, with the metropolitan regions of London, Paris, Achen and especially Eindhoven losing part of their immigrant inventors communities.

Figure A11: EU15 Main Immigrant Inventors Groups Patenting Activity

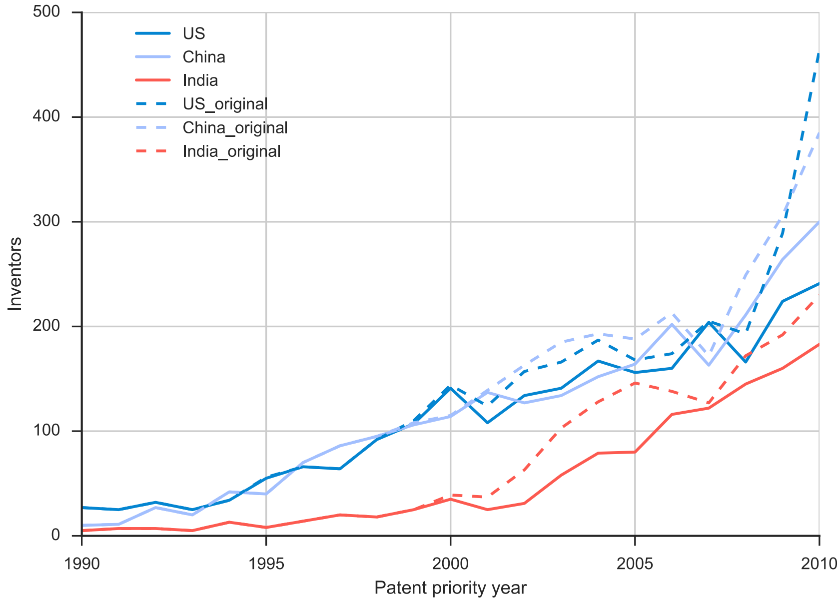
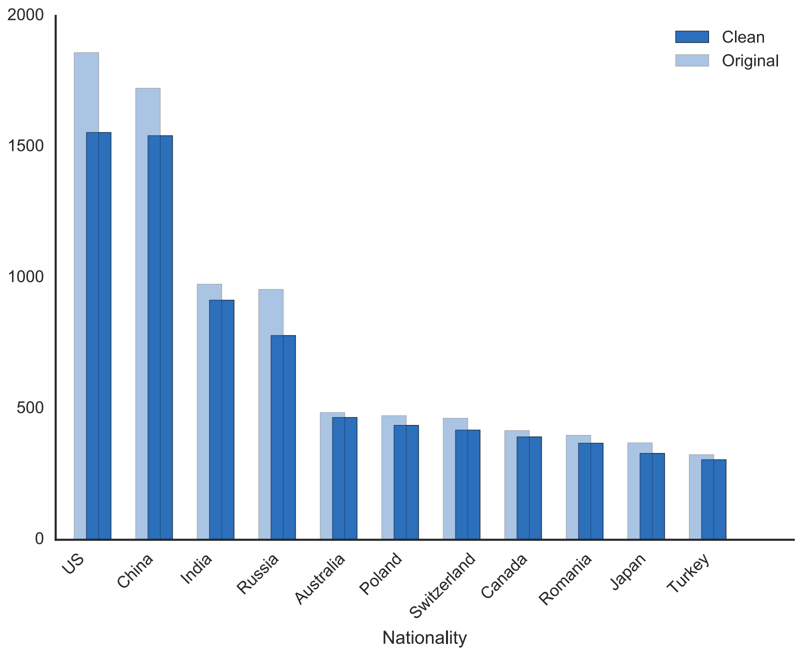


Figure A12: EU15 Main Immigrant Inventors Groups



The major critical addresses we found examining the US sample were 3,318. A great portion of it (2,196 observations) is related with inventors reporting Qualcomm's head-

Table A12: EU15 Immigrant Inventors by Main Metropolitan Areas

Clean			Regular		
Metropolitan Area	Inventors	Percent	Metropolitan Area	Inventors	Percent
London	1,004	9.73	*Eindhoven*	*1,674*	*14.35*
Paris	800	7.76	*London*	*1,076*	*9.22*
Eindhoven	514	4.98	*Paris*	*842*	*7.22*
Munich	390	3.78	Munich	390	3.34
Cambridge	384	3.72	Cambridge	384	3.29
Stockholm	356	3.45	Stockholm	356	3.05
Helsinki	285	2.76	Helsinki	285	2.44
Stuttgart	248	2.40	Stuttgart	248	2.13
Copenhagen	242	2.35	Copenhagen	242	2.07
Berlin	237	2.30	Berlin	237	2.03
Mannheim - Ludwigshafen	206	2.00	Mannheim - Ludwigshafen	206	1.77
Brussels	204	1.98	Brussels	204	1.75
Frankfurt	175	1.70	Frankfurt	175	1.50
Wien	148	1.43	*Achen*	*169*	*1.45*
Dublin	141	1.37	Vienna	148	1.27
Other	4,980	48.28	Other	5,028	43.11
Total	10,314	100.00	Total	11,664	100.00

quarters address in San Diego[26], while minor shares are associated with Merck Rahway research center[27] (613 observations), Philips office in Briaricliff Manor, NY[28] (246 observations), 3M corporate headquarters in Saint Paul, MN[29] (149 observations) and HP headquarters in Palo Alto[30] (114 observations).

Figure A13 and A14 show how not considering these inventors as true migrants would slightly decrease US main immigrant groups patenting activity and size. As in the EU15 case the second half of the sample experiences the biggest difference in terms of patenting activity gap.

Table A13 highlights how the CSAs associated with critical addresses would lose part of their immigrant inventors population. The case of San Diego's area is the more apparent which, previously inflated by inventors reporting Qualcomm's corporate addresses, loses roughly a third of its resident foreign nationality inventors.

[26]Qualcomm Inc, 5775 Morehouse Drive, San Diego, CA 92121, USA.
[27]Merck & Co, 90 E Scott Ave, Rahway, NJ 07065, USA.
[28]North American Philips Corporation, 345 Scarborough Rd, Briarcliff Manor, NY 10510, USA.
[29]3M Company, I-94 and McKnight Rd. St. Paul, 55144 MN, USA.
[30]HP Inc., 1501 Page Mill Road, Palo Alto, CA 94304, USA.

Figure A13: US Main Immigrant Inventors Groups Patenting Activity

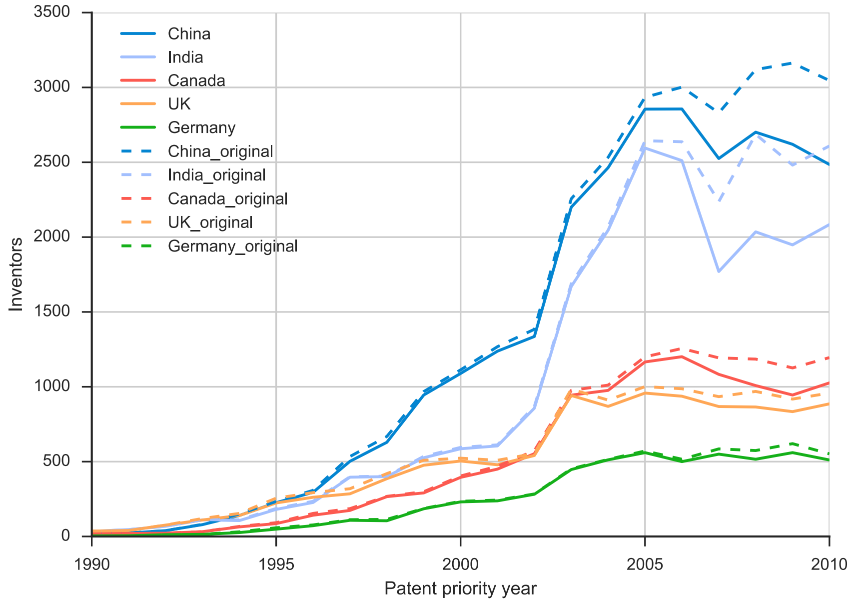


Figure A14: US Main Immigrant Inventors Groups

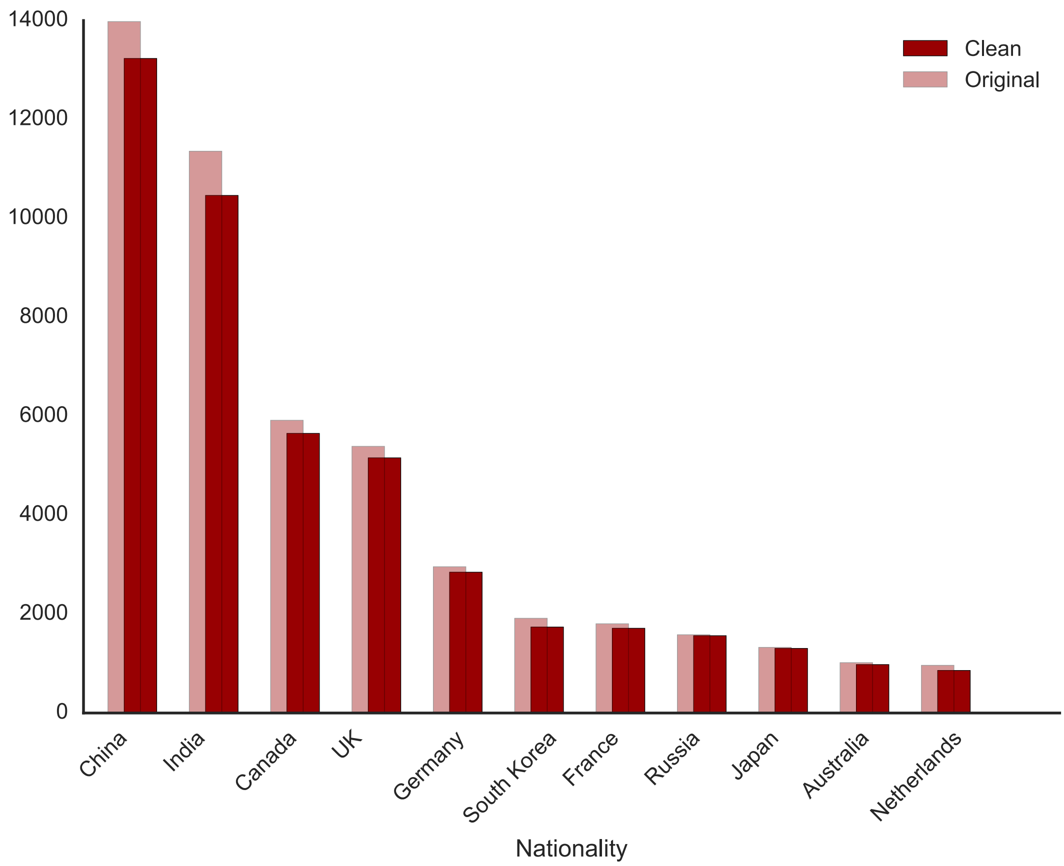


Table A13: US Immigrant Inventors by Main CSAs

Combined Statistical Area (CSA)	Inventors	Percent	Cumulative Percent
Clean			
San Jose - San Francisco - Oakland (CA)	11,234	19.29	19.29
New York - Newark - Bridgeport (NY-NJ-CT-PA)	7,173	12.32	31.61
Boston - Worcester - Manchester (MA-NH)	5,507	9.46	41.07
San Diego - Carlsbad - San Marcos (CA)	4,031	6.92	47.99
Los Angeles - Long Beach - Riverside (CA)	2,859	4.91	52.90
Philadelphia - Camden - Vineland (PA-NJ-DE-MD)	2,205	3.79	56.69
Washington - Baltimore - N. Virginia (DC-MD-VA-WV)	1,950	3.35	60.04
Chicago - Naperville - Michigan City (IL-IN-WI)	1,827	3.14	63.18
Houston - Baytown - Huntsville (TX)	1,573	2.70	65.88
Detroit - Warren - Flint (MI)	1,315	2.26	68.14
Raleigh - Durham - Cary (NC)	1,222	2.10	70.24
Seattle - Tacoma - Olympia (WA)	1,162	2.00	72.23
Minneapolis - St. Paul - St. Cloud (MN-WI)	979	1.68	73.91
Dallas - Fort Worth (TX)	888	1.53	75.44
Atlanta - Sandy Springs - Gainesville (GA-AL)	801	1.38	76.81
Other	13,608	23.19	100.00
Total	58,334	100.00	100.00
Regular			
Combined Statistical Area (CSA)	Inventors	Percent	Cumulative Percent
San Jose - San Francisco - Oakland (CA)	*1,348*	*18.44*	*18.44*
New York - Newark - Bridgeport (NY-NJ-CT-PA)	*8,031*	*13.05*	*31.49*
San Diego - Carlsbad - San Marcos (CA)	*6,226*	*10.12*	*41.60*
Boston - Worcester - Manchester (MA-NH)	5,507	8.95	50.55
Los Angeles - Long Beach - Riverside (CA)	2,859	4.65	55.20
Philadelphia - Camden - Vineland (PA-NJ-DE-MD)	2,205	3.58	58.78
Washington - Baltimore - N. Virginia (DC-MD-VA-WV)	1,950	3.17	61.95
Chicago - Naperville - Michigan City (IL-IN-WI)	1,827	2.97	64.92
Houston - Baytown - Huntsville (TX)	1,573	2.56	67.47
Detroit - Warren - Flint (MI)	1,315	2.14	69.61
Raleigh - Durham - Cary (NC)	1,222	1.99	71.59
Seattle - Tacoma - Olympia (WA)	1,162	1.89	73.48
Minneapolis - St. Paul - St. Cloud (MN-WI)	*1,128*	*1.83*	*75.32*
Dallas - Fort Worth (TX)	888	1.44	76.76
Atlanta - Sandy Springs - Gainesville (GA-AL)	801	1.30	78.06
Other	13,610	21.94	100.00
Total	61,652	100.00	100.00

Appendix B - Additional tables

Table B1: Negative Binomial estimation, inventor team diversity and patent quality

	cit3 US sample	cit5 US sample	cit US sample	cit3 EU sample	cit5 EU sample	cit EU sample
diversity	0.205***	0.190***	0.159***	0.415***	0.410***	0.391***
	(0.0141)	(0.0142)	(0.0147)	(0.0200)	(0.0196)	(0.0203)
team experience	0.00654***	0.00586***	0.00466***	0.00389***	0.00350***	0.00254***
	(0.000359)	(0.000368)	(0.000368)	(0.000203)	(0.000213)	(0.000224)
family size	0.0324***	0.0343***	0.0346***	0.0412***	0.0443***	0.0470***
	(0.000791)	(0.000799)	(0.000855)	(0.000728)	(0.000735)	(0.000780)
patent scope	0.0506***	0.0463***	0.0334***	0.0849***	0.0801***	0.0773***
	(0.00432)	(0.00415)	(0.00438)	(0.00426)	(0.00425)	(0.00469)
university	-0.0175	0.00405	0.00305	-0.289***	-0.256***	-0.225***
	(0.0124)	(0.0126)	(0.0138)	(0.0161)	(0.0162)	(0.0178)
triadic	0.346***	0.355***	0.364***	0.401***	0.435***	0.479***
	(0.00830)	(0.00839)	(0.00880)	(0.00821)	(0.00804)	(0.00857)
previous collaborations	0.0200*	0.00926	-0.00672	0.0595***	0.0489***	0.0361***
	(0.00858)	(0.00854)	(0.00876)	(0.00795)	(0.00786)	(0.00809)
team size	0.0712***	0.0752***	0.0761***	0.0588***	0.0614***	0.0607***
	(0.00238)	(0.00240)	(0.00249)	(0.00262)	(0.00262)	(0.00278)
NUTS3 regions	-0.0185***	-0.0206***	-0.0196***	-0.0172***	-0.0163***	-0.00758*
	(0.00383)	(0.00387)	(0.00403)	(0.00344)	(0.00348)	(0.00376)
Constant	-0.226***	-0.246***	-0.111***	-1.002***	-1.044***	-0.961***
	(0.0163)	(0.0164)	(0.0169)	(0.0179)	(0.0180)	(0.0185)
Constant	0.000284	0.104***	0.250***	0.141***	0.395***	0.688***
	(0.00558)	(0.00534)	(0.00531)	(0.00613)	(0.00512)	(0.00447)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	163572	163572	163572	246056	246056	246056
Log-likelihood	-412573.5	-457632.1	-516318.6	-453522.6	-521876.4	-618559.0

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B2: Negative Binomial estimation, inventor team weighted diversity and patent quality

	cit3 US sample	cit5 US sample	cit US sample	cit3 EU sample	cit5 EU sample	cit EU sample
weighted diversity	0.184***	0.166***	0.129***	0.561***	0.560***	0.550***
	(0.0178)	(0.0181)	(0.0186)	(0.0299)	(0.0299)	(0.0311)
team experience	0.00656***	0.00590***	0.00470***	0.00391***	0.00353***	0.00256***
	(0.000360)	(0.000369)	(0.000369)	(0.000203)	(0.000213)	(0.000225)
family size	0.0325***	0.0344***	0.0348***	0.0413***	0.0445***	0.0472***
	(0.000793)	(0.000800)	(0.000856)	(0.000730)	(0.000737)	(0.000782)
patent scope	0.0515***	0.0471***	0.0341***	0.0851***	0.0803***	0.0775***
	(0.00435)	(0.00417)	(0.00438)	(0.00426)	(0.00425)	(0.00469)
university	-0.0169	0.00495	0.00474	-0.292***	-0.260***	-0.228***
	(0.0124)	(0.0126)	(0.0138)	(0.0161)	(0.0162)	(0.0178)
triadic	0.346***	0.355***	0.364***	0.403***	0.437***	0.481***
	(0.00833)	(0.00840)	(0.00880)	(0.00836)	(0.00815)	(0.00865)
previous collaborations	0.0210*	0.0102	-0.00580	0.0588***	0.0481***	0.0352***
	(0.00860)	(0.00856)	(0.00878)	(0.00804)	(0.00793)	(0.00813)
team size	0.0735***	0.0774***	0.0781***	0.0610***	0.0635***	0.0626***
	(0.00239)	(0.00241)	(0.00250)	(0.00261)	(0.00261)	(0.00277)
NUTS3 regions	-0.0201***	-0.0221***	-0.0212***	-0.0200***	-0.0190***	-0.00998**
	(0.00385)	(0.00389)	(0.00405)	(0.00344)	(0.00349)	(0.00375)
Constant	-0.217***	-0.237***	-0.103***	-0.997***	-1.040***	-0.958***
	(0.0163)	(0.0164)	(0.0169)	(0.0180)	(0.0180)	(0.0186)
Constant	0.00158	0.105***	0.251***	0.143***	0.396***	0.689***
	(0.00559)	(0.00534)	(0.00531)	(0.00619)	(0.00515)	(0.00448)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	163151	163151	163151	245776	245776	245776
Log-likelihood	-411479.0	-456428.3	-514960.8	-453045.9	-521303.3	-617873.9

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B3: Negative Binomial estimation, foreign inventorship and patent quality.

	cit3 US sample	cit5 US sample	cit US sample	cit3 EU sample	cit5 EU sample	cit EU sample
foreign same	0.0832**	0.0747**	0.0682*	0.0540	0.0389	-0.0163
	(0.0284)	(0.0282)	(0.0297)	(0.0307)	(0.0308)	(0.0309)
foreign mix	0.150***	0.127***	0.0930***	0.145***	0.145***	0.142***
	(0.0194)	(0.0194)	(0.0199)	(0.0317)	(0.0309)	(0.0324)
locals foreign same	0.0736***	0.0702***	0.0595***	0.163***	0.165***	0.155***
	(0.00789)	(0.00792)	(0.00819)	(0.0103)	(0.0100)	(0.0103)
locals foreign mix	0.162***	0.150***	0.130***	0.308***	0.291***	0.280***
	(0.0121)	(0.0123)	(0.0127)	(0.0208)	(0.0210)	(0.0224)
team experience	0.00649***	0.00583***	0.00463***	0.00388***	0.00349***	0.00254***
	(0.000359)	(0.000368)	(0.000367)	(0.000203)	(0.000213)	(0.000225)
family size	0.0324***	0.0343***	0.0346***	0.0411***	0.0442***	0.0469***
	(0.000790)	(0.000799)	(0.000854)	(0.000728)	(0.000735)	(0.000780)
patent scope	0.0507***	0.0463***	0.0334***	0.0848***	0.0800***	0.0772***
	(0.00433)	(0.00416)	(0.00438)	(0.00426)	(0.00425)	(0.00470)
university	-0.0171	0.00430	0.00291	-0.286***	-0.254***	-0.223***
	(0.0124)	(0.0126)	(0.0138)	(0.0161)	(0.0162)	(0.0178)
triadic	0.346***	0.355***	0.364***	0.402***	0.435***	0.480***
	(0.00830)	(0.00839)	(0.00880)	(0.00824)	(0.00806)	(0.00858)
previous collaborations	0.0198*	0.00912	-0.00679	0.0597***	0.0491***	0.0362***
	(0.00858)	(0.00854)	(0.00876)	(0.00796)	(0.00787)	(0.00809)
team size	0.0664***	0.0707***	0.0720***	0.0539***	0.0569***	0.0565***
	(0.00251)	(0.00253)	(0.00261)	(0.00268)	(0.00268)	(0.00284)
NUTS3 regions	-0.0173***	-0.0195***	-0.0187***	-0.0166***	-0.0159***	-0.00747*
	(0.00384)	(0.00389)	(0.00405)	(0.00345)	(0.00350)	(0.00378)
Constant	-0.214***	-0.235***	-0.101***	-0.987***	-1.030***	-0.947***
	(0.0163)	(0.0164)	(0.0170)	(0.0180)	(0.0180)	(0.0186)
Constant	-0.000114	0.104***	0.250***	0.141***	0.395***	0.688***
	(0.00558)	(0.00534)	(0.00531)	(0.00615)	(0.00513)	(0.00447)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	163572	163572	163572	246056	246056	246056
Log-likelihood	-412552.2	-457617.3	-516307.7	-453534.9	-521890.8	-618570.5

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B4: Negative Binomial estimation on top cited patents. Dependent variable is cit3.

	Top 1% US sample	Top 10% US sample	Top 25% US sample	Top 1% EU sample	Top 10% EU sample	Top 25% EU sample
diversity	-0.0308	0.0128	0.0551***	0.119	0.124***	0.151***
	(0.0326)	(0.0163)	(0.0138)	(0.0758)	(0.0342)	(0.0245)
team experience	0.00357***	0.00518***	0.00556***	-0.0000478	0.000523	0.000977***
	(0.000337)	(0.000297)	(0.000304)	(0.000771)	(0.000276)	(0.000192)
family size	0.00195	0.00856***	0.0131***	0.00187	0.00665***	0.0106***
	(0.00163)	(0.000872)	(0.000758)	(0.00198)	(0.000863)	(0.000705)
patent scope	0.0312**	0.0287***	0.0310***	0.0480**	0.0419***	0.0434***
	(0.00988)	(0.00490)	(0.00414)	(0.0170)	(0.00628)	(0.00459)
university	-0.0131	-0.0407**	-0.0374**	-0.0431	-0.0545	-0.129***
	(0.0376)	(0.0148)	(0.0119)	(0.0986)	(0.0310)	(0.0190)
triadic	0.0381	0.0844***	0.123***	0.0911*	0.0868***	0.113***
	(0.0235)	(0.0102)	(0.00834)	(0.0420)	(0.0152)	(0.0104)
previous collaborations	-0.00761	-0.0223*	-0.0188*	0.0944**	0.0157	0.00820
	(0.0193)	(0.00953)	(0.00816)	(0.0358)	(0.0142)	(0.00983)
team size	0.00596	0.0209***	0.0313***	-0.00887	0.00652	0.0183***
	(0.00505)	(0.00268)	(0.00232)	(0.00933)	(0.00357)	(0.00275)
NUTS3 regions	0.0187*	-0.000715	-0.00684	-0.00667	-0.00883	-0.0147***
	(0.00789)	(0.00449)	(0.00383)	(0.0136)	(0.00510)	(0.00370)
Constant	2.459***	1.545***	1.122***	2.622***	1.529***	1.066***
	(0.0356)	(0.0188)	(0.0158)	(0.0682)	(0.0271)	(0.0192)
Constant	-2.132***	-1.676***	-1.345***	-2.268***	-2.067***	-1.844***
	(0.0581)	(0.0206)	(0.0133)	(0.104)	(0.0535)	(0.0313)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Tech dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	3285	26930	57798	850	12082	36213
ll	-12982.2	-89529.6	-177220.0	-3175.0	-36460.6	-97864.5

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B5: Cross-correlation table, EU15

Variables	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
[1]diversity	1.000											
[2]weighted diversity	0.848	1.000										
	(0.000)											
[3]citations	0.036	0.021	1.000									
	(0.000)	(0.000)										
[4]citations 5 years	0.061	0.043	0.947	1.000								
	(0.000)	(0.000)	(0.000)									
[5]citations 3 years	0.071	0.053	0.891	0.959	1.000							
	(0.000)	(0.000)	(0.000)	(0.000)								
[6]team size	0.095	0.076	0.081	0.099	0.105	1.000						
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)							
[7]university	0.064	0.076	-0.048	-0.043	-0.042	0.046	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)						
[8]team experience	0.038	0.031	0.079	0.109	0.123	0.259	-0.131	1.000				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
[9]previous collaborations	0.021	0.019	0.042	0.059	0.068	0.299	-0.064	0.550	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
[10]triadic	-0.013	-0.025	0.327	0.267	0.227	0.049	-0.055	0.056	0.036	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
[11]NUTS3 regions	-0.037	-0.022	0.038	0.034	0.033	0.531	-0.013	0.277	0.224	0.037	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
[21]family size	0.019	-0.005	0.258	0.244	0.230	0.110	-0.061	0.061	0.057	0.382	0.029	1.000
	(0.000)	(0.010)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

Table B6: Cross-correlation table, US

Variables	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
[1]diversity	1.000											
[2]weighted diversity	0.870	1.000										
	(0.000)											
[3]citations	0.047	0.034	1.000									
	(0.000)	(0.000)										
[4]citations 5 years	0.076	0.064	0.958	1.000								
	(0.000)	(0.000)	(0.000)									
[5]citations 3 years	0.091	0.080	0.906	0.966	1.000							
	(0.000)	(0.000)	(0.000)	(0.000)								
[6]team size	0.176	0.156	0.074	0.100	0.109	1.000						
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)							
[7]university	0.072	0.089	-0.011	-0.010	-0.018	-0.026	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)						
[8]team experience	0.171	0.167	0.115	0.152	0.165	0.190	-0.101	1.000				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
[9]previous collaborations	0.106	0.103	0.048	0.067	0.074	0.241	-0.042	0.496	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
[10]triadic	0.005	-0.001	0.347	0.306	0.265	0.024	-0.005	0.109	0.072	1.000		
	(0.058)	(0.623)	(0.000)	(0.000)	(0.000)	(0.000)	(0.044)	(0.000)	(0.000)			
[11]NUTS3 regions	-0.011	-0.021	0.032	0.038	0.035	0.483	-0.022	0.082	0.097	0.018	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
[12]family size	0.055	0.036	0.226	0.221	0.205	0.113	-0.046	0.177	0.113	0.365	0.051	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

Chapter 2

Migration and local innovation: post-USSR inventors in Europe and Israel

2.1 Introduction

Today most of the European countries are experiencing considerable influxes of migrants. After a short decline ensuing the financial crisis, the total number of permanent migrations to the OECD area as a whole is at its highest level since half a decade. Overall 4.8 million new permanent entries were registered in 2015 (almost a 10% increase over 2014), of which about the 11% is ascribed to permanent labour migration. Although permanent labour migration is declining, if we exclude Spain and Italy quotas, the figures for 2015 are higher than for 2014 (about +2%)[1].

According to the most recent OECD International Migration Outlook, recent migration waves have two characteristics . First of all employment rates among foreign-born individuals are increasing (and the same evidence is given by Eurostat data[2]). Second, the share of highly-educated, foreign-born individuals over the total working population is rising quite steadily since 2010; for the OECD area, the share of highly educated people for recent foreign newcomers (i.e. those who reached their destination country within the last five years) is around 32%, about twice the percentage observed for long-settled immigrants (OECD 2016).

Most strikingly, the provision of skilled workforce comes not only from economic migrants, but also from migrants with refugee status: the 15% of migrants who asked for

[1]Estimates in Bański 2013 suggests a polarization process to happen in Central Europe within the next decades: aging population, low birth rates, and severe climate changes in Sub-Saharan Africa may result in even higher net migration rates.

[2]Source: own elaborations on EU28 foreign population employment rates.
http://ec.europa.eu/eurostat/web/migrant-integration/data/database

asylum in France and Germany in 2015 had a tertiary degree[3].

Given these circumstances a solid, global response is still lacking. Some governments are reacting to the aforementioned dynamics with the introduction of new policies (or with a revision of existing practices), having acknowledged the implications of high-skilled migration in terms of employment and economic growth. France approved several modifications to the *Loi relative au droit des etrangers* in 2016, among which the introduction of a four-year "Talent Passport" for highly-skilled migrants and their families (Garcia-Peñalosa and Wasmer 2016). In 2015 Canada established the Express Entry procedure, a point-based system (where points are assigned based on individual features as work experience and level of education) which aims to accelerate the process of admission of foreign workers, in order to respond to specific employment needs. In the same year, and along the same principle, Ireland went through an extensive revision of its labor migration system: among several changes, a lower salary threshold was set for a number of highly-skilled occupations. In parallel, the above trends have aroused a growing interest among scholars toward the mechanics that regulate high-skilled migratory events and innovation, being the latter a vital component of countries' development and economic growth (Pradhan et al. 2016; Lucas 1988).

Overall, the literature suggest three mechanisms through which migrants, especially highly skilled ones, may enhance innovation in destination countries, namely:

- Increasing workforce diversity (ethnic or birthplace-driven) either in the host country or at the firm level.
- Positive self-selection of migrants: usually the best individuals in terms of skills and knowledge-base relocate achieve higher rewards.
- Knowledge transfer. The diffusion of knowledge embodied in high-skilled, foreign-born individuals has been addressed from the point of view of both the receiving and the sending country.

The present work partly relates to the third process. More specifically, I test whether inventors who migrated from the former Soviet Union (to whom I will refer as "ex-USSR inventors", in the rest of the paper) generate a positive knowledge externality for their native peers, at the country level. A formal, specific hypothesis can be formulated as follows:

HP: Inventors who migrated from the former Soviet Union (*ex-USSR* inventors in the rest of the paper) induce a positive knowledge externality for their native peers.

[3] Beyer 2016 documented a convergence process between natives' and foreigners' employment rates: interestingly, such catch-up mechanism is found to be faster for higher educated migrants.

I first examine the literature on highly-skilled migration from the former Soviet Republics towards Western Europe, Israel and the United States, from the collapse of the Soviet Union (in 1980 up to 2000). I will stress how this led to a considerable influx of ex-USSR scientists and academics.

I then use data on patent applications filed under the Patent Cooperation Treaty (Miguelez et al. 2013) to track the activity of local inventors active in Western Europe and Israel after the collapse of the Soviet Union, and show that the technological classes affected by such migratory shock outperformed the others in terms of local patent production.

I adopt two empirical strategies. In the first one I use an annual panel data set of EPO-PCT patent applications, and I estimate a model in which the number of patents produced by local inventors is explained by a dummy (that varies across years and technological classes) signaling the presence of active ex-USSR inventors in the focal class.

Given that the presence of ex-USSR inventors is likely to be endogenous, I move then to a classical DID approach. I collect all EPO-PCT patents signed prior to 1992 by inventors residing in one of the former Soviet Republics of which I identify the relevant technological classes, which constitute my treated group. After matching such group to a group of control technologies, I estimate the post 1991 treatment effect.

Results are solid and robust across specifications. Evidence supports the presence of a significant knowledge flow transmitted from ex-USSR to Europe and Israel. There is no evidence of a cohort effect. The first generation of former ex-Soviet migrants contributes to patent production almost as mush as the second one.

The rest of the paper is organized as follows. Section 2 briefly reviews the relevant literature on migration and innovation. Section 2.2 provides an overview of the main events that led to the 1991 breakdown (2.3.1), the USSR scientific system and its historical evolution (2.3.2), and a description of the Soviet IPR system (2.3.3). Section 2.4 describes the data and the collection methodology. Section 2.5 describes the sample and gives some descriptive over ex-USSR inventors' activity. Section 2.6 presents the empirical strategy while Section 2.7 gives an interpretation of results. Section 2.8 concludes.

2.2 Migration and innovation literature

Noteworthy studies (exhaustively reviews in Kerr 2013) posit immigrants as a crucial part of the US innovative and entrepreneurial processes. William Kerr documents the contribution to the US patenting system made by ethnic inventors: he estimates a decline in the contribution provided by Anglophone and European inventors, in lieu of an increased activity of Chinese and Indians (Kerr 2010, Kerr 2008). During the eighties and the nineties almost a quarte of Sylicon Valley ventures were run by Chinese and Indian entrepreneurs and executives (Saxenian 2002). Similarly, in 2006, the 25% of new US high-tech companies were started by immigrants (Wadhwa et al. 2007).

Such empirical literature, however, is rather heterogeneous in terms of testable hypotheses and migration episodes under study, mainly due to the lack of reliable and consistent sources of information, and to its multiple disciplinary origins (economic history, labour economics, innovation geography, among others; see Lissoni 2017). Still, three broad mechanisms linking migration to innovation can be envisaged:

- Diversity: migration increases the workforce diversity (ethnic or birthplace-driven) either in the host country or at the firm level, with such diversity being beneficial in terms of creativity and innovation through recombination (Alesina et al. 2016, Ozgen et al. 2013 ; see also chapters 1 and 3 in this dissertation).

- Positive self-selection: a large tradition of migration studies suggest positive self-selection to be the dominant sorting mechanisms between migrants and stayers in origin countries. It is the best individuals in terms of skills and knowledge-base who get the highest rewards by relocating (Orrenius and Zavodny 2005; Chiquiar and Hanson 2005; Chiswick 1999). These individuals improve the human capital of destination countries.

- Knowledge transfer: theoretical propositions such as trans nationalism and diaspora knowledge spillovers suggest that migrants may carry with the them specific knowledge either absent or in great demand in destination countries. The diffusion of knowledge embodied in high-skilled, foreign-born individuals has been addressed from the point of view of both the country of origin of migrants (Dos Santos and Postel-Vinay 2003, Kale et al. 2008, Wang 2015, Hornung 2014).

As for specific migration episodes from one country of origin to a region or country of destination, the second and the third mechanisms have attracted most of the attention. Coming to self-selection, Hunt 2011 finds that differences in patenting rates and entrepreneurial success between migrants and locals can be almost entirely explained by choices concerning their field of study or, more generally, their educational careers. Likewise in Kerr and Lincoln 2010 USPTO patents filed by Anglophone and non Anglo-Saxon inventors are found to be of similar quality (where quality is measured with the number of claims of each patent).

Such results suggest that foreign inventors are neither more productive than native ones, nor look like technological leaders. As such, they are in contrast with Stephan and Levin 2001, where immigrants are found to be over represented in the right tail distributions of a set of research-related performance indicators. As for the third mechanism (migration and knowledge diffusion), the historical literature has produced robust evidence concerning well-known cases of refugee migration from more to less technologically advanced countries. Examples include the migration of French Huguenots to Prussia in the XVII century (Hornung 2014) and that of German chemical scientists of Jewish origin to the US (Moser et al. 2014).

Several studies on the migration of Russian scientists to the US, after the collapse of the Soviet Union, follow in this tradition, both in terms of contents and methods. Ganguli 2015a instead focuses on Soviet immigrants in the US as a channel of knowledge diffusion among native scientists: using data on scientific publications, she documents a substantial knowledge flow toward US scientists. The strongest effects are ascribed to specific fields such as physics and life-sciences. Additionally, as expected ideas already accessible as codified knowledge are found to be more easily transmitted.

Borjas and Doran 2012a document instead a strong crowding-out effect for mathematical disciplines. After 1991, those American mathematicians whose specialization mostly overlapped with that of Soviet migrants suffered, on average, a reduction in productivity. They became also more likely to switch institution and found more difficulties publish in top-tier journals. Borjas and Doran do not deny the existence of hypothetical beneficial effects of Soviet mathematicians in terms of generation of new ideas; but claim that such positive externality could be outbalanced by other field-specific dynamics. The literature on destination countries other than the United States, as stressed in Nathan 2014b, Lissoni 2017 and Breschi et al. 2016, is much more limited, and circumscribed to firm-level evidence ozgen2013impact(Østergaard et al. 2011, Parrotta et al. 2012). The link between foreign inventors and patenting activity is explored in Nathan 2014a: for UK-based patents, inventors belonging to small ethnic groups are found to be spatially clustered, and the ethnic diversity of inventors teams is positively associated with individual patenting activity. Franzoni et al. 2012, using data from the GlobSci survey, find that migrant and returnee researchers operate through a larger network with respect to natives; moreover, the sizes of such migrants networks are positively correlated with the level of development of their countries of origins science base.

The hypothesis of a migration-innovation link through the mechanism of knowledge diffusion has been widely investigated by studies focusing on the countries of origin of migrants. Such studies focus on the the potential losses encountered by migrants homelands (the so called brain-gain effect), and on the compensating effects that knowledge feedback from the destination countries may produce/ Early studies (Bhagwati and Hamada 1974 and McCulloch and Yellen 1977 among the others) posit a welfare loss for the source countries, while more recent contributions tend to favor the hypothesis of a positive net effect: Mountford 1997 shows that an optimal level of emigration rate of skilled individuals exists, and under certain conditions (i.e. in case of endogenous human capital accumulation) such phenomenon may in principle increase long-run productivity and income equality of the source economy. Similarly Beine et al. 2008 finds that in countries with relatively low levels of human capital and low skilled emigration rates, brain drain produce net positive effects.

Related to the above findings, in Miguelez 2016 cross-country collaborations between developed and developing countries are found to be associated with the size of the skilled diaspora (measured as the number of fellow foreign inventors patenting in the same state) active in the host country. In Agrawal et al. 2011 the bulk of the analysis is restricted to the Indian diaspora of inventors. The main conclusion points to a detrimental effect of high-killed emigration: knowledge access for Indian residing inventors is undermined by the absence of fellow Indian expatriates ones.

2.3 Historical context

2.3.1 The breakdown of USSR

On December 26, 1991 the Union of Soviet Socialist Republics (USSR) was declared dissolved, with twelve former Soviet Republics forming the Commonwealth of Independent States (CIS)[4]. The world's largest country, built by the Russian Empire on the ashes of the 1917 Bolshevik Revolution had already started to fade away in the late 1980s, with the fall of the Berlin Wall and the constitution of the first political movements in the Baltic States, which eventually obtained independence in 1990 (Strayer 1998, Suny 1993).

Lithuania was the first republic to achieve its own independence in March 1990, after several months of bloody conflicts with Russian troops in the streets of Vilnius. A month later, the Supreme Council of the Republic of Georgia declared its own (although such declaration became effective only a year later).

After the 1991 breakdown the regime in the CIS and in the newly formed Baltic states rapidly shifted toward a more liberal regime, allowing for an unprecedented movement of people and goods toward Western regions (Sakwa et al. 2005).

Before the early 1990s, emigration from USSR was severely limited. Just a few ethnic groups, among which Jews and Germans, were allowed to leave in order to reunite with their families in the West. The Federal Republic of Germany provided asylum to ethnic Germans from the Soviet Republics with its 1949 Basic Law. Between 1950 and 1987 almost 2.6 millions German emigrated from Soviet Republics toward Germany(Dietz 2000, Kurthen 1995, Gokhberg and Nekipelova 2002). Starting from 1990, the Iron Curtain began to dissipate and emigration restrictions became less and less severe: official statistics set the number of Soviet emigrants for the 90-91 biennium at about 450,000.

The majority of Jewish migrants settled either in Israel or in the United States. The so called Post-Soviet Aliyah (literally *ascent*, i.e. return to Israel) began in the late 90's.

[4]CIS members: Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine and Uzbekistan.

Between 1989 and 2006, more than 1.5 million Soviet Jews left the former Soviet Union: of them, the 60%, migrated to Israel, while the rest were almost equally displaced between the United States and Western Europe (Tolts 2009, Latova and Savinkov 2012).

From a comparative point of view, Western countries other than Israel and West Germany registered much lower figures in terms of Soviet arrivals. Nonetheless, several narratives are well documented: the United Kingdom was one of the countries that benefited the most from the "elite" East-West migration during the 1990s, when many highly-educated professionals and entrepreneurs migrated from the former Soviet Union, and gave birth to a huge Russian-speaking diaspora (Pechurina 2017). Similar episodes involved the Netherlands, Greece, France and Portugal (Nicolaas and Sprangers 2001; Kopnina 2005; Voutira 2006).

2.3.2 Science and technology in the Soviet Union

This work highlights the role of Soviet scientists and engineers who emigrated from their homeland and were engaged in innovation activities in Western Europe and Israel. This makes it necessary to overview, albeit briefly, the scientific environment that trained these researchers.

The Soviet Union science system was a highly centralized one. Scientists were assigned by both the Central Committee (the main ruling body of the Communist Party) and the Academy of Sciences to academic institutes and R&D facilities (usually military-oriented), where they conducted research that was in line with the government agenda. With science and basic research being strongly subordinated to industrial and governmental needs, such system performed extremely well for some highly-prioritized objectives (i.e. hydroelectric power plants, atomic weapons and the space program, along with major advancements in fundamental areas such as mathematics and theoretical physics), but really poorly for others like computer and genetic engineering (Graham 1993, Fortescue 1986, Soyfer 2001).

By the end of the 1990s, the defense industry alone was responsible for almost the 20% of the total industry employment, with over 1.5 millions researchers assigned to military-related R&D. These scientists had access to relatively high wages, top quality equipment and almost unlimited financial resources for almost thirty years (Cooper 2013).

The collapse of the Soviet Union brought harsh changes. The vast demilitarization that occurred across the newly formed independent states, and the concurrent currency crisis produced a dramatic reduction in projects funding by the government (Dabrowski 2016). The effects of such budget reduction were twofold: former soviet researchers suddenly found themselves both with limited resources (such as foreign machinery, chemical reagents and periodical subscriptions) and much lower wages[5] (Saltykov 1997, Ganguli 2015b, Graham 1993). As a consequence, many of them decided to move to Israel, Europe and the United States. The phenomenon was remarkable, especially in the early 1990s, although accurate estimates are not yet available.Graham 1993 asserts that by the end of 1991, Israel received almost 6000 researchers from former Soviet Union Countries. Gokhberg and Nekipelova 2002 makes use of data provided by the Ministry of Internal Affairs of the Russian Federation, and evaluate the number of employees of the sector "Science and Scientific Services" that emigrated from Russia in the period 1990-2000 at about twenty thousand. The focus of the subsequent analysis will consist of several West European countries and Israel, as the recipients of such influx of scientists.

The repercussions of the exodus of Soviet science workforce on destination countries have been studied by, among others, Borjas and Doran 2012b, Paserman 2013 and Ganguli 2015b. The contexts under study, together with their results, are different.Borjas and Doran 2012b focuses on the net effects generated by the arrival of Soviet mathematicians in the US on their fellow natives. Using publication data, they describe multiple negative responses: for American mathematicians whose research fields lie over the Soviet ones, productivity decrease. In addition, they experienced a higher likelihood of switching between institutions (i.e. higher mobility) as a response to increased competition, and they also turned out to be less likely to produce a "home-rune" journal article.Paserman 2013, on the other hand, provide details on the Israelian manufacturing sector from 1970 to 1999. Higher concentrations of immigrants from former Soviet countries are found to be uncorrelated with differences in firms productivity. On the contrary, results suggest a negative association between such quantities with respect to low-technology sectors.

2.3.3 The Intellectual Property Rights system

Along with the progressive forces that invested the former Soviet countries after 1991, Russia adopted on September 23, 1992 the New Patent Law. Before that date, no such thing as private IP existed in Soviet Union: patent protection was technically available, but the main legal document conceived by the government was the Certificate of Authorship, which granted full ownership of the invention to the State (Feldbrugge et al. 1985). Any Soviet company could exploit the content of a Certificate, and the holder was usually compensated with prestige, career advancements and conspicuous monetary prizes (Graham 1993).

Certificates, contrary to patents, did not require the payment of any kind of fees, and

[5]In Russia, during 1992, the average salary of scientists was approximately the 65% of the national average, and researchers found in many cases that their remunerations were lower than the one of Moscow's taxi driver.

were indefinitely valid, while the standard duration of a patent was set to fifteen years[6]. Moreover, disputes over Certificates were allowed only within a year after the certificate was issued, and even though disputes over certificates were debated in ordinary judicial courts, the government had the privileged right to intervene whenever its interest was a stake, and to legislate over the court.

Such features determined the almost total absence of Soviet patents: from 1970 to 1975 only four patents were issued to Soviet inventors, compared to most than 200,000 Certificates (Portnova 1998, Soltysinski 1969).

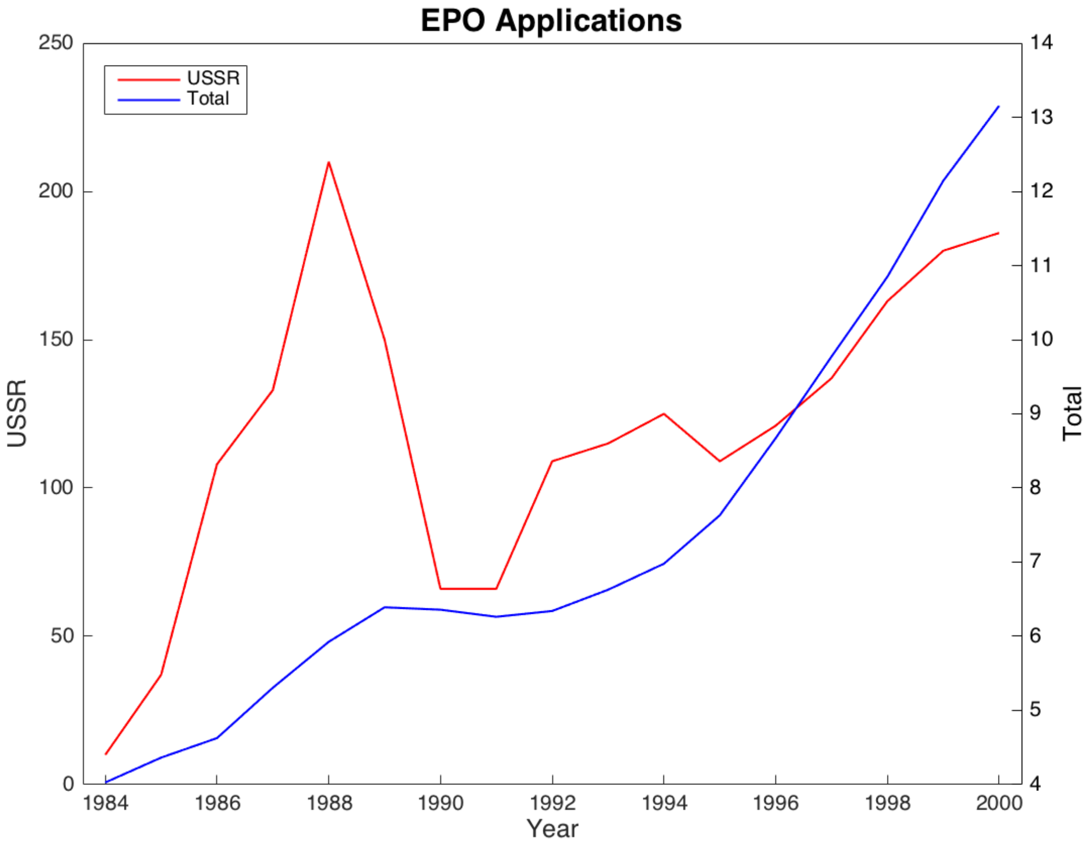
The New Patent Law made patents the only legitimate way to protect inventions in Russia. The reform moved the Russian system much closer to Western standards: since 1992 inventors could make profits on their discoveries for the entire life-time of their patents (usually twenty years), while the state was no longer the main owner of such inventions. The reform of the IP system was absolutely necessary in order to be admitted to the World Trade Organization, although the formal acceptance occurred only in August 2012. Several additional adjustments, mainly in terms of prosecutability for patent infringement, were necessary to finalize the process.

The regime shift can be observed in Figure 2.1. Since 1992 the number of patents with at least one Russian-based applicant filed at the European Patent Office, starts to recover after a massive decline occurred since 1989. Such decline can be mainly ascribed to factors highlighted in the previous Sections: a severe reduction in public spending, and the subsequent emigration of many scientists (that comprised as well a lot of Soviet inventors) toward West.

Still, post-1991 records of Soviet-based patent applications filed at the European Patent Office can serve to identify the main characteristics of inventive active of Russian and CIS inventors who were operating in their homeland after the Soviet Union breakdown: a comparison can provide insights on possible differences in patenting rates and on the main fields of activity.

[6]As Portnova 1998 reports, before the New Patent Law patents were almost entirely issued to prompt foreign companies to apply for Soviet legal protection[7]. After the Soviet Union became a contracting party of the Paris Convention, the number of foreign applications in USSR grew massively.

Figure 2.1: EPO patents by USSR applicants.



Notes: Patents with at least one Russian-based applicant filed at the European Patent Office on the left y-axis, thousands of EPO patents on the right y-axis.

2.4 Data and sample

This section is devoted to the description of the main data sources employed, and to the characterization of my observational units.The database used in this paper is constructed combining two main sources: the CRIOS-Patstat database (CPDB, Coffano and Tarasconi 2014) and the WIPO database (WDB).

The CPDB consists of patent applications filed at the European Patent Office, whose information contents have been harmonized over the years. A crucial advantage of the CPDB is that inventor are disambiguated (see Pezzoni et al. 2014). This is, to the best of my knowledge, the only comprehensive disambiguation process of inventors made for the entire collection of EPO applications [8].

The WDB, on the other hand, contains information on patent applications filed through the PCT route. These documents, thanks to a combination of patent offices legal rules, grant the nationality of each inventor for all applications with priority date

[8]Recently, Morrison et al. 2017 provided an alternative disambiguation strategy: it relies on geolocation data, and it performs well for both inventors' and assignees' names, although the analysis is limited to fewer applications with respect to the CPDB.

before 2011 and an extension to the US. Additional technical details can be found in Miguelez et al. 2013.

As anticipated, I will track a set of technological classes over time, and estimate their advancement levels as a function of the presence of inventors from the former Soviet Union. Accordingly, the unit of analysis is the single technological class, as defined by the International Patent Classification (IPC) classification scheme. The IPC is a taxonomy developed and supervised by WIPO, in accordance with the terms of international treaties dating back to 1971 (WIPO 2017). The first version of the IPC scheme dates back to 1968. Since 2010 the classification system is revised every year, while the current hierarchical structure was introduced with the major reform of 1999 (Wongel 2005). It is important to notice that most patent files are reclassified whenever a new edition comes into force. However, in order to exclude potential distortions coming from new technologies observed with new IPC classification editions, I will test my results imposing conditions on the years of creation of technological classes.

Some authors could argue, in general, that the way in which inventions are assigned to technologies is crucial. as other classification systems are available (Schmoch 2008). The Cooperative Patent Classification (CPC) system is an international standard jointly developed by the EPO and the USPTO; it contains all the IPC symbols plus a section dedicated to emerging technologies. I performed a manual search on Espacenet, and I noticed that the reclassification of old patent documents is still ongoing. Specifically I found that several PCT documents of the 1990s are assigned to fewer CPC codes than the IPC ones; this should not be the case given that CPC codes are way more numerous than IPC ones. In the end, given that a consistent share of my observations referred to EPO-PCT patents whose priority lies between 1990 and 2000, the IPC seems the best choice.

My observational units will be technological classes. One of the broadest and more recent definitions of technology is provided in Arthur 2007. He simply defines a technology as *a mean to fullfill a purpose*[9]. Besides, technology are defined as *recursive* objects: a large technology such as an hydroeletric power plant is made of smaller, functional components (i.e. turbines, generators, transformers etc.) that are in turn technologies themselves. Referring to the IPC taxonomy, if we exclude the broadest classification level, which breaks down the invention space into eight heterogeneous Sections, all the remaining level consists of elements that fit more than adequately into the definition of Arthur 2007.

I choose to operate at the fourth hierarchical level, namely the Main Group. To clarify on this point, Table 2.1 presents the IPC classification scheme for "Fast fission reactors", indicated by the full Subgroup code G21C1/02. As mentioned the level of observation will be the fourth one (which, in this case, corresponds to the Main group "Reactors", coded

[9]Refer to Arthur 2009 as well.

G21C1/00). My way of grouping patents based on their Main group classification results in grouping together patents that pertain to "Fast fission reactors" (G21C1/02), "Thermal reactors" (G21C1/04) and "Integral reactors" (G21C1/32), while mantaining separated reactors-related patents (G21C1/00, "Reactors") with patents that deals with shielding and protection items for nuclear facilities (G21C11/00, "Shielding, structurally associated with the reactor"). In order to rule out any possible threat of grouping-dependency, as part of my robustness analysis, I will construct alternative technological classes based on the Subclass definition and check the consistency of results.

Table 2.1: IPC classification of fast fission reactors.

Level	Code	Content
Section	G	Physics
Class	G21	Nuclear Physics; Nuclear Engineering
Sublcass	G21C	Nuclear Reactors
Main group	G21C1/00	Reactors
Subgroup	G21C1/02	Fast fission reactors

Notes: Figures refer to unique ex-USSR PCT inventors active over the period 1991-2000. *West-EU* encompasses EU15 countries, Switzerland and Israel.

The level of advancement of each class is expressed via patent production. The main downside of this approach, is that patent records cover just a portion of innovation activities, given that figures relative to R&D expenses in equipment, trade secrets etc. are not considered. Still, a similar approach has extensively been adopted in literature (Acs et al. 2002, J Acs and Audretsch 1989).

I will compute the yearly number of PCT patents signed in Europe and Israel for a period that ranges from 1980 to 2000. The geographical localization of patents will be made through CPDB inventors' addresses. The countries that I will consider in the subsequent analysis are Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Great Britain, Finalnd, Austria, Sweden, Switzerland and Israel. The list includes, besides Israel and Switzerland, all the countries that were members of the European Union in 1995.

Since my goal is to estimate the effect of migrant ex-USSR inventors on local technological advancement, I will only consider contributions (i.e. patent applications) made by local inventors. Both the local and the ex-USSR attribute are assigned based on the nationality record found in the WDB. Specifically, an inventor will be flagged as ex-USSR if his/her nationality belongs to one of the CIS states. Local inventors, on the other hand, are the ones whose nationality coincides with their country of residence.

The next section briefly describe ex-USSR inventors' activity in the chosen macro-region (i.e. EU15 plus Israel and Switzerland) up to 2000, and provides some descriptive on the available technological classes.

2.5 Descriptive evidence

Migrant inventors from the USSR were absent from patenting activities in Western countries before 1991, mainly due to emigration restrictions imposed by the Soviet government and to its peculiar IPR regime, as I highlighted in the previous Section. In contrast, during the first decade after December 26, 1991 almost 5000 citizens from the former Soviet Union countries have been involved in patenting activities. Of those, about the 70% were active in Russia (or in another CIS member state), while the others (respectively about 600 and 700 individuals) were residents in Western Europe, in the United States and other countries outside the former USSR boundaries[10].

The composition of destination countries of ex-USSR inventors throughout the 1990s is the following. The ones active in their homeland (that is Russia or one of the CIS countries) were about 4000, and accounted for approximately the 70%. The relative share of EU15[11] based individuals was 13% (723 inventors): the ones who were active in the US accounted for almost the 16%. As reported in Tab. 2.2, for both destination areas the majority of ex-USSR come from Russia. While there are substantial inflows from Ukraine as well, the number of inventors that come from the remaining CIS countries is limited.

Table 2.2: Number of ex-USSR inventors by country of origin and area of destination.

Country of origin	West-EU	United States
Russia	505	701
Ukraine	120	110
Belarus	21	19
Estonia	17	9
Moldova	13	3
Georgia	12	10
Armenia	12	7
Latvia	12	9
Lithuania	10	10
Uzbekistan	1	3

Notes: Figures refer to unique ex-USSR PCT inventors active over the period 1991-2000. *West-EU* encompasses EU15 countries, Switzerland and Israel.

This image is consistent with the one outlined in Graham 1993: one has to bear in mind that after 1991, within the galaxy of newly-formed independent states, the Soviet centralism was replaced by new types of authoritarian regimes. During early 1990s emigration from countries such as Azerbaijan, Moldova, Georgia and Uzbekistan was prohibited, and the little that occurred was illegal and limited to intra-CIS routes (Shevtsova 1992).

[10] cfr. demoscope.ru

[11] That is EU15, Switzerland and Israel.

For the period under analysis (i.e. from 1980 to 2000), there is no evidence, among ex-USSR inventors, of serial patenting activity: each inventor signed on average 1.17 patents.

In terms of collaborations, ex-USSR inventors tended to patent mainly with German ones (more than half of the collaborations), followed by British, Finnish and French. Such evidence points at a potential measurement issue in my database. As stressed in Dietz 2000 and Brubaker 1998, the 1953 German Refugee Law was introduced to provide a homeland for ethnic Germans from Eastern Europe and the Soviet Union, who had experienced forced resettlement, ethnic discrimination and expulsion during and after World War II. Ethnic Germans who left Eastern Europe in the aftermath of the 1991 Soviet Union collapse, were granted permission to settle in Germany and were given German citizenship almost automatically. As a result, several German inventors (especially the ones who collaborated with ex-USSR ones) are possibly individuals who left the Soviet Union and gained German citizenship. This automatically translate in a possible underestimation of ex-USSR activity: as a robustness check I will exclude Germany from my sample and provide alternative estimates.

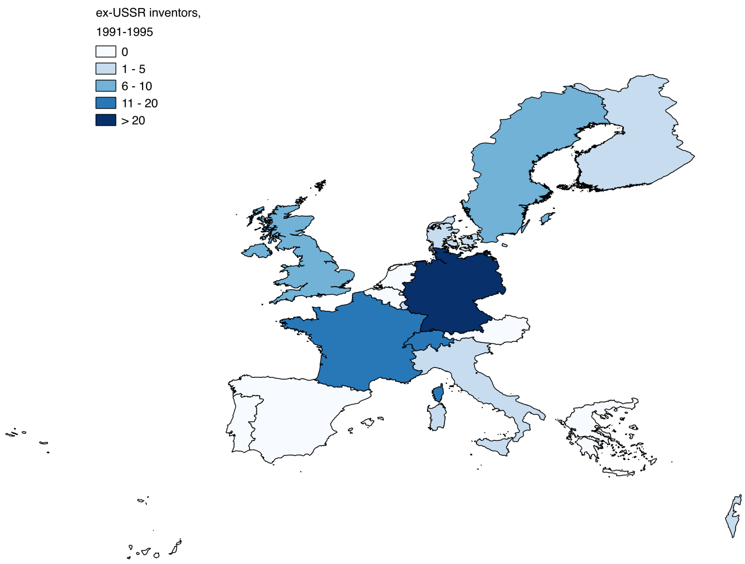
The central role played by Germany as a destination fro Soviet inventors is observed also in a more detailed geographical analysis . Fig.2.2 refers to the number of active ex-USSR inventors by country from 1991 to 1995: as we can see, Germany is the top targeted country, followed by France, Switzerland, Sweden and the United Kingdom. Within the subsequent five years (i.e. from 1996 to 2000, see Fig.2.3), the United Kingdom becomes a top destination together with Germany. Ex-USSR inventors' activity remains strong in Northern European countries, especially Finland and the Netherlands.

Fig.2.4 presents a comparison between local and ex-USSR activity by broad technological classes. The classification employed is the so called IPC35 proposed by Schmoch 2008 in order to overcome some of the limitations of another IPC-based based taxonomy (IPC ISI-OST-INPI) which dates back to 1992, namely the lack of some new technologies introduced with the 2008 revision, and a general imbalance that favors older, more consolidated technological classes.

As we can see, the classes in which ex-USSR inventors patented the most were Pharmaceuticals, Measurement (which contains several classes of Physics), Optics and Chemistry (specifically Organic Chemistry, Chemical Engineering and Biotechnology). Interestingly the distribution of ex-USSR inventors activity does not match that of local ones.

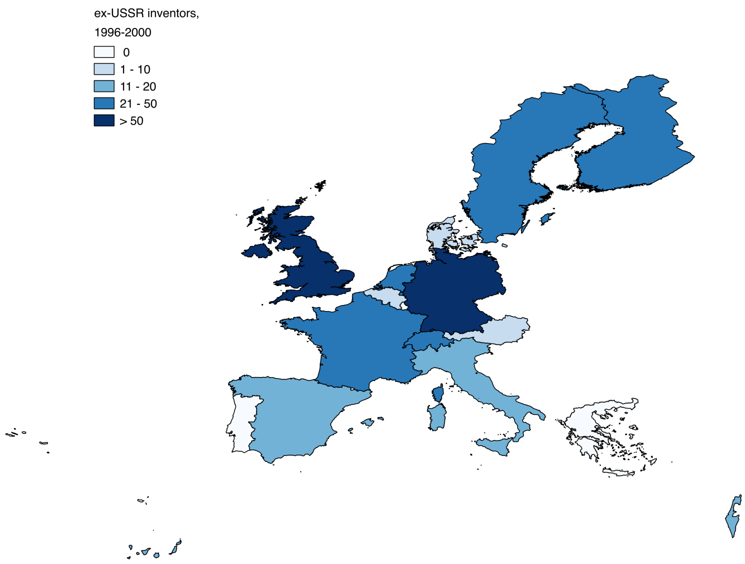
Last, I provide some evidence with respect to the level of maturity of technologies. In Fig.2.5 I classified the technological classes defined in the previous Section on the basis of their *birth year*, defined as the first year in which the first patent was granted in a technology class. The relative presence of Soviet inventors for the ten years that followed 1991 was highly concentrated in *Mature* and *Intermediate* technologies, with a

Figure 2.2: 1991-1995 patent production by country.



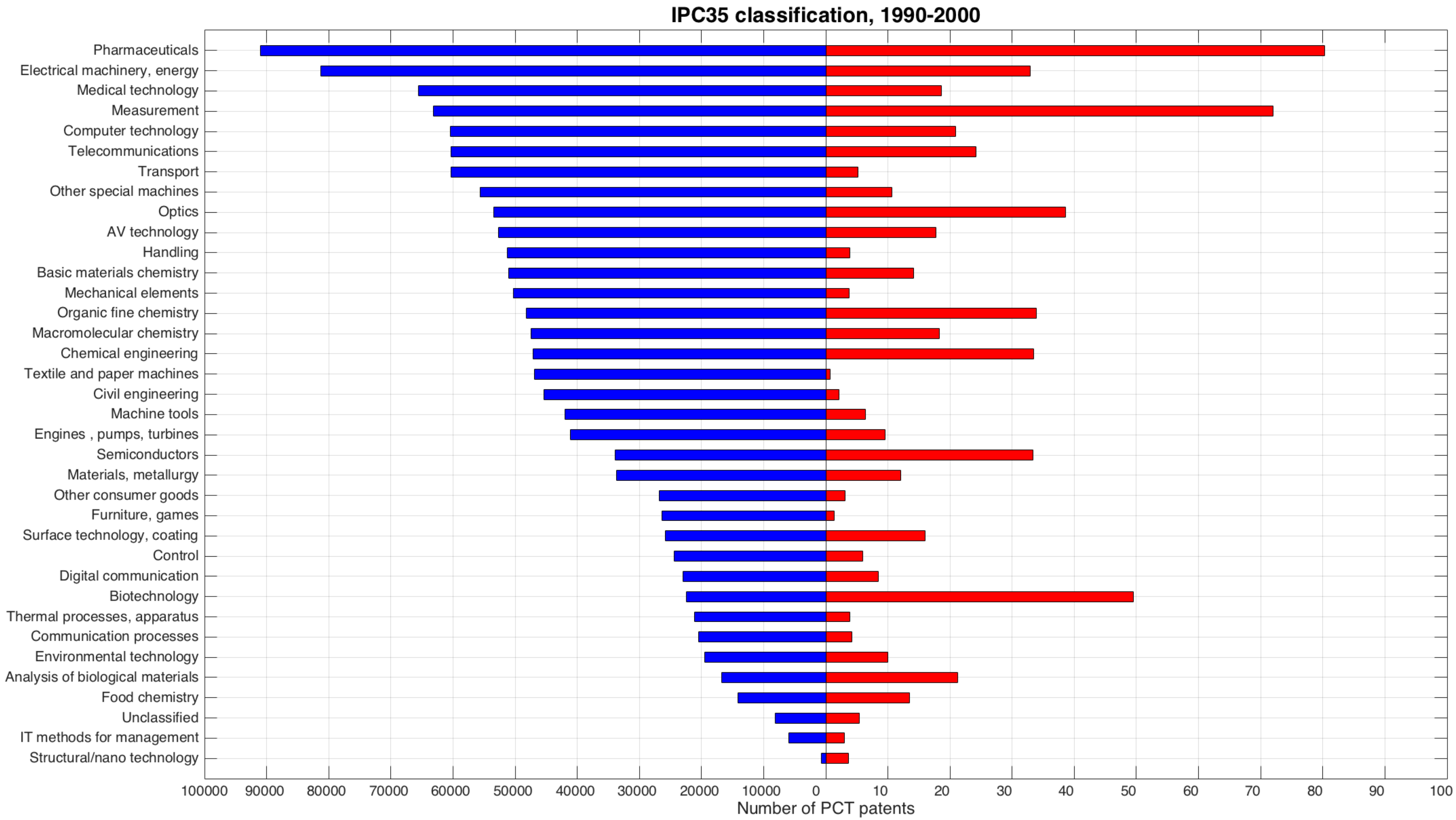
Different shades of blue indicate the number of ex-USSR inventors active between 1991 and 1995.

Figure 2.3: 1996-2000 patent production by country.



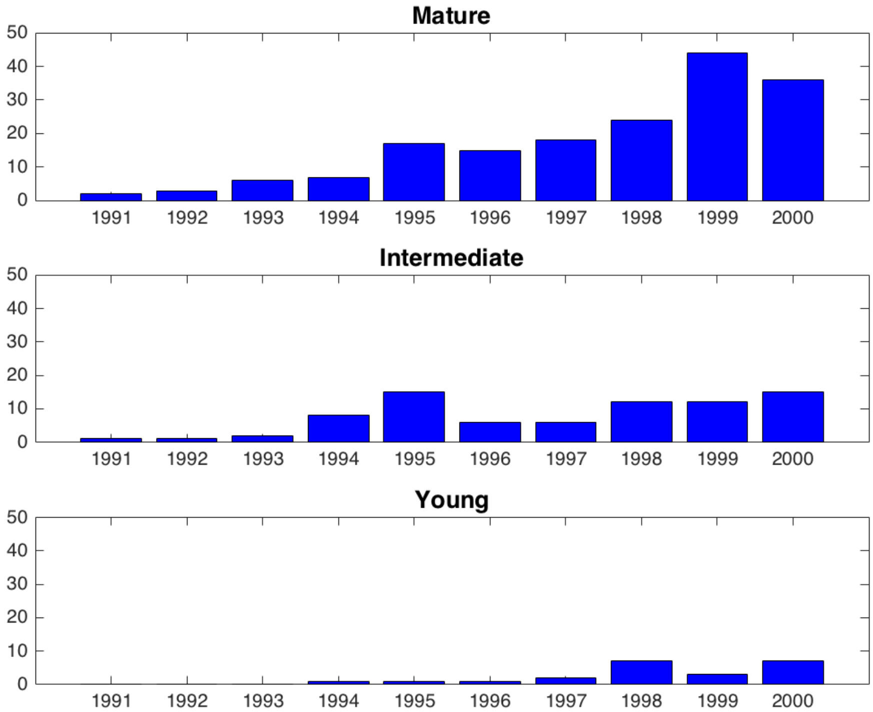
Different shades of blue indicate the number of ex-USSR inventors active between 1996 and 2000.

Figure 2.4: IPC35 classification



Notes: Counts refer to EPO applications with a corresponding PCT extension, signed by ex-USSR and local inventors (respectively in red and blue).

Figure 2.5: Yearly arrivals of ex-USSR inventors by age of the technological class.



Notes: Quantities refer to EPO application with a corresponding PCT extension, signed by inventors residing in one of the following countries: Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Great Britain, Finland, Austria, Sweden, Switzerland and Israel. Age brackets: Young is for technologies whose first applications was signed after 1985; Intermediate for technologies whose first applications was signed between 1981 and 1985; Mature is for technologies whose first applications was signed before 1981.

Table 2.3: Temporal distribution of technological classes' birth.

	IPC Subclass		IPC Main group	
Birth year	Freq.	Cum. Perc.	Freq.	Cum. Perc.
1977	167	5.0	299	5.3
1978	288	13.7	693	17.7
1979	304	22.9	659	29.4
1980	276	31.2	526	38.8
1981	276	39.6	529	48.2
1982	224	46.3	381	55.0
1983	178	51.7	285	60.1
1984	209	58.0	342	66.2
1985	201	64.1	305	71.7
1986	152	68.7	206	75.3
1987	136	72.8	211	79.1
1988	121	76.4	179	82.3
1989	128	80.3	198	85.8
1990	119	83.9	157	88.6
1991	120	87.5	157	91.4
1992	78	89.9	99	93.2
1993	70	92.0	84	94.7
1994	63	93.9	77	96.0
1995	55	95.6	65	97.2
1996	47	97.0	51	98.1
1997	39	98.2	41	98.8
1998	32	99.1	34	99.4
1999	12	99.5	14	99.7
2000	15	100.0	15	100.0

Notes: Subclass and Main group refer respectively to the third and fourth level of the IPC hierarchical classification system. I define the year of birth of technological class i as the priority year of the first patent application that has been assigned to technological class i.

really marginal contribution brought to *Young* ones[12]. A detailed overview of technology "births" by year is given in Tab. 2.3.

2.6 Empirical strategy

This study uses patent data to proxy the generation of innovative ideas. The empirical analysis intends to evaluate the impact of the arrival of ex-USSR inventors in the European and Israelian labor market on the innovative performance of local inventors. I estimate the

[12] *Young* technologies are the ones whose first applications was signed after 1985; *Intermediate* is for technologies whose first applications was signed between 1981 and 1985; *Mature* is for technologies whose first applications was signed before 1981.

change in patent production at the technological class with two alternative approaches.

2.6.1 Model 1: post-1991 evidence

The first procedure relies on ex-USSR inventors' patenting activity after 1991. I estimate the following regression:

$$Y_{i,t} = \beta_0 + \beta_1 exUSSR_{i,t-1} + \beta_x X_{i,t} + \eta_i + \eta_t + \eta_i * \eta_t + \varepsilon_{i,t} \quad (2.1)$$

The dummy variable $exUSSR_{i,t-1}$ takes value one if technology i benefitted from the contribution of an ex-USSR inventor in year *t-1*. The dependent variable $Y_{i,t}$ quantifies the patent production of each technological class that is attributed to local inventors (that is, inventors whose nationality and country of residence coincide). It consists of two alternative measures:

- *patents*: the yearly fractional count of patent applications by technological class. A single patent is usually assigned to multiple IPC Main Groups: for all EPO-PCT applications, the average number of Main Groups per patent is 2.7. The fractional count assigns to technological classes that are listed on the same patent an equal portion of it. For each patent application, only the share relative to local inventors is taken into account [13]. The log-transformation of the fractional count is used throughout the anlysis.
- *cit_patents*. This alternative dependent variable intends to control for quality differentials across patent applications. The above fractional count is weighted by the number of forward citations (computed at the DOCDB family level, see Martínez 2010) received within three years after the priority date. This variable is log-transformed as well.

Controls in $X_{i,t}$ include the age of technological class i, expressed as a second degree polynomial, and two variables (either categorical or continuous) signaling the presence of foreign inventors other than ex-USSR in class i and year t, and controlling for the quantity of local inventors in the previous year. The model includes a full set of year and technological class dummies, and it is estimated via OLS.

I hypothesize that a technological class that experienced the contribution of an ex-USSR expatriate after 1991 will, on average in subsequent years, perform better compared to the opposite circumstance of no ex-USSR inventors' input. The absence of

[13] For instance, if a patent is assigned to four classes, each one of them will be given 0.25 patents. Fractional count is widely used, see Brusoni et al. 2005, Perko and Narin 1997, De Rassenfosse et al. 2013. For each patent, I count only the portion signed by local inventors in the following way: If a patent is signed by three locals and one foreigner, I split 0.75 among the technological classes to which it is assigned.

patent records signed by USSR based inventors, on the contrary, forces me to identify my treated sub-sample solely with the variable $exUSSR$: such variable expresses the time variant shock that originates from the Soviet Union collapse, and it is likely to be endogenous. First of all, immigrants may not choose their technological destinations randomly, as unobserved factors correlated with the outcome of interest could lead ex-USSR migrants to enter labor markets, resulting in an biased estimate of β_1. To be more specific, the technological field of activity of ex-USSR expatriates after they settled in Europe (and Israel), was presumably chosen according to a combination of individual- and technological-specific attributes. While the former consisted essentially of educational background and previous working experience, the latter comprise several observable features. Additionally, the sudden influx of ex-USSR workforce may lead to displacement of natives as a consequence of increased labor market competition: as pointed out in Ganguli 2015a[14]

In order to address these endogeneity concerns I use a difference-in-differences analysis at the technology level, based on pre-1991 evidence, to provide causal evidence on the knowledge diffusion hypothesis.

2.6.2 Model 2: pre-1992 evidence

As noticed in Section 3.3, there is no evidence on the activity of ex-USSR inventors before they start patenting in EU15, Switzerland and Israel, making it impossible to preliminary identify the technological classes that, in a classical DID framework, should be flagged as "treated".

A standard approach would in fact be the one of selecting the technological classes in which Soviet based inventors were mainly active before 1992, and compare the patent production ascribed to local inventors in such classes with respect to the ones that were of little to no interest for the Soviet research program. To do so, I exploit the information provided by inventors' residences, as reported in patent documents. I identify all the PCT patents in which at least one inventor was residing in the Soviet Union and whose priority date is before 1992. This sample is made of 758 patents, corresponding to about a thousand technological classes. Given that most of them appear only a few times and in order to avoid overestimating the technological scope of pre-1992 Soviet activity, I produce a score that follows the fractional count principle highlighted in the previous sections. I assign to each technology listed on the same patent an equal portion of it. The final technology score is simply the sum of such portions, and I retain only those classes whose score is greater or equal than one, for a total of 294 classes. This "treated" sample accounts for the 4.9% of all the technologies present in the period under analysis.

[14]Although here the proposition refers to geographical dislocation, the same rationale can explain technological displacement.

$$Y_{i,t} = \beta_0 + \beta_1 USSR_i * post91_t + \beta_x X_{i,t} + \eta_i + \eta_t + \eta_i * \eta_t + \varepsilon_{i,t} \tag{2.2}$$

where the variable $USSR_i$ equals 1 if technology class i includes at least one patent with a Soviet address that was made before 1992; the indicator variable $post91$ equals 1 starting with the collapse of the Soviet Union. The dependent variable is either *patents* or *cit_patents* as in Eq.2.1). Control variables in $X_{i,t}$ are also the same.

The calendar year fixed effect η_t is included to take care of yearly fluctuation in patenting that affect technological classes in the same way, while η_i is a technology fixed effect which soaks up all unobserved time- invariant characteristics. Specific technology-year shocks are controlled for by the $\eta_i * \eta_t$ term. I estimate Eq.2.2 using OLS with robust standard errors clustered at the technology level. This takes into account the possible serial correlation in technologies outcomes over time.

In order to overcome the potential non-randomness of the treatment, I select a sample of control technologies that closely match the "treated" ones with respect to several dimensions: the number of patents filed per year at the EPO (with a corresponding PCT extension), averaged over the first five years of each class, and the number of active inventors per year, averaged over the same five years period. As an additional covariate, I include the year of first patent of each class. I match comparable technological classes using Coarsened Exact Matching (CEM, see Iacus et al. 2012 and Blackwell et al. 2009). The whole sample consists of 6335 technological classes and 294 of them experienced some ex-USSR contribution between 1991 and 2000: these units are matched to comparable classes for which such contribution didn't occur. The resulting sample is comprises 546 technological classes, given that to each "treated" class correspond a unique control class, and that some technologies are lost in the in the matching process.

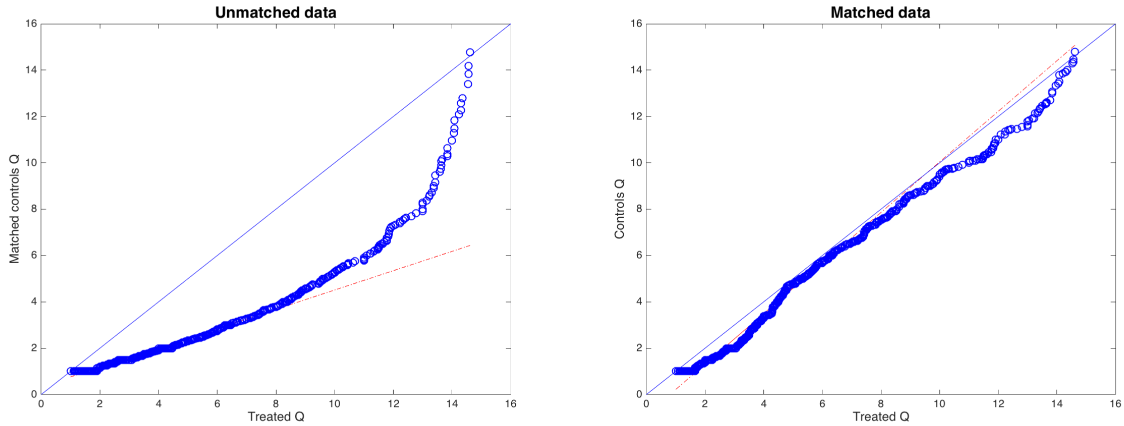
The CEM process reduces imbalance between treated and untreated classes along the four aforementioned covariates. Imbalance is measured by the L_1 statistics introduced in Iacus et al. 2008: such measure computes the absolute differences, in the quantities of interest, between treated and control units. The CEM algorithm produced a reduction of imbalance of almost 58%, where L_1 decreased from 0.52 to 0.22[15].

Figure 2.6 shows the imbalance reduction for the yearly patent production via QQ plots[16]: clearly the average patent production is disproportionately higher for the technological classes marked as treated. After the matching procedure, the two distributions are much more similar as the points lie around the diagonal. A similar pattern can be observed in Figure 2.7: however in this case, although the imbalance reduction is clear, for what concerns higher values the average number of active inventors is greater for the

[15]For a technical discussion of the CEM alghortim refer to Blackwell et al. 2009. Empirical contributions that employ a similar strategy are, among the others, Azoulay et al. 2011, Lewis et al. 2014 and Nall 2015.

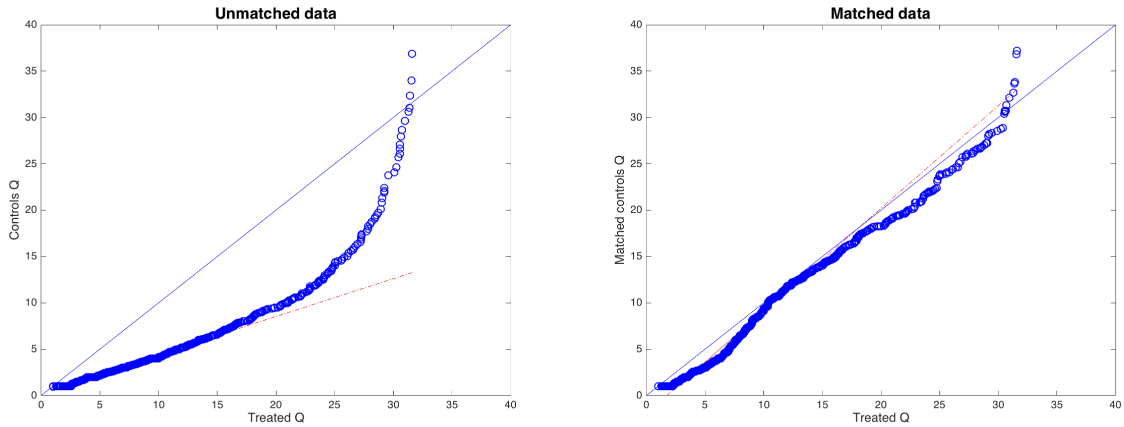
[16]A QQ plot is a graphical method of comparing two distributions by plotting their quantiles against each other.

Figure 2.6: Average patent production.



Notes: Comparison of QQ-plots before and after the CEM match procedure. The dashed red line is a linear fit. Treated quantiles on the x-axis, control quantiles on the y-axis.

Figure 2.7: Average active inventors.



Notes: Comparison of QQ-plots before and after the CEM match procedure. The dashed red line is a linear fit. Treated quantiles on the x-axis, control quantiles on the y-axis.

control group.

Overall the above results emphasize a significant reduction in the potential bias of $USSR$ coefficient. In the next Section I will discuss the main econometric results, referring to both the approaches outlined above.

2.7 Results

I first discuss the evidence relative to the descriptive model of Eq.2.1. The coefficients are reported in Table 2.4, where the dependent variable is *patents*. Column (1) relates to the whole set of technological classes, Column (2) to those classes whose first patent was filed before 1980, and Column (3) to the ones whose first application was filed before 1985 . The coefficients of $exUSSR$ range from 0.65 to 0.67. Such results holds even when I account for the age of the technological class: in Columns (4) to (6) I introduce a second order polynomial degree of class age (i.e. $Age2 = Age + Age^2$) and find very slightly lower coefficients. Given that the dependent variable *patents* is expressed in logs, a unitary change in *exUSSR* results, on average, produces a $(e^{\beta} - 1) * 100$ percentage change in the patent quantity made by local inventors: specifically, the activity of ex-USSR inventors for the whole set of technological classes (i.e. Column (1)) results in a 93% increase patent production.

Table 2.5 accounts for differences in the quality of PCT patents using forward citations at the family level[17] by making use of of *cit_patents* as alternative dependent variable. Coefficients are higher[18], but still with little variation across sub samples.

The coefficients of $exUSSR$ could be affected by omitted variable bias. In particular, the higher patent production may be the result of foreign inventors activity in general. In order to properly separate the two effects, in Table 2.6 I add the variable *non-exUSSR* $foreign_{i,t}$, a dummy that takes value one if technology i benefited, in year t, from the activity of at least one foreign inventor who was not flagged as ex-USSR. The coefficients of $exUSSR$ although slightly lower do not change much (whereas for the whole set of technologies, I have an 80% increase in the number of patent).

As for Model 1, I finally control for the actual number of inventors active between 1980 and 2000, rather than just including a dichotomous variable. In Table 2.7 I include three variables for the number of ex-USSR inventors, the number of locals, and the number of foreign, non ex-USSR ones (respectively *# exUSSR*, *# locals* and *# non exUSSR*), all expressed in logs and referring to $t - 1$. Results confirm the conclusions drawn from the previous analysis.

Issues about the non-random assignment of ex-USSR inventors in more productive, more populated technologies were discussed in the previous Section. I now turn to the examination of results pertaining to Model 2 (Eq. 2.2). Table 2.8 refers to the whole sample: patent production for the USSR-flagged technologies is, on average, more than double with respect to the non-USSR ones (a coefficient of 0.848 as in column (4) is equal to a plus 133% effect in the dependent variable).

However, non-random assignment may play a crucial role: ex-USSR inventors may be more likely to patent in highly-populated classes. In Table 2.9, the group of control technological classes consists of those successfully matched on the basis of the CEM procedure. The coefficients for $USSR * post91$ are still positive but lower: treated classes exhibit 62% more patents. As in the first specification, I control again for the total number of local and foreign (other than ex-USSR) which lowers the coefficient of interest to 0.233 (i.e. 26%, column (4) of Table 2.10).

[17]More on citations as an indicator of the quality of patented inventions see Trajtenberg 1990 and Moser et al. 2011.

[18]This difference in coefficients intensity, when patent count is weighted by the number of forward citations, is in line with Moser et al. 2014.

Table 2.4: Model 1.

	(1)	(2)	(3)	(4)	(5)	(6)
exUSSR	0.658***	0.655***	0.677***	0.637***	0.639***	0.654***
	(0.0365)	(0.0382)	(0.0298)	(0.0380)	(0.0280)	(0.0307)
Constant	0.312***	0.464***	0.350***	0.0874***	0.121	0.0712**
	(0.0170)	(0.0114)	(0.0121)	(0.0178)	(0.0622)	(0.0260)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Class dummies	Yes	Yes	Yes	Yes	Yes	Yes
Age2	No	No	No	Yes	Yes	Yes
Observations	61054	31491	51035	61054	31491	51035
Classes	6335	2191	4136	6335	2191	4136
R^2	0.398	0.450	0.409	0.412	0.455	0.423

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS specification including a full set of technological class, age and year dummies. Robust standard errors, clustered at the level of the technological class, are reported in the parentheses. The dependent variable is the yearly fractional count of patents signed by local inventors expressed in logs.

Table 2.5: Model 1, controlling for patent quality.

	(1)	(2)	(3)	(4)	(5)	(6)
exUSSR	0.841***	0.807***	0.864***	0.819***	0.790***	0.838***
	(0.0393)	(0.0486)	(0.0480)	(0.0437)	(0.0602)	(0.0524)
Constant	0.135***	0.268***	0.164***	-0.111**	-0.0821	-0.156***
	(0.0188)	(0.0156)	(0.0153)	(0.0344)	(0.0935)	(0.0396)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Class dummies	Yes	Yes	Yes	Yes	Yes	Yes
Age2	No	No	No	Yes	Yes	Yes
Observations	61054	31491	51035	61054	31491	51035
Classes	6335	2191	4136	6335	2191	4136
R^2	0.366	0.414	0.376	0.373	0.417	0.384

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS specification including a full set of technological class, age and year dummies. Robust standard errors, clustered at the level of the technological class, are reported in the parentheses. The dependent variable is the citation weighted yearly fractional count of patents signed by local inventors expressed in logs.

Table 2.6: Model 1, controlling for non-exUSSR foreign activity.

	(1)	(2)	(3)
exUSSR	0.604***	0.610***	0.623***
	(0.0420)	(0.0484)	(0.0463)
non-exUSSR foreign	0.298***	0.309***	0.305***
	(0.00607)	(0.00899)	(0.00710)
Constant	0.0914***	0.139**	0.0822**
	(0.0198)	(0.0488)	(0.0258)
Year dummies	Yes	Yes	Yes
Class dummies	Yes	Yes	Yes
Age2	Yes	Yes	Yes
Observations	61054	31491	51035
Classes	6335	2191	4136
R^2	0.428	0.468	0.433

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS specification including a full set of technological class, age and year dummies. Robust standard errors, clustered at the level of the technological class, are reported in the parentheses.

Table 2.7: Model 1, Controlling for previous inventor activity.

	(1)	(2)	(3)
# exUSSR	0.439***	0.389***	0.418***
	(0.0226)	(0.0306)	(0.0249)
# non exUSSR	0.142***	0.168***	0.153***
	(0.00348)	(0.00476)	(0.00409)
# locals	0.126***	0.151***	0.141***
	(0.00652)	(0.00742)	(0.00686)
Constant	-0.228***	-0.336***	-0.279***
	(0.0248)	(0.0382)	(0.0259)
Year dummies	Yes	Yes	Yes
Class dummies	Yes	Yes	Yes
Age2	Yes	Yes	Yes
Observations	61054	31491	51035
Classes	6335	2191	4136
R^2	0.487	0.536	0.508

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS specification including a full set of year dummies. Robust standard errors, clustered at the level of the technological class, are reported in the parentheses. The samples include: all the technological classes (1), the ones born before 1980 (2), the ones born before 1985 (3). The variable *# exUSSR* is the log count of ex-ussr inventors active in class i year $t-1$. *# locals* and *# non exUSSR* refer, respectively, to the log count of local and foreign (other than the ones from USSR) inventors.

Table 2.8: Model 2, USSR-based PCT technologies. Whole sample.

	(1)	(2)	(3)	(4)
USSR∗post91	0.897***	0.864***	0.863***	0.848***
	(0.0611)	(0.0573)	(0.0578)	(0.0574)
[1] Year dummies	Yes	Yes	Yes	Yes
[2] Class dummies	No	Yes	Yes	Yes
[1] ∗ [2]	No	No	Yes	Yes
Age2	No	No	No	Yes
Observations	61054	61054	61054	61054
Classes	6335	6335	6335	6335
R^2	0.153	0.207	0.223	0.226

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS specification including a full set of technology-year dummies (i.e. [1], [2] and [1] ∗ [2]). Robust standard errors, clustered at the level of the technological class, are reported in the parentheses. The variable $USSR_i$ equals 1 if technology class i includes a reasonable amount of patents with a Soviet address signed before 1992 (for a detailed description of the procedure, see Section 6.2); the indicator variable post91 equals 1 starting with the collapse of the Soviet Union.

Table 2.9: Model 2, USSR-based PCT technologies. Matched sample.

	(1)	(2)	(3)	(4)
USSR∗post91	0.505***	0.503***	0.547***	0.486***
	(0.0810)	(0.0755)	(0.0811)	(0.0790)
[1] Year dummies	Yes	Yes	Yes	Yes
[2] Class dummies	No	Yes	Yes	Yes
[1] ∗ [2]	No	No	Yes	Yes
Age2	No	No	No	Yes
Observations	7740	7740	7740	7740
Classes	546	546	546	546
R^2	0.221	0.335	0.375	0.392

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS specification including a full set of technology-year dummies (i.e. [1], [2] and [1] ∗ [2]). Robust standard errors, clustered at the level of the technological class, are reported in the parentheses. The variable $USSR_i$ equals 1 if technology class i includes a reasonable amount of patents with a Soviet address signed before 1992 (for a detailed description of the procedure, see Section 6.2); the indicator variable post91 equals 1 starting with the collapse of the Soviet Union.

Next, I check the robustness of my result across IPC Sections, identified with the first IPC classification digit. There is a total of eight Sections that span from Human Necessities (a quite heterogeneous groups, consisting of items such as Agriculture, Foodstuff, Tobacco and Wearing Appearel among the others) to Electricity. Table 2.11 reports the coefficients of Eq. 2.2 obtained with separate regressions for each one of them.

While there is no effect of ex-USSR inventors for Constructions, Mechanical Engineering and Electricity, previous results are mainly driven by Human Necessities, Operations/Transporting, Chemistry and Physics. That is consistent with anecdotal evidence regarding Soviet technological advancements[19].

Building on historical evidences of East-to-West migratory patterns, I last check for some cohort effect. Beside the considerable incentives for migration out of the Soviet Union resulting along with the economic downturn that occurred in the 1990s, reforms were made to adhere to international migration standards, and in 1993 the internal passport system was officially abolished. Several sources document a considerable rise in outflows rates between 1995 and 1996 (Buckley 1995, Heleniak 2002, Dietz 2000).

I check the variation of the coefficient of interest by years. Results of fig. 2.8 depict a rather flat trend. The coefficient relative to 1992 is almost not significant, which could point to the fact that the true outcome of migrant inventors on local patenting activity may have become observable with some delay. Possibly, ex-USSR inventors started engaging in new collaborations with their European peers across 1991 and 1992, and the knowledge transfer process took a while to produce some appreciable effect. Still the stationary of the coefficient is rather surprising, as one would expect a strong self-selection process for the first wave of migrants.

To sum up, the estimates partially go in favor of the hypothesis of knowledge diffusion: as explained, the Soviet Union was at the forefront in several scientific disciplines right before the 1991 collapse, while the 1990s were characterized by a progressive deterioration of the research and education systems Graham and Dezhina 2008.

[19] In fields such as physics (whose technological classes encompass the ones relative to mathematics as well), geology and soil science (that cover technological classes in both Physics and Human Necessities Section) and chemistry(Graham 1993).

Table 2.10: Model 2, USSR-based PCT technologies Matched sample. Controlling for previous inventor activity.

	(1)	(2)	(3)	(4)
USSR∗post91	0.249***	0.216***	0.232***	0.233***
	(0.0412)	(0.0371)	(0.0399)	(0.0409)
# non exUSSR	0.160***	0.164***	0.158***	0.159***
	(0.0177)	(0.0167)	(0.0185)	(0.0185)
# locals	0.390***	0.395***	0.397***	0.397***
	(0.0158)	(0.0140)	(0.0150)	(0.0150)
Age2	No	No	No	Yes
Observations	7740	7740	7740	7740
Classes	546	546	546	546
R^2	0.640	0.676	0.694	0.694

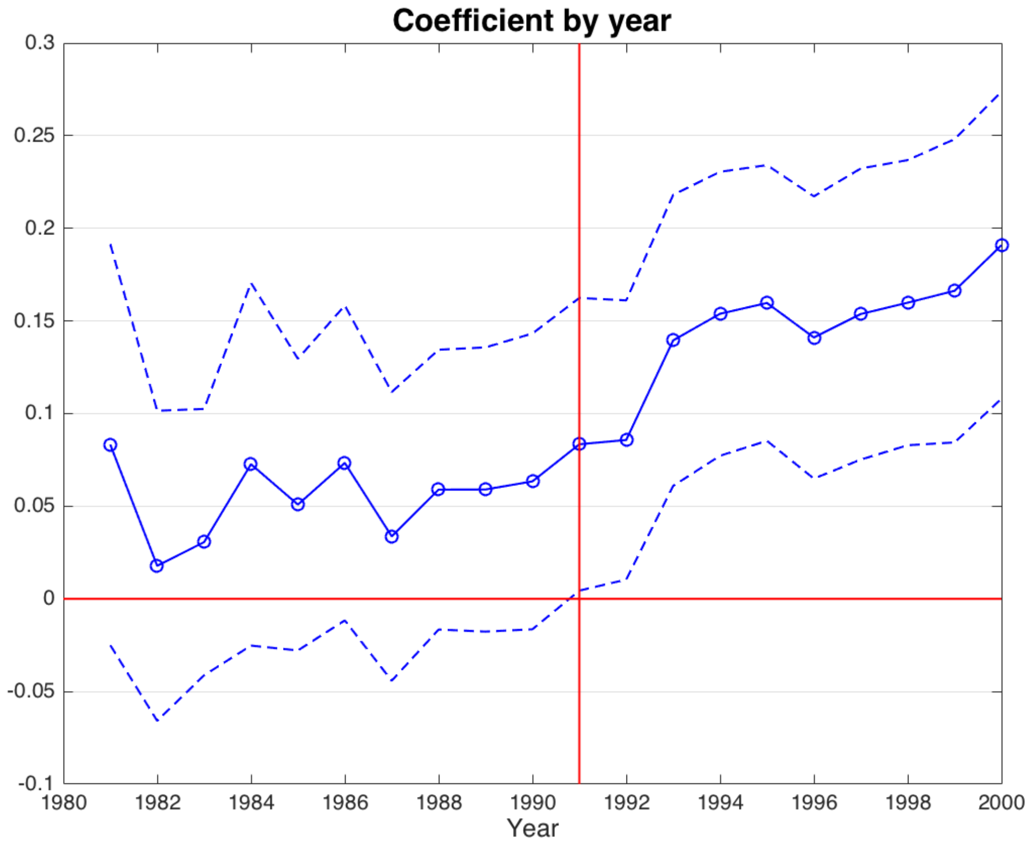
Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS specification including a full set of technology-year dummies (i.e. [1], [2] and [1]∗[2]). Robust standard errors, clustered at the level of the technological class, are reported in the parentheses. The variable $USSR_i$ equals 1 if technology class i includes a reasonable amount of patents with a Soviet address signed before 1992 (for a detailed description of the procedure, see Section 6.2); the indicator variable post91 equals 1 starting with the collapse of the Soviet Union.

Table 2.11: Model 2, USSR-based EPO technologies Matched sample. IPC Sections.

IPC Section	A	B	C	E	F	G	H
USSR∗post91	0.268*	0.219**	0.267***	0.217	0.135	0.123***	0.0863
	(0.125)	(0.0714)	(0.0742)	(0.128)	(0.0887)	(0.0320)	(0.128)
# locals	0.489***	0.373***	0.332***	0.358***	0.264***	0.522***	0.414***
	(0.0469)	(0.0268)	(0.0241)	(0.0540)	(0.0491)	(0.0319)	(0.0483)
# non exUSSR	0.0444	0.165***	0.228***	0.141	0.125*	0.224***	0.0950
	(0.0421)	(0.0383)	(0.0361)	(0.0719)	(0.0500)	(0.0504)	(0.0498)
Observations	1176	1747	1735	327	765	847	733
Classes	70	108	107	20	50	50	43
R^2	0.727	0.559	0.707	0.538	0.351	0.738	0.613

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS specification including a full set of technology-year dummies (i.e. [1], [2] and [1] ∗ [2]). Robust standard errors, clustered at the level of the technological class, are reported in the parentheses. The indicator variable post91 equals 1 starting with the collapse of the Soviet Union. IPC Section D (Textiles) is not reported due to the low number of observations.
IPC codes: A - Human necessities, B - Operations, Transporting C - Chemistry , D - Textiles , E - Constructions, F - Mechanical engineering , G - Physics , H - Electricity.

Figure 2.8: OLS estimates, annual coefficients.



Notes: The y-axis reports the coefficient β_t in the regression: $patents_{i,t} = \beta_0 + \beta_t EPOUSSR * year_t + \beta_x X_{i,t} + \eta_i + \eta_t + \eta_i * \eta_t + \varepsilon_{i,t}$. The reference year is 1980. Standard errors are clustered at the level of technological classes.

2.8 Concluding remarks

International migrants are maintaining a significant presence in both Europe and the US. A consistent and growing share of them is highly skilled and participate to the innovation process. Still we do not know much of the extent of their contribution (whether they increase the innovation rate of destination countries), nor of its type. In this latter respect, in particular, two hypotheses stand out: one, the diffusion hypotheses, suggest that highly skilled migrants may transfer knowledge along with them, from origin to destination; the other, the selection hypothesis, contends that highly skilled migrants are positively self-selected, up to the point when they hold superior skills compared to locals. In terms of knowledge production the effects for the receiving countries are still poorly understood, and the consequences for domestic workforce remains unclear.

Using a panel dataset of technological classes (identified according to the WIPO-IPC classification), I document the activity of Soviet inventors who arrived in EU15 and Israel after the 1991 collapse of USSR. The presence of many scientists and inventors among the Soviet migrating population persisted across the 1990s (Graham 1993), some of them being specialized in fields where the USSR lead internationally. This provides an interesting experimental setting to investigate whether a knowledge diffusion or a selection process was in place. Post-1991 evidence indicates that while they concentrated in developed technologies, ex-USSR inventors were beneficial to the patent production of locals. Additional results based on the pre-1991 patenting activity in the Soviet Union confirm the presence of a the knowledge diffusion mechanism, even after controlling for characteristics of technologies and year-specific technological shocks.

The increase in patent production attributed to local inventors is around 20%, and varies across broad scientific disciplines: higher coefficients are found for Chemistry- and Human Necessities-related technologies. I observe no strong cohort effect, as the average effects are constant over time starting from 1993. This piece of evidence goes against a self-selection hypothesis: migrant inventors who engaged in patenting activities across Europe and Israel may have not possessed above-average skills with respect to the entire pool of Soviet Scientists. Consequently results could underestimate the true effect: focusing on a clearly self-selected population of migrant scientists Moser et al. 2014 estimate a 60% increase in patents-by-technology produced by domestic inventors.

While contributing to the advancement of the literature on migration and invention, this paper suffers of limitations I hope to address soon. First, it lacks a detailed analysis of inventive activities in the Soviet Union before 1991, due to the difficulty of accessing the relevant archival data concerning the Certificates of Authorship, which we know to be the closest title to patents in communist Russia. Such data could improve the analysis under two aspects First the identification of top Soviet technologies would be more precise. Second, it would allow to identify the inventors who were active in their homelands before 1991. This, in turn, would give the possibility to track the individuals who moved out of the Soviet Union after 1991, thus observing directly the migratory event, instead of deducing it from the nationality of inventors or representing it by means of a time-related treatment.

Chapter 3

Internationalization of Business Research Teams: How Cultural Diversity affects Research Performance?

3.1 Introduction

Contemporary organizations increasingly build their strategies in relation to the business environment affecting their lines of activity. Even some organizations once operating in sectors characterized by a national focus are now embedded in a global network of exchanges and personal relationships. Higher education sector is a good example, as the trend toward corporatization (Kim 2009) and marketization (Brown and Carasso 2013; Molesworth et al. 2010) has led academic institutions to increasingly adhere to international business models of operation. This is especially true for business schools (including management departments within universities, whether public or private) which do not only strive to be financially successful and globally competitive, but increasingly pursue the mission to educate business leaders for multinational companies (Johansson and Śliwa 2014).

Accreditation bodies (e.g. AACSB, EQUIS) contribute to this trend by encouraging business schools to increase workforce diversity and promote international research collaborations. As a result, the cultural and ethnic diversity of the faculties has increased considerably (Kim 2009; Richardson and Zikic 2007). In the United Kingdom, for example, non-national academics constitute 24% of all faculty (Lambert and Usher 2013); in 2013, almost 3 million international students were enrolled in OECD countries' higher education institutes, with about 1.4 million students enrolled in EU institutions. Focusing on They made up, on average, for the 8% of total tertiary enrollments across the whole

OECD area, with higher proportions (between 16% and 18%) registered in Australia, Austria, New Zeland, Switzerland and the United Kingdom (OECD 2016).

Besides getting more and more internationalized, business schools are increasingly research-oriented. The scholars they attract collaborate within teams with both a diverse cultural background and a direct involvement in research and innovation activities,. They produce a truly globalized “business knowledge” (Maznevski and DiStefano 2000, Ravlin et al. 2000). Business research team members work across temporal and spatial boundaries, most often in the absence of face-to-face interaction, to coordinate their activities toward the attainment of common goals from different locations around the globe. Globalization and mobility, along with technological development, give rise to multinational and multicultural distributed teams (Connaughton and Shuffler 2007), transnational team (Haas 2006), or multinational workgroups (Hambrick et al. 1998).

Faced to these trends, we do not fully understand how internationalization reflected in business research teams affects research performance.

We investigate this issue by drawing upon diversity literature, with a focus on visible forms of diversity such as ethnic or gender diversity, and less emphasis on diversity across cultures, values and attitudes or even ideological orientations (Stahl et al. 2010).

Even though diversity is critical for contemporary organizations or teams to unlock their full potential, it is often unclear why some benefit from their diversity, yet others fail to do so. Drawing on a variety of theoretical and empirical grounds, this study extends prior research on diversity and performance by examining the conditions under which teams benefit from their diversity. Accordingly, we investigate when the relationship between cultural diversity and research performance becomes strengthened or weakened by considering its boundary conditions. In particular, we focus on the geographic dispersion of members within research teams: as it relates to the frequency of face-to-face interactions among team members, the degree of geographic dispersion is extensively acknowledged as a critical component of team performance (Zenun et al. 2007; Patrashkova-Volzdoska et al. 2003; Boschma 2005). In recent years major improvements in ICT resulted in a disproportionate increase of virtual teams (that is, teams whose members are obliged to collaborate across space using technology); latest developments conjecture that self-categorization may operate also at the geographical level, given that relationships among co-located members could profit from in-person communication as well as from close physical proximity in general (Polzer et al. 2006; Cramton 2001; Espinosa et al. 2003; Griffith and Neale 2001).

In particular, we test two hypothesis: whether a research team has published in top academic journals or not (Jensen and Wang 2017); and how many citations were received by the published article (forward citations) (Judge et al. 2007). We take the ranking of the journal as a measure of research ”quality”, while we consider the number of forward citations as a measure of research ”impact”. The main reason for distinguishing between the two relates to the different visibility of team diversity at different stages of the publications process. The publication of a paper depends chiefly on a blind peer review system, which obscures any information on the authors' ethnic and gender diversity to the reviewers. Hence, acceptance for publication ought to be based only on the intrinsic merit of the paper (quality). On the contrary, the scholars who read and decide whether to cite the paper have access to the ethnic and gender information, which may affect their decision in one way or another.

We offer three important contributions to the literature. First, we improve the understanding of the mechanisms and contextual conditions under which cultural diversity affects team performance both theoretically and empirically. Second, research team in management literature was is limited to contexts such as corporate R&D management. Previous studies have investigated how variance in scientific and technical teams' expertise enhances innovation outcomes (Singh and Fleming 2010) and contributes to higher rates of innovation breakthroughs (Fleming et al. 2007; Hargadon 2005). Whereas organizational research has focused on capital-intensive industries and therefore investigated their financial, technological, and market performance, this study differentiates itself by exploring the antecedents of research quality and research impact performance. Lastly, the context of our study does not only help us to test the relation between cultural diversity and team performance, but also understand how and when the notion of universalism and inclusiveness implemented in business academia bears fruit.

Next section presents a review of the relevant literature on diversity and team performance we capitalize on in order to develop testable hypotheses. This is followed by an explanation of research methodology and the empirical results. Finally, we provide implications drawn from the findings and future directions of the research.

3.2 Literature review

3.2.1 Diversity in organizations and its performance implications

Diversity is a characteristic of groups that refers to demographic differences such as gender, race, ethnicity, or nationality, all of which potentially contribute to a cultural identity that stems from membership in socio-culturally distinct demographic groups (McGrath et al. 1995).

Members of groups tend to share certain world views, norms, values, goals, priorities, and sociocultural heritage (Alderfer and Smith 1982; Ely and Thomas 2001). Their cultural markers can be realized through similarity in communication style, rules, meanings, and even language (Larkey 1996), which are shared within cultural identity groups, but differ across them, and are central to knowledge exchange processes.

In light of recent political events, and considering the migratory trends observed through the last decades, academic research in economics and management is inquiring the phenomenon to provide precise recommendations. Specifically, what can we expect from the growing ethnic mix with respect to innovation and knowledge creation? The literature of course is broad and ranges from large units of observation such as countries or regions (Ozgen et al. 2014; Alesina et al. 2016; Parrotta et al. 2014), to smaller subjects such as teams within organizations, which ultimately constitute the focus of this paper. At the macro level, recent literature (which is mostly limited to developed economies) overall surmises a benefit from ethnic/cultural diversity in terms of innovation. Ozgen et al. 2014 found a positive impact of ethnic diversity on innovation in Dutch and German regions. On the same line, Alesina et al. 2016 conclude in favor of a positive correlation, across European regions, between birthplace diversity and income, productivity and patenting rates. Bosetti et al. 2012, by means of panel of 20 European countries, find birthplace diversity to be positively correlated with patent production.

When we examine smaller observational units such as firms, results are far more mixed: from the point of view of Watkins and Ferrara 2005, ethnic diversity may boost productivity and/or innovation when it comes as a provision of new knowledge, or whenever the skills of foreigners integrate the ones of native workers (Ottaviano and Peri 2006; Trax et al. 2015). On the other hand, diversity may hamper innovation in light of cultural and linguistic barriers among a more diverse workforce (Parrotta et al. 2012). Lee and Nathan 2010 find that workforce diversity, measured for 2300 London firms, is positively correlated with both product and process innovation.

Bassett-Jones 2005 acknowledge a dual mechanism: on the one hand diversity can foster creativity and innovation, at the same time it can cause misunderstandings, low morale and loss of competitiveness.

Usually the aforementioned evidences hold across very heterogeneous firms. There has been a growing interest to medium and small-sized organizational units, primarily focusing on top management teams. This strand of literature mainly rests on upper echelon theories: choices made by top executives forge both characteristics and outcomes of their own organizational units; these decisions heavily hinge on individuals' observable experiences and psychological features (Hambrick and Mason 1984; Prahalad and Bettis 1986; Finkelstein and Hambrick 1996).

All the subsequent research that built on upper echelons theory surmises that perceptible demographic characteristics can be used as a substitutive measure of executive orientation, defined as a cognitive schema embodying one's personal ideology about human nature, viewpoints and perceptions of one's most proximate reality (Tetlock 2000), or to quote (Henderson et al. 2006, pag. 448), of "how the environment behaves, what options are feasible, and how the organization should be run". Bantel and Jackson 1989 analyze innovations (in-house developed new products, programs and/or services) in a sample of 199 US banks, as a function the heterogeneity stemming from differences in functional background and expertise. They find a positive correlation between diversity and the development of new items; also they find that more diverse teams are on average subject to higher turnover rates. Arguably this can be regarded as an instance of the dual mechanism described above: diversity stimulates innovation processes, but it also increases tension among team members due to communication struggles and inflated conflicts, , which result in lower stability of teams.

Innovation in corporate strategies is found to be enhanced by heterogeneity in educational specialization of top management team members for a random sample of 100 Fortune500 firms, supporting the theory that the interaction of diverse cognitive schemas promotes adaptation (Wiersema and Bantel 1992). Murray 1989 tested the relationship between firm performance, rather than firm innovation, and diversity, focusing on top management team members for a sample of 85 Fortune500 firms (limiting himself to oil and food companies). Interestingly, he diversifies between short and long term performance (where the latter is proxies by measures such as earnings to sales and earnings to total capital, while the former is computed as the ratio of stock price to earnings). His results support the hypothesis of a negative impact of diversity on short-term performance, while higher top management team heterogeneity results in better results in the long run (although such results are limited to the oil industry).

Broadly speaking, the impact of top management team diversity on organizational outcomes has been of great interest for management scholars throughout early 90's (Grimm and Smith 1991; Keck 1991; Hambrick and D'Aveni 1992; McLeod and Lobel 1992).

While multiple efforts failed to provide precise indications, but mainly led to mixed results, subsequent research has partially shifted the debate toward moderating factors, i.e. outer determinants that may potentially explain the two-fold behavior of organizational outcome terms of team diversity. Evidence of successful diversity in small teams has been provided by Hong and Page 2004 through a broad theoretical framework and Watson et al. 2002 by means of an empirical study. Communication barriers and lower workforce cohesion are the main arguments among the set of studies that claim an overall negative outcome in terms of produced innovation (Williams and O'Reilly III 1998; Zajac et al. 1991).

Most of the cited scholars now agree upon the intrinsic difficulty in consistently measure organizational units' performances; furthermore the presence of several co-founding factors usually makes it quite challenging to robustly estimate causal effects. Through a meta-analysis of most of the research conducted up to 2009, Joshi and Roh 2009 estimate the direct effect of diversity, once the role of contextual factors such as sector and size measures, supporting the "negative" view of diversity as a predictor of performance.

3.2.2 Business knowledge production landscape

The context of our study is business knowledge produced and diffused at a global scale. Business knowledge is usually produced by academic scholars and practitioners mainly through informal and non-market mediated mechanisms, mostly through team-based temporary collaborations (Wuchty et al. 2007). Scholars form research teams to take advantage of knowledge or information held by colleagues and students, within and across organizations, so to enhance the quality of knowledge they produce and secure multiple diffusion channels. Such teams capitalize on the specialized skills of their members by bringing together a group of people who are unfamiliar with one another skills, but must work interdependently on complex tasks (Bechky 2006).

Team-based research is not peculiar to business, as it is increasing across nearly all fields, which suggests that the process of knowledge creation has fundamentally changed. Wuchty et al. 2007 found that by the most recent period, a team-authored paper has a higher probability of being extremely highly cited which is the result derived from 19.9 million papers over 5 decades. Likewise, teams now produce exceptionally high impact research. Furthermore, the production of these intellectual artifacts in business and management research areas heavily relies on temporary team collaborations among the members originating from different cultural background. In fact, Financial Times Business School ranking takes into account the proportion of international faculty by calculating the diversity of faculty by citizenship and the percentage whose citizenship differs from their country of employment. Likewise, internationalization of business and management research activities is legitimized by authorities and accreditation bodies. Some studies show that knowledge production outputs carried out by internationalized research teams are more valuable and important than the outputs generated by research teams with homogeneous demographic in terms of national and geographic origin (Singh and Fleming 2010; Bercovitz and Feldman 2011).

A unique characteristic of research collaborations is their temporary nature, as they involve groups of people who are temporarily associated around specific tasks, after whose completion the team disbands and may or may not collaborate again in different compositions (Bakker et al. 2013). In the case of of business and management research, scholars collaborate until they publish their research outcomes in journal outlets, whose ranking we take as indicator of research quality (e.g. FT 45, UT Dallas).

We also investigate the legacy of such research outcomes in terms of forward citations, which measure the visibility and impact of a research. Citations received by each paper are considered in many previous studies as a measure of research visibility (Mingers and Xu 2010; Lee and Bozeman 2005) which is a metric for assessing the impact of an article and can also be used to evaluate the impact of research teams, academic institutions, universities, and journals (Judge et al. 2007).

3.3 Hypotheses development

3.3.1 Cultural diversity and team research performance

Cultural diversity involves differences in the knowledge structure of team members that reflects variability in values, cognition, and behavioral patterns (Earley 1993) as well as cultural differences in task-related experiences and task design (Dahlin et al. 2005; Durand and Jacqueminet 2015; Erez 2010; Hinds et al. 2011; Stahl and Tung 2015). Cultural diversity encompasses both deep-level diversity, such as values, norms, and beliefs, and surface-level ethnic diversity (Harrison et al. 2002).

As for explaining the role of cultural diversity in team performance, there are three theoretical streams of literature. First, according to *Similarity Attraction Theory*, people are attracted to working with and cooperating with those they find similar in terms of values, beliefs, and attitudes (Williams and O'Reilly III 1998). Second, according to *Social Identity* and *Social Categorization Theory* (Tajfel 1982), people tend to categorize themselves into specific groups, and categorize others as outsiders or part of other groups. People treat members of their own group with favoritism, and may judge members of other groups according to group traits (e.g., stereotyping). These first two perspectives suggest a negative effect of diversity on team performance, as diversity makes social processes more difficult. Such framework surmise potential communication and coordination failures, that results from increasing diversity: individuals tend to prioritize interactions and exchanges with members of the same category (Greenwald and Banaji 1995). We assert however that these two theories fit better to situations in which the workplace setting, and therefore therefore the choice of collaborators among colleagues, as well as the modality of interaction with them are pre-established by a higher authority, which both assemble the teams and decides on division of labour. Team members unsatisfied with the team composition have limited options to express dissent, which, paraphrasing Hirschman 1970, we can sum up as exit (undermining the team performance) or loyalty to the higher authority (collaborating as requested). Room for voice, such as proposing to reshape the team composition or renegotiating the division of labour within it, is scarce or non-existent[1]. Such a stark choice may not occur in the absence of an authority imposing both the composition of teams and its members' tasks, which is often the case for highly skilled individuals working in research.

There are two further reasons for resisting the hypothesis of a negative link between team diversity and performance, in the case of research teams. The first one pertains to the establishment/consolidation of personal relationships (that is, in a working envi-

[1]The impossibility to voice dissent is even more limited when the choice between exit and loyalty to the team implies a reverse choice with respect to the ethnic or cultural group, so that undermining the team performance is a sign of loyalty to the group, while collaborating as requested implies an exit from it (for example through ostracization.

ronment, to teammates' choice), not to the outcome that results from a collaborative process. As mentioned in (Adler and Gundersen 2007), the relation between diversity and performance heavily depends on the task assigned to the team. In the case of highly routinized assignments, usually marked by low levels of discretion and few interactions among group members, group outcome and performance should not significantly benefit from increased diversity. These conditions are extremely far from our observational units (business research teams), which, focused as they are on the production of novel contents, are exposed to trial-and-error processes and need to operationalize highly complex tasks.

Second, we can think of business scholars as of individuals whose training expose them to highly internationalized environments. In this sense, an alternative theoretical setting (i.e. *information-processing Theory*) is better aligned with the context of our study, as it argues that heterogeneous viewpoints, perspectives and to some extent personal values are found to stimulate not only group performance, but also commitment, satisfaction, intent to remain in the group and overall morale (Jehn et al. 1999). Although a detrimental effect of diversity in terms of workplace practices and operations may still occur, this should be outbalanced by larger information sets and higher levels of creativity.

A further element in favor of adopting an information process theoretical perspective for our case at hand, is the temporary nature of business research teams, whose composition changes over time according to the research interests of the group, the availability of funding, and personal career trajectories. Similarly ranked faculty across organizations collaborate as long as they find it useful for their careers and sustainable economically. PhD students and junior faculty have transient relationships with more senior researchers, before moving to other organizations or developing their own lines of research. This gives our specific group of highly skilled individuals ample latitude when it comes to building or entering a team, and negotiating tasks within them.

Taking into consideration the aforementioned arguments, and reshaping them to the context under analysis, we argue that the positive repercussions stemming from information theory will prevail on the negative ones, resulting in a net positive effect of the cultural diversity on two key measure of performance, namely the quality and impact of the team's scientific publications. We propose the following hypothesis:

H1.a Cultural diversity of business research teams is positively associated with their research quality (publishing in highly-ranked journals)

H1.b Cultural diversity of business research teams is positively associated with their research impact (receiving citations)

3.3.2 Geographic dispersion of teams

We have already stressed how few studies exist on cultural diversity and team performance, especially for highly skilled jobs. Even fewer studies have examined the mechanisms that may enable any positive effect of culturally diverse teams on team performance (Stahl et al. 2010; Van Knippenberg et al. 2004). Whereas previous studies drawing upon social categorization theory focus on the negative effect of diversity on performance, literature adopting information or decision-making processes perspective predicts and shows that diverse groups outperform homogeneous groups (Williams and O'Reilly III 1998). Our study intends to extend the previous contributions by identifying and examining the role of geographic dispersion on the relation between cultural diversity and research team performance. This is an important dimension to consider, as more research is needed to identify the mechanisms that activate task-relevant communication dynamics, that is the processes of sharing and understanding new pieces of knowledge that enable diverse teams to leverage their large pool of resources (Stahl et al. 2010). Due to our interest in cultural diversity as an outcome of international highly skilled migration, assessing the geographical dispersion of teams has also an operational value. In fact, geographically dispersed may not result from any migration process (such as when they result from the association of scholars who never moved into each others country, but met at meetings or through scientific societies or personal acquaintances) or result from it only indirectly (as when the scholars' acquaintanceship derives from past migration episodes of one or another team member).

Geographic dispersion measures the extent at which team members are spread across different locations. There are a number of reasons to presume it to interfere with the effects of diversity on team performance. First and foremost, dispersed teams are limited in the frequency and modes of interactions. They meet in person only once in a while and often communicate by e-mail, videoconferencing or other ICT devices, which hampers the decision making process and limit the possibility to appreciate in full the personality of other team members, including their cultural traits (Gibson and Gibbs 2006; Maznevski and Chudoba 2000; Stahl et al. 2010). Shrinking face-to-face interactions may as well reduce the hints on which individuals capitalize on for the social categorization process (Carte and Chidambaram 2004; Mortensen and Hinds 2001): as such, dispersion can hamper the resolution of team conflicts (Hinds and Mortensen 2005).

Beside acknowledging the relevance of authors' physical displacement in terms of the diversity-performance relation, there are no further theoretical arguments for arguing a specific effect of geographic dispersion. Hence we will limit to consider it a moderating factor of the otherwise positive relationship :

H2.a The relationship between cultural diversity of business research teams and research quality is moderated by the geographical dispersion of team members.

H2.b The relationship between cultural diversity of business research teams and research impact is moderated by the geographical dispersion of team members.

3.4 Research Design

3.4.1 Sample

We collected data on business journals from the Clarivate (former ISI) Web of Science. The data contain detailed information on articles published in peer-review academic journals such as authors' surnames, article titles, year of publication, name of journal outlets, authors' affiliations, annual number of citations, etc. We restricted our sample to publications whose authors started publishing from 1997 - to properly account for historical outputs - and before 2010 to ensure enough time for papers to be cited. The final core sample used for our analyses comprises 43,379 articles appeared in 319 journals, co-authored by 133,072 different authors who are affiliated to 13,460 different institutions worldwide. The Clarivate Web of Science database does not contain any information on the authors' cultural background, country of origin or nationality. To fill this gap we rely upon the IBM-GNR database, which is part of a service package developed by IBM (the IBM InfoSphere Global Name Recognition) to manage, search, analyze, and compare multicultural name data sets. The IBM-GNR databases associate a large number of names and surnames to one or (more often) several countries of likely origin, based on information produced by the United States' immigration authorities of the United States of America in the first half of the 1990s, which registered all names and surnames of all foreign citizens entering the US, along with their nationality. When fed with either a name or a surname or both, IBM-GNR returns a list of "countries of association" and two main scores:

- *frequency*, which indicates to which percentile of the frequency distribution of names or surnames the name or surname belongs to, for each country of association[2];
- *significance*, which approximates the frequency distribution of the name or surname across all countries of association[3].

The main limitation of IBM-GNR is the absence of the United States among the likely countries of origin of the recorded names and surnames (information on US citizens was not collected). This drawback makes it impossible to associate a country of origin to authors with an Anglo-Saxon surname. Similar issues arise whenever we try to associate

[2]For example an extremely common Vietnamese surname such as Nguyen will be associated both to Vietnam and to France, which hosts a significant Vietnamese minority; but in Vietnam it will get a frequency value of 90, while it France it will get only, say, 50, the Vietnamese being just a small percentage of the population

[3]continuing with the previous example, the highest percentage of inventors names Nguyen will be found in Vietnam, followed by the US and France

authors to countries of origin whose official languages are not exclusive to those countries only (e.g. Germany, Austria and Switzerland, which share German language; or several Arab-speaking countries) or host large linguistic minorities or immigrant groups. This forces us to abandon the country as unit of reference and opt for larger "clusters" of countries with similar linguistic and cultural traits. The management literature we discussed in section 2 offers several examples of such clusterization. We adopt the one proposed by Ronen and Shenkar 2013. Building on an extensive set of previous studies, they group 62 countries in 11 cultural clusters (see Tab.3.1). We adapted this taxonomy to our data by proceeding in two steps. In the first step we assigned each author in our dataset to one and only one "country of origin", based on the information provided by IBM-GNR. In the second step, we assign such country of origin to one and only one cluster.

As for the first step, we performed it by multiplying the significance and frequency associated with authors' surnames. We then selected the country of origin with the top score. In order to get an idea of the quality of our results, we tested them against information on names, surnames and nationality of around 160000 inventors of USPTO patents from the WIPO-PCT database (Miguelez et al. 2013). Our assignment procedure of country of origin produces a matching rate of 64%. However, when we converted nationalities in cultural groups, following Ronen and Shenkar 2013, the matching rate rise to almost 78%[4]. Finally, we complemented our dataset by referring to specific categorization and ranking lists provided by ABS[5] and ABDC and UT Dallas Business School in order to construct certain variables.

3.4.2 Econometric framework and variables

The units of analysis in our study are research teams, which are composed of academic business researchers that have co-authored a paper published in peer-reviewed journal. We measure the research team performance with two indicators referring to their collaboration: research quality and research impact.

- Research quality: Indicator variable that takes on a series of ordinal values. Higher value indicates higher journal ranking based on ABDC (Australian Business Deans Council) classification of 2013 (Chan et al. 2012).
- Research impact: Total number of citations a paper received until 2013 from all the journals in the ISI Web of Science database. Self-citations are excluded.

[4]Nationality is not an ideal test for the assignment of a country of origin to an individual. Migrants from certain country of origin may acquire the nationality of the country of destination over the course of their life. When considering highly skilled individuals such as business scholars (from our database) or inventors (from the WIPO-PCT database) the incidence of these cases may not be trivial. Hence, a 78% matching is not unsatisfactory.

[5]http://www.arc.gov.au/australian-and-new-zealand-standard-research-classification-anzsrc

Table 3.1: Cultural groups.

United Arab Emirates	Netherlands	Germany	USA	Guatemala	Turkey	Italy	Belarus	South Africa	Philippines	Taiwan
Morocco	Norway	Switzerland	Canada	Brazil	Greece	Portugal	Slovakia	Nigeria	Malaysia	China
Kuwait	Denmark		Australia	Argentina		Spain	Estonia		Thailand	Singapore
	Sweden		Ireland	Chile		France	Czech Republic		Iran	Hong Kong
	Iceland		UK	Uruguay		Belgium	Poland		Indonesia	
	Finland			Mexico			Ukraine		Jamaica	
				Ecuador			Romania		Pakistan	
				Peru			Russia		Zimbabwe	
				Costa Rica			Slovenia			
				Bolivia			Hungary			
				El Salvador			Bulgaria			
				Venezuela						
				Colombia						

Such measures will be regressed on two explanatory variables, namely Team cultural Diversity and Geographic dispersion, plus a set of controls:

- Team cultural diversity: Once the surnames of all the authors in our dataset are assigned to one of the eleven cultural groups, we computed a Jaccard coefficient of similarity expressed as the number of overlapping cultural groups of the team members divided by the number of all the cultural groups at the team-level. Our measure of team cultural diversity is simply equal to 1- Jaccard coefficient of similarity (Jaccard distance).
- Team geographic dispersion: We computed the Jaccard distance at team-level of the differences between authors regarding their countries of affiliations, that is, we subtract from one the ratio between the number of overlapping countries of affiliations of the co-authors and the number of all the different countries of affiliations of these co-authors.

We include three levels of control variables, that may influence research quality and impact: paper's characteristics, team's characteristics and journal's characteristics:

- Paper age: The number of years between the publication date and 2013. The paper age squared is also included in our analyses since older articles have already accumulated a large number of citations and/or have a greater opportunity to attract citations than the recent ones, to a certain extent, then this attractiveness may decrease as the paper becomes obsolete.
- Team research impact: Cumulative number of citations received by other papers since 1997 by the focal paper's co-authors.
- Team knowledge depth: Yearly lagged cumulative number of publications of all the authors composing the team since 1997.
- Team size: Number of coauthors.
- Team tenure: Average number of years since the first publication by each co-author.
- Top institution: Dummy variable indicating whether the focal paper includes at least one author affiliated to an elite institution based on the Top 100 UT Dallas Business School Research Ranking of institutions worldwide. In fact, the University of Texas at Dallas has created a database to track publications in 24 leading business journals in order to provide a top 100 business school rankings since 1990 based on the total contributions of faculty in research which corresponds the context of our study.
- Team collaborations: Total number of previous collaborations among the coauthors of the focal paper up to the publication year.

- Team FT45 stock: Total number of papers previously published on FT45 journals by the coauthors.

- FT45: Dummy variable indicating whether the journal that published the focal paper has ever been listed on the FT45 ranking.

Besides these covariates, we also controlled for some additional variables that could influence both research quality and impact. Since the number of papers increases across years as well as the number of journals within the Web of Science database, we take into account the year fixed effects which capture the changes of practices and behaviors over time. Moreover, as some research domains garner more attention and research activity than others, we used the discipline fixed effects based on the ABS classification.

To address the relationships between the research performance (quality and impact) and the aforementioned explanatory and control variables, we use two different econometric models. We estimate the research quality equation with ordered logistic, due to the ordinal nature of the dependent variable. Research quality is in fact made of four categories, with higher values indicating higher journal quality based on the ABDC ranking. The general equation of the model is the following:

$$Res\_Quality_i = \beta_{0,i} + \beta_1 Cult\_Div_i + \beta_2 Geo\_Disp_i + \beta_c Controls_i + Pub\_Year_i + Discipline_i + \epsilon_i \tag{3.1}$$

Where $\beta_{0,i}$ are the cut points.

Turning our attention to research impact, we adopted a negative binomial regression to identify the factors that facilitate the knowledge diffusion, since the number of citations received by a paper is a count variable. A negative binomial regression is more appropriate than a Poisson model because the former can better deal with over-dispersion issues commonly found in the latter (Wooldridge 2010). Following, is the general equation of the model:

$$Res\_impact_i = \beta_0 + \beta_1 Cult\_Div_i + \beta_2 Geo\_Disp_i + \beta_c Controls_i + Pub\_Year_i + Discipline_i + \epsilon_i \tag{3.2}$$

Tables 3.3 and 3.2 presents descriptive statistics and correlations of the variables.

Table 3.2: Summary statistics.

Variable	Mean	Std. Dev.	Min.	Max.
Research impact	25.6	27.898	0	145
Team cultural diversity	0.669	0.42	0	1
Geographic dispersion	0.821	0.33	0	1
Paper Age	8.492	3.702	4	16
Paper age squared	85.813	70.954	16	256
Team knowledge depth	2.996	2.24	1	17
Team research impact	278.971	395.319	0	6919
Tem collaborations	21.938	21.63	2	273
FT45	0.265	0.441	0	1
Team FT45 stock	3.682	5.472	0	51
Team tenure	11.763	3.442	5	17
Team size	2.558	0.8	2	10
Top institution	0.439	0.496	0	1
Research Quality	**Observations**		**Perc.**	
A*	16804		38.74%	
A	19594		45.17%	
B	3817		8.8%	
C	3164		7.29%	
Observations	43379			

Table 3.3: Correlation matrix

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
[1] Research impact	1.00												
[2] Research quality	0.32***	1.00											
[3] Team cultural diversity	0.04***	0.07***	1.00										
[4] Geographic dispersion	-0.00	0.02***	0.03***	1.00									
[5] Team knowledge depth	0.04***	0.06***	0.01	0.04***	1.00								
[6] Team research impact	0.42***	0.23***	0.07***	0.01***	0.37***	1.00							
[7] Team collaborations	0.12***	0.11***	0.04***	0.03***	0.62***	0.55***	1.00						
[8] FT45	0.29***	0.81***	0.10***	0.01	0.06***	0.26***	0.11***	1.00					
[9] Team FT45 stock	0.24***	0.33***	0.11***	0.03***	0.39***	0.65***	0.59***	0.57***	1.00				
[10] Team tenure	0.25***	0.17***	0.04***	-0.00	0.09***	0.28***	0.14***	0.12***	0.20***	1.00			
[11] Team size	0.01*	-0.01***	0.00	0.05***	0.15***	0.16***	0.47***	-0.01	0.14***	-0.09***	1.00		
[12] Top institution	0.14***	0.18***	0.09***	0.09***	0.11***	0.23***	0.17***	0.22***	0.30***	0.15***	0.08***	1.00	
[13] Paper Age	0.24***	0.08***	-0.01	-0.02***	-0.25***	0.01**	-0.14***	0.02***	-0.08***	0.76***	-0.09***	0.03***	1.00
[14] Paper age squared	0.21***	0.07***	-0.01	-0.02***	-0.25***	-0.00	-0.14***	0.02***	-0.08***	0.74***	-0.08***	0.03***	0.99***

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

For the purpose of computing correlations, the *Research Quality* variable is coded on a 1 to 4 scale.

3.5 Empirical Results

We first present the results relative to Research quality. Tab. 3.4 contains the estimates of Eq.3.1. Column (3) includes the controls discussed in the previous section (with the exclusion of FT45, as such variable is strongly correlated with Research quality and would probably result in biased estimates): the effect of Team cultural diversity is positive and remains significant, meaning that more diverse team on average produce higher quality papers. The effect of geographic dispersion becomes not significant when we include fixed effects for disciplines.

The remaining control variables are all positively associated with Research quality (as expected), with the exclusion of Team research impact: the total number of citations received by the coauthors is a negative predictor of paper quality. Given that self-citations were not excluded, this may be due to the fact that scholars who publish more on lower quality journals, self-cite to grant visibility and promote their own work. The marginal effect of the variable is in fact slightly positive at all quality categories other than A^*.

Of course in case of non-standard models such as the ordinal logistic one, where the dependent variable is expressed on an ordinal scale, the interpretation of stand-alone coefficients is far from being straightforward. Tab.3.5 reports the marginal effects of all the covariates for each outcome of Research quality. An increase in cultural diversity increase the probability of being publish in a top journal, where A^* stands for the highest rank provided by the UTDallas classification.

While HP1a seems to be confirmed, HP2a predicted a moderating effect of team geographic dispersion on the relationship between team cultural diversity and research performance. In Tab. 3.6 we replicate the previous set of estimates with an interaction term between Cultural diversity and Geographic dispersion. While the sign and significance of variables do not vary, the interaction term is positive (Columns (1) to (3)), indicating a moderating effect of geographical dispersion: the positive link between cultural diversity and the probability of being published on a top journal is strengthen with more geographically dispersed teams.

Looking at the moderating effect across different outcomes of Research quality (see Fig.3.1), we observe that the moderating behavior of Geographic dispersion only operates for papers published on A^* journals.

These results combined (that is the ones in Tab.3.5 and Fig.3.1) offer a clear picture: cultural diversity seems to operate only at the top level of journal quality. It does not seem to be a crucial factor when publishing in average and below-average quality journals. This could signal the presence of some unobserved set of skills of coauthors: it could be that top researchers possess an intrinsic predisposition to multicultural environments, or that they are willing to cross cultural-borders in order to collaborate with scholars who

Table 3.4: Research quality, ordinal logit model.

	(1)	(2)	(3)
Team cultural diversity	0.367***	0.142***	0.122***
	(0.0220)	(0.0231)	(0.0240)
Geographic dispersion	0.133***	0.0748*	0.0556
	(0.0280)	(0.0295)	(0.0305)
Team knowledge depth		-0.0338***	-0.0165**
		(0.00573)	(0.00608)
Team research impact		-0.00302***	-0.00369***
		(0.000199)	(0.000205)
Team collaborations		0.0238***	0.0293***
		(0.00173)	(0.00184)
Team FT45 stock		0.385***	0.458***
		(0.00565)	(0.00642)
Team tenure		0.0761***	0.0679***
		(0.00577)	(0.00602)
Team size		0.0216	0.00274
		(0.0179)	(0.0187)
Top institution		0.323***	0.320***
		(0.0219)	(0.0228)
Cutpoint 1	-2.885***	-3.020***	-4.386***
	(0.0544)	(0.0682)	(0.0917)
Cutpoint 2	-1.718***	-1.810***	-3.123***
	(0.0513)	(0.0658)	(0.0896)
Cutpoint 3	0.553***	0.772***	-0.260**
	(0.0505)	(0.0650)	(0.0880)
Discipline dummies	No	No	Yes
Publication year dummies	Yes	Yes	Yes
Observations	43379	43379	43379
Log-likelihood	-45523.7	-41164.8	-38129.7
χ^2	1656.9	9374.7	15444.9
Pseudo R^2	0.088	0.102	0.168

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Estimated parameters. Odds ratios correspond to exponentiated coefficients, i.e. e^β

Table 3.5: Research quality: marginal effects of Model (3).

	C	B	A	A*
Team cultural diversity	-0.006	-0.005	-0.005	0.016
Geographic dispersion	-0.006	-0.005	-0.006	0.018
Team knowledge depth	0.001	0.001	0.001	-0.002
Team research impact	0.000^+	0.000^+	0.000^+	-0.001
Team collaborations	-0.002	0.002	0.002	0.005
Team FT45 stock	0.027	0.024	0.025	0.076
Team tenure	-0.005	-0.004	-0.004	0.013
Team size	-0.001	-0.001	-0.001	0.003
Top institution	-0.020	-0.017	-0.018	0.055

Marginal effects are computed keeping the other covariates fixed at their mean value. They can be interpreted as the marginal effect on the absolute probability for a paper to fall in each category.

Table 3.6: **Research quality, testing for moderating effect. Ordinal logit.**

	(1)	(2)	(3)
[1] Team cultural diversity	0.243***	0.1407***	0.217***
	(0.0542)	(0.0366)	(0.0192)
[2] Geographic dispersion	0.0345	-0.00506	-0.0266
	(0.0483)	(0.0504)	(0.0519)
[1] * [2]	0.153***	0.125**	0.134***
	(0.0412)	(0.0439)	(0.0260)
Team knowledge depth		-0.0339***	-0.0166**
		(0.00573)	(0.00608)
Team research impact		-0.00302***	-0.00369***
		(0.000199)	(0.000205)
Team collaborations		0.0238***	0.0293***
		(0.00173)	(0.00184)
Team FT45 stock		0.385***	0.458***
		(0.00565)	(0.00642)
Team tenure		0.0762***	0.0680***
		(0.00577)	(0.00602)
Team size		0.0219	0.00296
		(0.0179)	(0.0187)
Top institution		0.322***	0.319***
		(0.0219)	(0.0228)
Cutpoint 1	-2.966***	-3.085***	-4.453***
	(0.0632)	(0.0759)	(0.0980)
Cutpoint 2	-1.798***	-1.875***	-3.190***
	(0.0606)	(0.0737)	(0.0960)
Cutpoint 3	0.473***	0.708***	-0.327***
	(0.0598)	(0.0729)	(0.0945)
Discipline dummies	No	No	Yes
Publication year dummies	Yes	Yes	Yes
Observations	43379	43379	43379
log-likelihood	-45520.6	-41162.9	-38127.8
χ^2	1663.1	9378.5	15448.7
Pseudo R^2	0.097	0.108	0.16

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

posses complementary or field-specific skills.

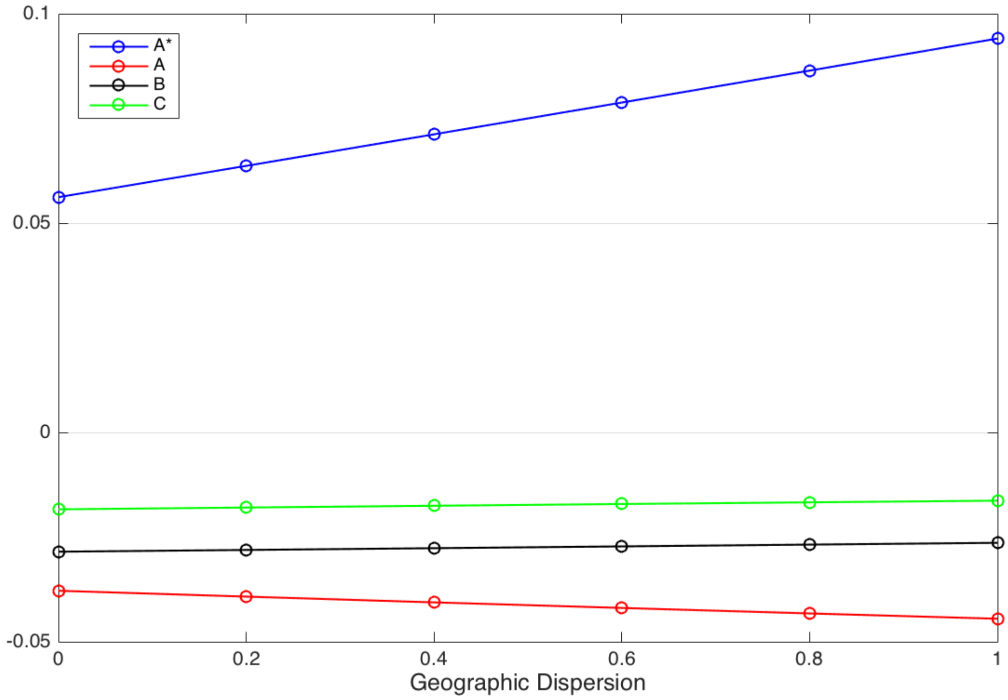
Turning to HP1b, results in Tab.3.7 examine the main effects of team cultural diversity on research impact . Model 3 includes the full set of control variables and fixed effects for field and publication year.

Results reveal that teams exhibiting higher cultural diversity are more likely to be cited, consistent with the hypothesis HP1b. Specifically, a one percentage positive shift in *Team cultural diversity* results in about 4% more forward citations[6]. On average, considering the mean value of the dependent variable, the absolute increment correspond to an additional forward citation.

The variable *FT45* should account for the overall quality of research expressed by the

[6]Within a negative binomial framework in fact, a coefficient of 0.0386 correspond to an incidence rate ratio of 1.039.

Figure 3.1: Research quality, interaction. Marginal effects.



team, through the visibility of the journal: the inclusion of such control should in principle make the coefficient of *Team cultural diversity* robust to heterogeneity in terms of quality of research. The estimate indicate that papers published on more "in sight" journals are on average more cited, while geographically dispersed teams receive fewer citations.

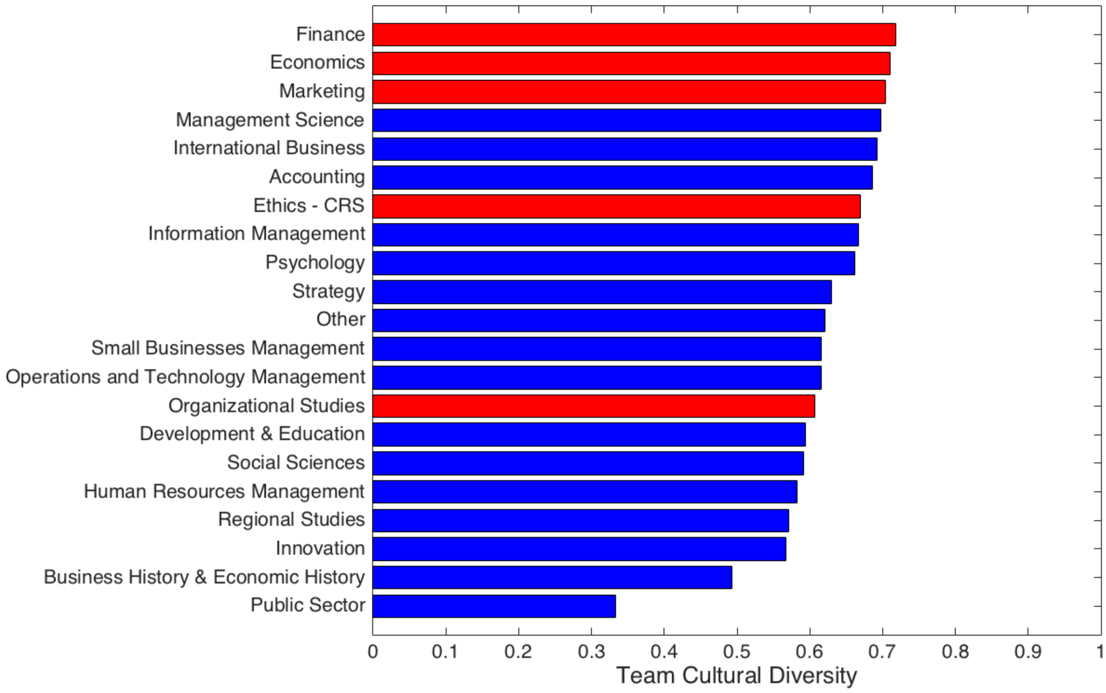
Also the relation between the number of forward citations and paper age is non-linear and concave.

No major effect is found for *Team knowledge depth*, while the fact that one or more authors belong to a top institution increases the number of citations by almost the 10%. Tab.3.8 present the estimates of the full model for the most populated fields of our database, that is Finance, Ethics - Corporate Social Responsibility, Marketing, Organization Studies and Economics[7]. Just to have an idea of the behavior of *Team cultural diversity* across Management disciplines, Fig.3.2 shows the mean values of such variable across fields. We can see that there is little variation across the 5 fields under analysis, and all of them show high levels of cultural diversity (between 0.66 and 0.71).

The effect of *Team cultural diversity* varies across disciplines: the highest are found for Organization and Economics (respectively a 13% and 12% increase in the number of forward citations), while lower values are found for Finance and Marketing, while for the Ethics-CRS field there is no appreciable effect. Field-specific unobservable dynamics are likely to be in place: however we limit ourselves to remark that the effects are in general

[7]Each paper is assigned to la single field.

Figure 3.2: Average Team Cultural diversity.



positive or at most null, and we leave a deeper investigation of such dynamics for future research.

Last, as for the first set of hypothesis, we include an interaction term between Cultural diversity and Geographic dispersion to test HP2b (Tab.3.9). Column (3) shows a positive coefficient for the interaction term (with no effect for the standalone variable): indicating that cultural diversity enables research impact (as measured by the number of citations a paper received) when teams members are highly dispersed in space. This may be due to the fact that when cultural diversity is higher coauthors have access to different networks of academics, which makes their work more visible and increases its likelihood to be cited.

Table 3.7: Research impact.

	(1)	(2)	(3)
Team cultural diversity	0.140***	0.0218*	0.0386***
	(0.0114)	(0.0104)	(0.0103)
Geographic dispersion	-0.0238*	-0.0135	-0.0470**
	(0.0115)	(0.0101)	(0.0143)
Paper Age	0.406***	0.329***	0.323***
	(0.0321)	(0.0295)	(0.0291)
Paper age squared	-0.0163***	-0.0124***	-0.0122***
	(0.00158)	(0.00145)	(0.00143)
Team knowledge depth		0.00355	-0.00212
		(0.00279)	(0.00281)
Team research impact		0.00154***	0.00144***
		(0.0000219)	(0.0000215)
Team collaborations		-0.00143***	-0.000956**
		(0.000365)	(0.000363)
FT45		0.583***	0.633***
		(0.0126)	(0.0129)
Team FT45 stock		-0.0380***	-0.0360***
		(0.00141)	(0.00142)
Team tenure		-0.0275***	-0.0213***
		(0.00235)	(0.00233)
Team size		-0.0332***	-0.0357***
		(0.00653)	(0.00647)
Top institution		0.0779***	0.0938***
		(0.00924)	(0.00915)
Constant	-0.0837***	-0.285***	-0.318***
	(0.00671)	(0.00699)	(0.00704)
Discipline dummies	No	No	Yes
Publication year dummies	Yes	Yes	Yes
Observations	43379	43379	43379
Log-likelihood	-181508.2	-176948.1	-176228.0
χ^2	6661.1	15781.3	17221.6
Pseudo R^2	0.1180	0.1427	0.1466

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.8: Research impact.

	ECONOMICS	ETHICS-CRS	FINANCE	MARKETING	ORGANIZATION
Team cultural diversity	0.118**	-0.0227	0.0891**	0.0669*	0.123***
	(0.0392)	(0.0297)	(0.0281)	(0.0279)	(0.0315)
Geographic dispersion	-0.0562	-0.178***	-0.0689*	-0.0853*	-0.0800*
	(0.0476)	(0.0348)	(0.0345)	(0.0334)	(0.0393)
Paper Age	0.424***	0.425***	0.447***	0.490***	0.442***
	(0.0205)	(0.0157)	(0.0139)	(0.0143)	(0.0163)
Paper age squared	-0.0174***	-0.0184***	-0.0187***	-0.0208***	-0.0190***
	(0.000945)	(0.000764)	(0.000674)	(0.000685)	(0.000723)
Team knowledge depth	0.0155	0.0142	0.0140	0.0118	-0.0188*
	(0.0134)	(0.00905)	(0.00808)	(0.00733)	(0.00900)
Team research impact	0.00350***	0.00118***	0.00140***	0.00108***	0.00158***
	(0.000156)	(0.0000569)	(0.0000631)	(0.0000473)	(0.0000722)
Tem collaborations	-0.0149***	-0.00199	-0.00138	-0.00439***	0.000809
	(0.00203)	(0.00105)	(0.00108)	(0.000950)	(0.00127)
FT45	0.576***	0.451***	0.885***	0.548***	0.786***
	(0.0289)	(0.0305)	(0.0374)	(0.0344)	(0.0382)
Team FT45 stock	-0.0800***	-0.0284***	-0.0365***	-0.0237***	-0.0423***
	(0.00998)	(0.00398)	(0.00453)	(0.00295)	(0.00447)
Team tenure	-0.00853	0.0149*	-0.00788	0.00385	0.00307
	(0.00884)	(0.00647)	(0.00551)	(0.00566)	(0.00698)
Team size	0.104***	0.111***	0.0811***	0.111***	0.00842
	(0.0267)	(0.0177)	(0.0186)	(0.0168)	(0.0170)
Top institution	0.121***	0.114***	0.0914***	0.0668**	0.0836**
	(0.0351)	(0.0272)	(0.0248)	(0.0250)	(0.0279)
Constant	-0.242***	-0.105***	-0.218***	-0.464***	-0.248***
	(0.0269)	(0.0190)	(0.0181)	(0.0203)	(0.0208)
Observations	3074	5830	6656	5098	4929
Log-likelihood	-11557.0	-23894.2	-25872.8	-21483.5	-19451.5
χ^2	28120.7	59461.5	65848.7	80694.6	52150.3
Pseudo R^2	0.1201	0.1312	0.1136	0.0987	0.1242

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.9: Research impact, testing for moderating effect.

	(1)	(2)	(3)
[1] Team cultural diversity	0.0145	0.0716**	0.0483
	(0.0286)	(0.0255)	(0.0251)
[2] Geographic dispersion	-0.0953***	-0.0885***	-0.0744***
	(0.0257)	(0.0229)	(0.0226)
[1] ∗ [2]	0.169***	0.116***	0.108***
	(0.0324)	(0.0288)	(0.0283)
Paper Age	0.402***	0.329***	0.323***
	(0.0326)	(0.0295)	(0.0291)
Paper age squared	-0.0163***	-0.0124***	-0.0122***
	(0.00160)	(0.00145)	(0.00143)
Team knowledge depth		0.00349	-0.00215
		(0.00279)	(0.00280)
Team research impact		0.00154***	0.00144***
		(0.0000219)	(0.0000215)
Team collaborations		-0.00144***	-0.000970**
		(0.000365)	(0.000363)
FT45		0.583***	0.633***
		(0.0126)	(0.0129)
Team FT45 stock		-0.0380***	-0.0360***
		(0.00141)	(0.00142)
Team tenure		-0.0274***	-0.0212***
		(0.00235)	(0.00233)
Team size		-0.0329***	-0.0354***
		(0.00653)	(0.00647)
Top institution		0.0775***	0.0933***
		(0.00924)	(0.00915)
lnalpha			
Constant	-0.0466***	-0.285***	-0.318***
	(0.00666)	(0.00699)	(0.00704)
Discipline dummies	No	No	Yes
Publication year dummies	Yes	Yes	Yes
Observations	43379	43379	43379
Log-likelihood	-182384.1	-176940.0	-176220.7
χ^2	4909.4	15797.6	17236.1
Pseudo R^2	0.1133	0.1427	0.1469

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.6 Discussion and conclusions

This chapter mostly builds the on management literature, namely the one on team functioning, attributes and performance. Diversity in teams has been extensively investigated in recent years, mainly due to the increasing heterogeneity observed in the labour force. Previous research addressing the role of ethnic/cultural diversity in teams has often shown contradictory results (Watkins and Ferrara 2005, Ottaviano and Peri 2006, Trax et al. 2015, Parrotta et al. 2012 Lee and Nathan 2010).

In order to advance our understanding on such relationship, we exploit a database on scientific publications. Specifically we focus on academics who published in business-related fields between 1997 and 2009. Through a surname analysis we assign a cultural/ethnic group to each scholar, and study the effect of diversity, that stems from collaborations between individuals belonging of different cultural groups, on two measures of team performance.

We test four hypotheses that rest on Social Identity and Information theory. Following recent contributions, we check the presence of confounding factors by looking at the effect of geographical dispersion of team members on the aforementioned diversity-performance connection. Our result supports the view of information-processing theory, which argues that diversity brings different perspectives and a wide range of network-mediated resources to enhance problem-solving, creativity, and innovation in team.

In terms of research quality, as measured with the UT Dallas ranking of the journal in which a team manage to publish its paper, we find a positive association with team cultural diversity, although high levels of heterogeneity seem to be beneficial only in order to get published on journals belonging to the top UT Dallas category. The effect strengthened in teams that are more dispersed in space: common values and a shared cultural background may overcome the lack of geographical proximity among researchers.

To quantify team research impact we use the number of forward citations received by a paper since the publication date. We find a positive effect of cultural diversity, which however holds only in geographically dispersed teams: when team members are concentrated in space, such effect is null.

Future developments include an enrichment of the dataset, with the inclusion of instituion- and country-level variables, and some robustness checks on the surname analysis procedure.

Bibliography

Acs, Z. J., Anselin, L., and Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research policy*, 31(7):1069–1085.

Adler, N. J. and Gundersen, A. (2007). *International dimensions of organizational behavior*. Cengage Learning.

Agrawal, A., Kapur, D., and McHale, J. (2008). How do spatial and social proximity influence knowledge flows? evidence from patent data. *Journal of urban economics*, 64(2):258–269.

Agrawal, A., Kapur, D., McHale, J., and Oettl, A. (2011). Brain drain or brain bank? the impact of skilled emigration on poor-country innovation. *Journal of Urban Economics*, 69(1):43–55.

Alcacer, J. and Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. *The Review of Economics and Statistics*, 88(4):774–779.

Alderfer, C. P. and Smith, K. K. (1982). Studying intergroup relations embedded in organizations. *Administrative Science Quarterly*, pages 35–65.

Alesina, A., Harnoss, J., and Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2):101–138.

Alesina, A. and La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *Journal of public Economics*, 89(5):897–931.

Almeida, P., Phene, A., and Li, S. (2014). The influence of ethnic community knowledge on indian inventor innovativeness. *Organization Science*, 26(1):198–217.

Altonji, J. G. and Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, pages 201–234. University of Chicago Press.

Arslan, C., Dumont, J.-C., Kone, Z., Moullan, Y., Ozden, C., Parsons, C., and Xenogiani, T. (2015). A new profile of migrants in the aftermath of the recent economic crisis.

Arthur, W. B. (2007). The structure of invention. *Research policy*, 36(2):274–287.

Arthur, W. B. (2009). *The nature of technology: What it is and how it evolves*. Simon and Schuster.

Azoulay, P., Zivin, J. S. G., and Sampat, B. N. (2011). The diffusion of scientific knowledge across time and space: Evidence from professional transitions for the superstars of medicine. Technical report, National Bureau of Economic Research.

Bacchiocchi, E. and Montobbio, F. (2009). Knowledge diffusion from university and public research. a comparison between us, japan and europe using patent citations. *The Journal of Technology Transfer*, 34(2):169–181.

Baixauli-Soler, J. S., Belda-Ruiz, M., and Sanchez-Marin, G. (2015). Executive stock options, gender diversity in the top management team, and firm risk taking. *Journal of Business Research*, 68(2):451–463.

Bakker, R. M., Boroş, S., Kenis, P., and Oerlemans, L. A. (2013). It's only temporary: time frame and the dynamics of creative project teams. *British Journal of Management*, 24(3):383–397.

Bański, J. (2013). Modern-day demographic processes in central europe and their potential interactions with climate change. *Papers on Global Change IGBP*, 20(1):85–99.

Bantel, K. A. and Jackson, S. E. (1989). Top management and innovations in banking: Does the composition of the top team make a difference? *Strategic management journal*, 10(S1):107–124.

Bassett-Jones, N. (2005). The paradox of diversity management, creativity and innovation. *Creativity and innovation management*, 14(2):169–175.

Bechky, B. A. (2006). Gaffers, gofers, and grips: Role-based coordination in temporary organizations. *Organization Science*, 17(1):3–21.

Beine, M., Docquier, F., and Özden, Ç. (2011). Diasporas. *Journal of Development Economics*, 95(1):30–41.

Beine, M., Docquier, F., and Rapoport, H. (2008). Brain drain and human capital formation in developing countries: winners and losers. *The Economic Journal*, 118(528):631–652.

Bellini, E., Ottaviano, G. I., Pinelli, D., and Prarolo, G. (2013). Cultural diversity and economic performance: evidence from european regions. In *Geography, institutions and regional economic performance*, pages 121–141. Springer.

Belot, M. V. and Hatton, T. J. (2012). Immigrant selection in the oecd. *The Scandinavian Journal of Economics*, 114(4):1105–1128.

Bercovitz, J. and Feldman, M. (2011). The mechanisms of collaboration in inventive teams: Composition, social networks, and geography. *Research Policy*, 40(1):81–93.

Beyer, R. (2016). The labor market performance of immigrants in germany. *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Immigration and Labor Markets, No. D22-V3.*

Bhagwati, J. and Hamada, K. (1974). The brain drain, international integration of markets for professionals and unemployment: a theoretical analysis. *Journal of Development Economics*, 1(1):19–42.

Blackwell, M., Iacus, S. M., King, G., and Porro, G. (2009). cem: Coarsened exact matching in stata. *The Stata Journal*, 9(4):524–546.

Boeri, T. (2012). *Brain drain and brain gain: The global competition to attract high-skilled migrants*. Oxford University Press.

Borjas, G. J. (1987). Immigrants, minorities, and labor market competition. *ILR Review*, 40(3):382–392.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, 118(4):1335–1374.

Borjas, G. J. and Doran, K. B. (2012a). Cognitive mobility: Labor market responses to supply shocks in the space of ideas.

Borjas, G. J. and Doran, K. B. (2012b). The collapse of the soviet union and the productivity of american mathematicians. *HKS Faculty Research Working Paper Series.*

Borjas, G. J., Doran, K. B., and Shen, Y. (2017). Ethnic complementarities after the opening of china: How chinese graduate students affected the productivity of their advisors. *Journal of Human Resources*, pages 0516–7949R.

Boschma, R. (2005). Role of proximity in interaction and performance: conceptual and empirical challenges.

Bosetti, V., Cattaneo, C., and Verdolini, E. (2012). Migration, cultural diversity and innovation: A european perspective.

Breschi, S., Lissoni, F., Miguelez, E., et al. (2015). Foreign inventors in the us: Testing for diaspora and brain gain effects.

Breschi, S., Lissoni, F., and Tarasconi, G. (2017). Inventor data for research on migration & innovation: The ethnic-inv pilot database. *The International Mobility of Talent and Innovation: New Evidence and Policy Implications.*

Breschi, S., Lissoni, F., and Temgoua, C. N. (2016). 23. migration and innovation: a survey of recent studies. *Handbook on the Geographies of Innovation*, page 382.

Brown, R. and Carasso, H. (2013). *Everything for sale?: the marketisation of UK higher education*. Routledge.

Brubaker, R. (1998). Migrations of ethnic unmixing in the" new europe". *The International migration review*, 32(4):1047–1065.

Brusoni, S., Criscuolo, P., and Geuna, A. (2005). The knowledge bases of the worlds largest pharmaceutical groups: what do patent citations to non-patent literature reveal? *Economics of Innovation and New Technology*, 14(5):395–415.

Buckley, C. (1995). The myth of managed migration: Migration control and market in the soviet period. *Slavic Review*, 54(4):896–916.

Carte, T. and Chidambaram, L. (2004). A capabilities-based theory of technology deployment in diverse teams: Leapfrogging the pitfalls of diversity and leveraging its potential with collaborative technology. *Journal of the Association for Information Systems*, 5(11):4.

Carter, D. A., D'Souza, F., Simkins, B. J., and Simpson, W. G. (2010). The gender and ethnic diversity of us boards and board committees and firm financial performance. *Corporate Governance: An International Review*, 18(5):396–414.

Cerna, L. and Chou, M.-H. (2014). The regional dimension in the global competition for talent: Lessons from framing the european scientific visa and blue card. *Journal of European Public Policy*, 21(1):76–95.

Cerna, L. and Czaika, M. (2016). European policies to attract talent: The crisis and highly skilled migration policy changes. In *High-Skill Migration and Recession*, pages 22–43. Springer.

Chaloff, J. and Lemaitre, G. (2009). Managing highly-skilled labour migration.

Chan, K. C., Tong, J. Y., and Zhang, F. F. (2012). Accounting journal rankings, authorship patterns and the author affiliation index. *Australian Accounting Review*, 22(4):407–417.

Chellaraj, G., Maskus, K. E., and Mattoo, A. (2006). Skilled immigrants, higher education, and us innovation. *International migration, remittances and the brain drain*, pages 245–59.

Chiquiar, D. and Hanson, G. H. (2005). International migration, self-selection, and the distribution of wages: Evidence from mexico and the united states. *Journal of political Economy*, 113(2):239–281.

Chiswick, B. R. (1999). Are immigrants favorably self-selected? *The American economic review*, 89(2):181–185.

Cipolla, C. M. (1972). The diffusion of innovations in early modern europe. *Comparative Studies in Society and History*, 14(1):46–52.

Coffano, M. and Tarasconi, G. (2014). Crios-patstat database: sources, contents and access rules. *Center for Research on Innovation, Organization and Strategy, CRIOS Working Paper*, (1).

Connaughton, S. L. and Shuffler, M. (2007). Multinational and multicultural distributed teams: A review and future agenda. *Small group research*, 38(3):387–412.

Cooper, J. (2013). The russian economy twenty years after the end of the socialist economic system. *Journal of Eurasian Studies*, 4(1):55–64.

Cramton, C. D. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. *Organization science*, 12(3):346–371.

Dabrowski, M. (2016). Currency crises in post-soviet economiesa never ending story? *Russian Journal of Economics*, 2(3):302–326.

Dahlin, K. B., Weingart, L. R., and Hinds, P. J. (2005). Team diversity and information use. *Academy of Management Journal*, 48(6):1107–1123.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application*, volume 1. Cambridge university press.

De Rassenfosse, G., Dernis, H., Guellec, D., Picci, L., and de la Potterie, B. v. P. (2013). The worldwide count of priority patents: A new indicator of inventive activity. *Research Policy*, 42(3):720–737.

Desmet, K., Weber, S., and Ortuño-Ortín, I. (2009). Linguistic diversity and redistribution. *Journal of the European Economic Association*, 7(6):1291–1318.

Dietz, B. (2000). German and jewish migration from the former soviet union to germany: background, trends and implications. *Journal of Ethnic and Migration Studies*, 26(4):635–652.

Docquier, F., Marfouk, A., et al. (2006). International migration by educational attainment, 1990-2000.

Dos Santos, M. D. and Postel-Vinay, F. (2003). Migration as a source of growth: the perspective of a developing country. *Journal of Population Economics*, 16(1):161–175.

Dumont, J.-C., Spielvogel, G., and Widmaier, S. (2010). International migrants in developed, emerging and developing countries.

Durand, R. and Jacqueminet, A. (2015). Peer conformity, attention, and heterogeneous implementation of practices in mnes. *Journal of International Business Studies*, 46(8):917–937.

Earley, C. P. and Mosakowski, E. (2000). Creating hybrid team cultures: An empirical test of transnational team functioning. *Academy of Management Journal*, 43(1):26–49.

Earley, P. C. (1993). *Culture, self-identity, and work*. Oxford University Press on Demand.

Ely, R. J. and Thomas, D. A. (2001). Cultural diversity at work: The effects of diversity perspectives on work group processes and outcomes. *Administrative science quarterly*, 46(2):229–273.

Erez, M. (2010). Culture and job design. *Journal of Organizational Behavior*, 31(2-3):389–400.

Espinosa, J. A., Cummings, J. N., Wilson, J. M., and Pearce, B. M. (2003). Team boundary issues across multiple global firms. *Journal of Management Information Systems*, 19(4):157–190.

Feldbrugge, F. F. J. M., Van den Berg, G. P., and Simons, W. B. (1985). *Encyclopedia of Soviet law*. Number 28. Brill.

Finkelstein, S. and Hambrick, D. C. (1996). *Strategic leadership: Top executives and their effects on organizations*. West Publishing Company.

Fleming, L., Mingo, S., and Chen, D. (2007). Collaborative brokerage, generative creativity, and creative success. *Administrative science quarterly*, 52(3):443–475.

Fortescue, S. (1986). *The Communist Party and Soviet Science*. Springer.

Franzoni, C., Scellato, G., and Paula, S. (2012). Foreign-born scientists: Mobility patterns for 16 countries. *Nature Biotechnology*, 30(12):1250–53.

Freeman, R. B. (2010). Globalization of scientific and engineering talent: international mobility of students, workers, and ideas and the world economy. *Economics of Innovation and New Technology*, 19(5):393–406.

Freeman, R. B. and Huang, W. (2015). Collaborating with people like me: Ethnic coauthorship within the united states. *Journal of Labor Economics*, 33(3 pt 2).

Ganguli, I. (2015a). Immigration and ideas: What did russian scientists bring to the united states? *Journal of Labor Economics*, 33(S1):S257–S288.

Ganguli, I. (2015b). Who leaves and who stays? Evidence on immigrant selection from the collapse of soviet science. *Global mobility of research scientists: The economics of who goes where and why*, page 133.

Garcia-Peñalosa, C. and Wasmer, É. (2016). Preparing france for the increasing international mobility of talents. *Notes du conseil danalyse économique*, (4):1–12.

Gavigan, J. P., Ottitsch, M., and Greaves, C. (1999). *Demographic and social trends panel report*. Institute for Prospective Technological Studies.

Gibson, C. B. and Gibbs, J. L. (2006). Unpacking the concept of virtuality: The effects of geographic dispersion, electronic dependence, dynamic structure, and national diversity on team innovation. *Administrative Science Quarterly*, 51(3):451–495.

Gokhberg, L. and Nekipelova, E. (2002). International migration of scientists and engineers in russia. *International mobility of the highly skilled*, pages 177–187.

Graham, L. R. (1993). *Science in Russia and the Soviet Union: A short history*. Cambridge University Press.

Graham, L. R. and Dezhina, I. (2008). *Science in the new Russia: Crisis, aid, reform*. Indiana University Press.

Greenwald, A. G. and Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, 102:4.

Griffith, T. L. and Neale, M. A. (2001). 8. information processing in traditional, hybrid, and virtual teams: From nascent knowledge to transactive memory. *Research in organizational behavior*, 23:379–421.

Grimm, C. M. and Smith, K. G. (1991). Research notes and communications management and organizational change: A note on the railroad industry. *Strategic Management Journal*, 12(7):557–562.

Haas, M. R. (2006). Acquiring and applying knowledge in transnational teams: The roles of cosmopolitans and locals. *Organization Science*, 17(3):367–384.

Hambrick, D. C., Cho, T. S., and Chen, M.-J. (1996). The influence of top management team heterogeneity on firms' competitive moves. *Administrative science quarterly*, pages 659–684.

Hambrick, D. C. and D'Aveni, R. A. (1992). Top team deterioration as part of the downward spiral of large corporate bankruptcies. *Management Science*, 38(10):1445–1466.

Hambrick, D. C., Davison, S. C., Snell, S. A., and Snow, C. C. (1998). When groups consist of multiple nationalities: Towards a new understanding of the implications. *Organization studies*, 19(2):181–205.

Hambrick, D. C. and Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of management review*, 9(2):193–206.

Hargadon, A. (2005). Technology brokering and innovation: linking strategy, practice, and people. *Strategy & Leadership*, 33(1):32–36.

Harhoff, D., Narin, F., Scherer, F. M., and Vopel, K. (1999). Citation frequency and the value of patented inventions. *Review of Economics and statistics*, 81(3):511–515.

Harrison, D. A., Price, K. H., Gavin, J. H., and Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface-and deep-level diversity on group functioning. *Academy of management journal*, 45(5):1029–1045.

Hatton, T. J. (2014). The economics of international migration: A short history of the debate. *Labour Economics*, 30:43–50.

Heleniak, T. (2002). Migration dilemmas haunt post-soviet russia. *Migration Information Source*.

Henderson, A. D., Miller, D., and Hambrick, D. C. (2006). How quickly do CEOs become obsolete? industry dynamism, ceo tenure, and company performance. *Strategic Management Journal*, 27(5):447–460.

Hinds, P., Liu, L., and Lyon, J. (2011). Putting the global in global work: An intercultural lens on the practice of cross-national collaboration. *Academy of Management Annals*, 5(1):135–188.

Hinds, P. J. and Mortensen, M. (2005). Understanding conflict in geographically distributed teams: The moderating effects of shared identity, shared context, and spontaneous communication. *Organization science*, 16(3):290–307.

Hirschman, A. O. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*, volume 25. Harvard university press.

Hong, L. and Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46):16385–16389.

Hornung, E. (2014). Immigration and the diffusion of technology: The huguenot diaspora in prussia. *The American Economic Review*, 104(1):84–122.

Hunt, J. (2011). Which immigrants are most innovative and entrepreneurial? distinctions by entry visa. *Journal of Labor Economics*, 29(3):417–457.

Hunt, J. (2015). Are immigrants the most skilled us computer and engineering workers? *Journal of Labor Economics*, 33(S1):S39–S77.

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

Iacus, S. M., King, G., and Porro, G. (2008). Matching for causal inference without balance checking.

Iacus, S. M., King, G., and Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political analysis*, 20(1):1–24.

J Acs, Z. and Audretsch, D. B. (1989). Patents as a measure of innovative activity. *Kyklos*, 42(2):171–180.

Jaffe, A. B. and De Rassenfosse, G. (2016). Patent citation data in social science research: Overview and best practices. Technical report, National Bureau of Economic Research.

Jehn, K. A., Northcraft, G. B., and Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. *Administrative science quarterly*, 44(4):741–763.

Jensen, M. and Wang, P. (2017). Not in the same boat: How status inconsistency affects research performance in business schools. *Academy of Management Journal*, pages amj–2015.

Johansson, M. and Śliwa, M. (2014). Gender, foreignness and academia: An intersectional analysis of the experiences of foreign women academics in uk business schools. *Gender, Work & Organization*, 21(1):18–36.

Jones, B. F. (2009). The burden of knowledge and the death of the renaissance man: Is innovation getting harder? *The Review of Economic Studies*, 76(1):283–317.

Joshi, A. and Roh, H. (2009). The role of context in work team diversity research: A meta-analytic review. *Academy of Management Journal*, 52(3):599–627.

Judge, T. A., Cable, D. M., Colbert, A. E., and Rynes, S. L. (2007). What causes a management article to be citedarticle, author, or journal? *Academy of Management Journal*, 50(3):491–506.

Kale, D., Wield, D., and Chataway, J. (2008). Diffusion of knowledge through migration of scientific labour in india. *Science and Public Policy*, 35(6):417–430.

Keck, S. (1991). Top management team structure: does it matter anyway? In *Academy of management meeting, Miami, FL.*

Kemeny, T. and Cooke, A. (2015). Spillovers from immigrant diversity in cities. *Journal of Economic Geography.*

Kerr, S. P., Kerr, W., Özden, Ç., and Parsons, C. (2016). Global talent flows. *The Journal of Economic Perspectives*, 30(4):83–106.

Kerr, W. R. (2008). The ethnic composition of us inventors. *Harvard Business School Working Paper*.

Kerr, W. R. (2010). Breakthrough inventions and migrating clusters of innovation. *Journal of urban Economics*, 67(1):46–60.

Kerr, W. R. (2013). Us high-skilled immigration, innovation, and entrepreneurship: Empirical approaches and evidence. Technical report, National Bureau of Economic Research.

Kerr, W. R. and Lincoln, W. F. (2010). The supply side of innovation: H-1b visa reforms and us ethnic invention. *Journal of Labor Economics*, 28(3):473–508.

Kerr, W. R. and Mandorff, M. (2015). Social networks, ethnicity, and entrepreneurship. Technical report, National Bureau of Economic Research.

Kim, T. (2009). Transnational academic mobility, internationalization and interculturality in higher education. *Intercultural Education*, 20(5):395–405.

Kirkman, B. L., Lowe, K. B., and Gibson, C. B. (2006). A quarter century of culture's consequences: A review of empirical research incorporating hofstede's cultural values framework. *Journal of international business studies*, 37(3):285–320.

Kopnina, H. (2005). *East to West Migration Russian Migrants in Western Europe*.

Kurthen, H. (1995). Germany at the crossroads: national identity and the challenges of immigration. *International Migration Review*, pages 914–938.

Lakha, F., Gorman, D. R., and Mateos, P. (2011). Name analysis to classify populations by ethnicity in public health: validation of onomap in scotland. *Public health*, 125(10):688–696.

Lambert, J. and Usher, A. (2013). *The pros and cons of internationalization: How domestic students experience the globalizing campus*. Higher Education Strategy Associates.

Lanjouw, J. O. and Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators*. *The Economic Journal*, 114(495):441–465.

Larkey, L. K. (1996). Toward a theory of communicative interactions in culturally diverse workgroups. *Academy of Management Review*, 21(2):463–491.

Latova, N. V. and Savinkov, V. I. (2012). The influence of academic migration on the intellectual potential of russia. *European Journal of Education*, 47(1):64–76.

Lee, N. and Nathan, M. (2010). Knowledge workers, cultural diversity and innovation: evidence from london. *International Journal of Knowledge-Based Development*, 1(1-2):53–78.

Lee, S. and Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social studies of science*, 35(5):673–702.

Lewis, B. W., Walls, J. L., and Dowell, G. W. (2014). Difference in degrees: Ceo characteristics and firm environmental disclosure. *Strategic Management Journal*, 35(5):712–722.

Lissoni, F. (2017). International migration and innovation diffusion: an eclectic survey. *Regional Studies*, 0(0):1–13.

Lissoni, F., Mairesse, J., Montobbio, F., and Pezzoni, M. (2011). Determinants of scientific productivity: A study on italian and french physicists. *Industrial and Corporate Change*, 20(1):253–294.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, 22(1):3–42.

Mahroum, S. (2001). Europe and the immigration of highly skilled labour. *International Migration*, 39(5):27–43.

Martínez, C. (2010). Patent families: when do different definitions really matter? *Scientometrics*, 86(1):39–63.

Maskus, K. E. (2000). *Intellectual property rights in the global economy*. Peterson Institute.

Maznevski, M. L. and Chudoba, K. M. (2000). Bridging space over time: Global virtual team dynamics and effectiveness. *Organization science*, 11(5):473–492.

Maznevski, M. L. and DiStefano, J. J. (2000). Global leaders are team players: Developing global leaders through membership on global teams. *Human resource management*, 39(2, 3):195.

McCulloch, R. and Yellen, J. L. (1977). Factor mobility, regional development, and the distribution of income. *Journal of Political Economy*, 85(1):79–96.

McGrath, J. E., Berdahl, J. L., and Arrow, H. (1995). Traits, expectations, culture, and clout: The dynamics of diversity in work groups.

McLeod, P. L. and Lobel, S. A. (1992). The effects of ethnic diversity on idea generation in small groups. In *Academy of Management Proceedings*, volume 1992, pages 227–231. Academy of Management.

Melitz, J. and Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics*, 93(2):351–363.

Miguelez, E. (2016). Inventor diasporas and the internationalization of technology. *The World Bank Economic Review*, page lhw013.

Miguelez, E., Fink, C., et al. (2013). Measuring the international mobility of inventors: A new database. *World Intellectual Property Organization-Economics and Statistics Division*.

Mingers, J. and Xu, F. (2010). The drivers of citations in management science journals. *European Journal of Operational Research*, 205(2):422–430.

Molesworth, M., Scullion, R., and Nixon, E. (2010). *The marketisation of higher education*. Routledge.

Morrison, G., Riccaboni, M., and Pammolli, F. (2017). Disambiguation of patent inventors and assignees using high-resolution geolocation data. *Scientific Data*, 4.

Mortensen, M. and Hinds, P. J. (2001). Conflict and shared identity in geographically distributed teams. *International Journal of Conflict Management*, 12(3):212–238.

Moser, P., Ohmstedt, J., and Rhode, P. (2011). Patents, citations, and inventive outputevidence from hybrid corn. *NBER Working Paper*, (w21443).

Moser, P., Voena, A., and Waldinger, F. (2014). German jewish émigrés and us invention. *The American Economic Review*, 104(10):3222–3255.

Mountford, A. (1997). Can a brain drain be good for growth in the source economy? *Journal of development economics*, 53(2):287–303.

Murray, A. I. (1989). Top management group heterogeneity and firm performance. *Strategic management journal*, 10(S1):125–141.

Nall, C. (2015). The political consequences of spatial policies: How interstate highways facilitated geographic polarization. *The Journal of Politics*, 77(2):394–406.

Nathan, M. (2014a). Same difference? minority ethnic inventors, diversity and innovation in the uk. *Journal of Economic Geography*, 15(1):129–168.

Nathan, M. (2014b). The wider economic impacts of high-skilled migrants: a survey of the literature for receiving countries. *IZA Journal of Migration*, 3(1):4.

Nicolaas, H. and Sprangers, A. (2001). Migration motives of non-dutch immigrants in the netherlands. In *Paper for EAPS Population Conference, Helsinki, Finland. June*, pages 7–9.

Niebuhr, A. (2010). Migration and innovation: Does cultural diversity matter for regional r&d activity? *Papers in Regional Science*, 89(3):563–585.

Nielsen, B. B. and Nielsen, S. (2011). The role of top management team international orientation in international strategic decision-making: The choice of foreign entry mode. *Journal of World Business*, 46(2):185–193.

Nielsen, B. B. and Nielsen, S. (2013). Top management team nationality diversity and firm performance: A multilevel study. *Strategic Management Journal*, 34(3):373–382.

Nijkamp, P. and Poot, J. (2015). Cultural diversity: a matter of measurement. *The Economics of Cultural Diversity*, pages 17–51.

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

OECD (2016). International migration outlook 2016.

Orrenius, P. M. and Zavodny, M. (2005). Self-selection among undocumented immigrants from mexico. *Journal of Development Economics*, 78(1):215–240.

Østergaard, C. R., Timmermans, B., and Kristinsson, K. (2011). Does a different view create something new? the effect of employee diversity on innovation. *Research Policy*, 40(3):500–509.

Ottaviano, G. I. and Peri, G. (2006). The economic value of cultural diversity: evidence from us cities. *Journal of Economic Geography*, 6(1):9–44.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, 10(1):152–197.

Özden, Ç., Parsons, C. R., Schiff, M., and Walmsley, T. L. (2011). Where on earth is everybody? the evolution of global bilateral migration 1960–2000. *The World Bank Economic Review*, 25(1):12–56.

Ozgen, C., Nijkamp, P., and Poot, J. (2011). Immigration and innovation in european regions. *Tinbergen Institute Discussion Paper*.

Ozgen, C., Nijkamp, P., and Poot, J. (2013). The impact of cultural diversity on firm innovation: evidence from dutch micro-data. *IZA Journal of Migration*, 2(1):18.

Ozgen, C., Peters, C., Niebuhr, A., Nijkamp, P., and Poot, J. (2014). Does cultural diversity of migrant employees affect innovation? *International Migration Review*, 48(s1):S377–S416.

Parrotta, P., Pozzoli, D., and Pytlikova, M. (2012). Does labor diversity affect firm productivity?

Parrotta, P., Pozzoli, D., and Pytlikova, M. (2014). The nexus between labor diversity and firms innovation. *Journal of Population Economics*, 27(2):303–364.

Paserman, M. D. (2013). Do high-skill immigrants raise productivity? evidence from israeli manufacturing firms, 1990-1999. *IZA Journal of Migration*, 2(1):6.

Patrashkova-Volzdoska, R. R., McComb, S. A., Green, S. G., and Compton, W. D. (2003). Examining a curvilinear relationship between communication frequency and team performance in cross-functional project teams. *IEEE Transactions on Engineering Management*, 50(3):262–269.

Pechurina, A. (2017). Post-soviet russian-speaking migration to the uk: The discourses of visibility and accountability. In *Post-Soviet Migration and Diasporas*, pages 29–45. Springer.

Perko, J. S. and Narin, F. (1997). The transfer of public science to patented technology: A case study in agricultural science. *The Journal of Technology Transfer*, 22(3):65–72.

Pezzoni, M., Lissoni, F., and Tarasconi, G. (2014). How to kill inventors: testing the massacrator© algorithm for inventor disambiguation. *Scientometrics*, 101(1):477–504.

Polzer, J. T., Crisp, C. B., Jarvenpaa, S. L., and Kim, J. W. (2006). Extending the faultline model to geographically dispersed teams: How colocated subgroups can impair group functioning. *Academy of Management Journal*, 49(4):679–692.

Portnova, M. (1998). Ownership and enforcement of patent rights in russia: Protecting an invention in the existing environment. *Indiana International & Comparative Law Review*, 8(2):505–549.

Pradhan, R. P., Arvin, M. B., Hall, J. H., and Nair, M. (2016). Innovation, financial development and economic growth in eurozone countries. *Applied Economics Letters*, 23(16):1141–1144.

Prahalad, C. K. and Bettis, R. A. (1986). The dominant logic: A new linkage between diversity and performance. *Strategic management journal*, 7(6):485–501.

Ravlin, E. C., Thomas, D. C., and Ilsev, A. (2000). Beliefs about values, status, and legitimacy in multicultural groups: Influences on intra-group conflict. *Innovations in international and cross-cultural management*, pages 17–51.

Richard, O. C., Kirby, S. L., and Chadwick, K. (2013). The impact of racial and gender diversity in management on financial performance: How participative strategy making features can unleash a diversity advantage. *The International Journal of Human Resource Management*, 24(13):2571–2582.

Richardson, J. and Zikic, J. (2007). The darker side of an international academic career. *Career Development International*, 12(2):164–186.

Ronen, S. and Shenkar, O. (2013). Mapping world cultures: Cluster formation, sources and implications. *Journal of International Business Studies*, 44(9):867–897.

Sakwa, R. et al. (2005). *The rise and fall of the Soviet Union*. Routledge.

Saltykov, B. G. (1997). The reform of russian science. *Nature*, 388(6637):16–18.

Saxenian, A. (2002). Silicon valleys new immigrant high-growth entrepreneurs. *Economic development quarterly*, 16(1):20–31.

Schettino, F., Sterlacchini, A., and Venturini, F. (2013). Inventive productivity and patent quality: Evidence from italian inventors. *Journal of Policy Modeling*, 35(6):1043–1056.

Schmoch, U. (2008). Concept of a technology classification for country comparisons. *Final report to the world intellectual property organisation (wipo), WIPO*.

Scott Long, J. (1997). Regression models for categorical and limited dependent variables. *Advanced quantitative techniques in the social sciences*, 7.

Shevtsova, L. (1992). Post-soviet emigration today and tomorrow. *International Migration Review*, pages 241–257.

Simonton, D. K. (1988). *Scientific genius: A psychology of science*. Cambridge University Press.

Singh, J. and Fleming, L. (2010). Lone inventors as sources of breakthroughs: Myth or reality? *Management science*, 56(1):41–56.

Soltysinski, S. (1969). New forms of protection for intellectual property in the soviet union and czechoslovakia. *The Modern Law Review*, 32(4):408–419.

Soyfer, V. N. (2001). The consequences of political dictatorship for russian science. *Nature Reviews Genetics*, 2(9):723–729.

Stahl, G. K., Mäkelä, K., Zander, L., and Maznevski, M. L. (2010). A look at the bright side of multicultural team diversity. *Scandinavian Journal of Management*, 26(4):439–447.

Stahl, G. K. and Tung, R. L. (2015). Towards a more balanced treatment of culture in international business studies: The need for positive cross-cultural scholarship. *Journal of International Business Studies*, 46(4):391–414.

Stephan, P. E. and Levin, S. G. (2001). Exceptional contributions to us science by the foreign-born and foreign-educated. *Population research and Policy review*, 20(1):59–79.

Sternitzke, C. (2009). Defining triadic patent families as a measure of technological strength. *Scientometrics*, 81(1):91–109.

Sterzi, V. (2013). Patent quality and ownership: An analysis of uk faculty patenting. *Research Policy*, 42(2):564–576.

Strayer, R. (1998). *Why did the Soviet Union collapse?: understanding historical change*. ME Sharpe.

Stuen, E. T., Mobarak, A. M., and Maskus, K. E. (2012). Skilled immigration and innovation: evidence from enrolment fluctuations in us doctoral programmes. *The Economic Journal*, 122(565):1143–1176.

Suny, R. (1993). *The revenge of the past: Nationalism, revolution, and the collapse of the Soviet Union*. Stanford University Press.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual review of psychology*, 33(1):1–39.

Tetlock, P. E. (2000). Cognitive biases and organizational correctives: Do both disease and cure depend on the politics of the beholder? *Administrative Science Quarterly*, 45(2):293–326.

Tolts, M. (2009). Post-soviet aliyah and jewish demographic transformation. In *15th World Congress of Jewish Studies, Jerusalem, Israel*, pages 2–6. Citeseer.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *The Rand Journal of Economics*, pages 172–187.

Trax, M., Brunow, S., and Suedekum, J. (2015). Cultural diversity and plant-level productivity. *Regional Science and Urban Economics*, 53:85–96.

Van Knippenberg, D., De Dreu, C. K., and Homan, A. C. (2004). Work group diversity and group performance: an integrative model and research agenda. *Journal of applied psychology*, 89(6):1008.

Voutira, E. A. (2006). Post-soviet diaspora politics: The case of the soviet greeks. *Journal of modern Greek studies*, 24(2):379–414.

Wadhwa, V., Saxenian, A., Rissing, B. A., and Gereffi, G. (2007). America's new immigrant entrepreneurs: Part i.

Waldinger, R. D. (1986). *Through the eye of the needle: Immigrants and enterprise in New York's garment trades*. New York: New York University Press.

Wang, D. (2015). Activating cross-border brokerage: Interorganizational knowledge transfer through skilled return migration. *Administrative Science Quarterly*, 60(1):133–176.

Watkins, A. and Ferrara, E. L. (2005). Ethnic diversity and economic performance. *Journal of economic literature*, 43(3):762–800.

Watson, W. E., Johnson, L., and Zgourides, G. D. (2002). The influence of ethnic diversity on leadership, group process, and performance: An examination of learning teams. *International Journal of Intercultural Relations*, 26(1):1–16.

Wiersema, M. F. and Bantel, K. A. (1992). Top management team demography and corporate strategic change. *Academy of Management journal*, 35(1):91–121.

Williams, K. Y. and O'Reilly III, C. A. (1998). Demography and diversity in organizations. a review of 40 years of research. *Research in organizational behavior*, 20:77–140.

WIPO (2017). Guide to the international patent classification, version 2017.

Wongel, H. (2005). The reform of the ipc: consequences for the users. *World Patent Information*, 27(3):227–231.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wuchty, S., Jones, B. F., and Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827):1036–1039.

Zajac, E. J., Golden, B. R., and Shortell, S. M. (1991). New organizational forms for enhancing innovation: The case of internal corporate joint ventures. *Management Science*, 37(2):170–184.

Zenun, M., Loureiro, G., and Araujo, C. (2007). The effects of teams co-location on project performance. *Complex systems concurrent engineering*, pages 717–726.