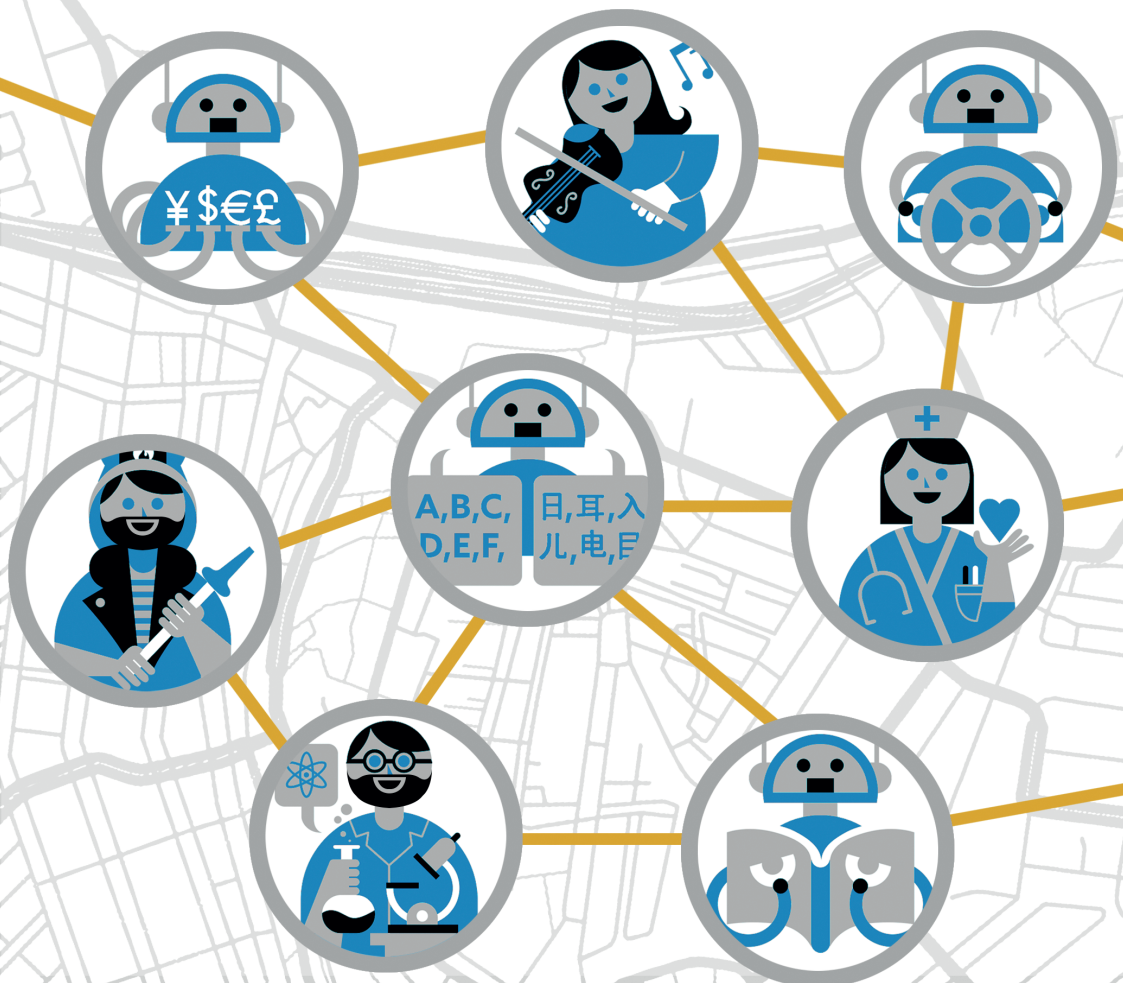


THE EVOLVING GEOGRAPHY OF JOBS

HOW RELATEDNESS
SHAPES LABOUR DYNAMICS



TERESA FARINHA

THE EVOLVING GEOGRAPHY OF JOBS:
HOW RELATEDNESS SHAPES LABOUR DYNAMICS

Teresa Farinha

Doctoral Dissertation

Farinha, T.

The evolving geography of jobs: how relatedness shapes labour dynamics

DOI: <https://doi.org/10.33540/413>

ISBN: 978-94-6421-097-2

Inside layout: Lara Leijtens, persoonlijkproefschrift.nl

Printed and published by: Ipskamp Printing, proefschriften.net

Cover: Tiago Albuquerque and Adriano Lameira (<http://tiagoalbuquerque.blogspot.com>)

THE EVOLVING GEOGRAPHY OF JOBS:

HOW RELATEDNESS SHAPES LABOUR DYNAMICS

**De evolutie in de geografie van banen: hoe
gerelateerdheid de arbeidsmarktdynamiek vorm geeft**

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 14 december 2020 des middags te 12.45 uur

door

Teresa Margarida Antunes Farinha Fernandes

geboren op 25 januari 1982
te Lissabon, Portugal

Promotor:

Prof. dr. R.A. Boschma

Copromotoren:

Dr. A. Morrison

Dr. P.M.A. Balland

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van MIT MediaLab,
Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

Committee: Prof. Dr. Carolina Castaldi
Prof. Dr. Koen Frenken
Dr. César Hidalgo
Prof. Dr. Simona Iammarino
Prof. Dr. Frank van Oort

“(…) it will be those peoples, who can keep alive,
and cultivate into a fuller perfection, the art of life itself
and do not sell themselves for the means of life, who will be able
to enjoy the abundance when it comes”
(Keynes, 1930, pp. 358-373)

CONTENTS

| | |
|---|----|
| Acknowledgements | 10 |
| Abstract | 12 |
| 1. Introduction | 15 |
| 1. Motivation | 16 |
| 2. Research questions | 18 |
| 3. Outline of this thesis | 21 |
| 2. Urban occupational structures as information networks: the effect on network density of increasing number of occupations | 23 |
| 1. Introduction | 24 |
| 2. Modeling and estimation frameworks | 26 |
| 3. Results | 28 |
| 4. Discussion | 32 |
| 5. Materials and methods | 34 |
| 5.1. Data | 34 |
| 5.2. Occupational interdependence | 36 |
| 5.3. Density and size of urban occupational networks | 39 |
| 3. What drives the geography of jobs in the US: Unpacking relatedness | 43 |
| 1. Introduction | 44 |
| 2. Regional diversification in jobs: three mechanisms | 45 |
| 3. Occupational Data and Network analysis | 47 |
| 3.1. Occupational data | 47 |
| 3.2. Unpacking relatedness | 49 |
| 3.3. The Job Space – a descriptive analysis | 52 |
| 4. Relatedness Dimensions and the Renewal of the Job-Space | 53 |
| 4.1. Variables and descriptives | 53 |
| 4.2. Entry and Exit Models – Only Geographical Relatedness measure | 57 |
| 4.3. Entry and Exit Models – All Dimensions of Relatedness | 59 |
| 5. Concluding remarks | 64 |
| 4. Impacts from automation diffuse locally: a novel approach to estimate jobs risk in US cities | 67 |
| 1. Introduction | 68 |
| 2. What determines a job's risk of automation? | 69 |
| 2.1 Technical feasibility | 69 |
| 2.2 Local feasibility | 70 |

| | | |
|------|--|-----|
| 3. | Employment data | 74 |
| 3.1 | The Job Space Under Automation | 75 |
| 3.2 | Relatedness to high-risk-jobs | 78 |
| 4. | Impact of automation on employment growth in US cities | 80 |
| 4.1 | Variables and descriptives | 80 |
| 4.2 | Job-City employment growth under automation | 83 |
| 4.3 | Robustness analysis | 87 |
| 5. | Concluding remarks | 88 |
| 5. | EU Incentives to Business Increased Job Quality In Portuguese Firms: Evidence from Linked Employer-Employee Data and Natural Experiment Conditions | 101 |
| 1. | Introduction | 102 |
| 2. | Previous Literature | 104 |
| 3. | POE/PRIME Program, Data and Descriptive Statistics | 107 |
| 4. | Empirical Strategy | 110 |
| 5. | Results | 116 |
| 6. | Sensitivity Analysis | 117 |
| 7. | Concluding remarks and policy implications | 119 |
| 6. | Conclusion | 123 |
| 1. | Overview and research contributions | 124 |
| 1.1. | On the “magnet effects” of relatedness | 125 |
| 1.2. | On the “diffusion effects” of relatedness | 127 |
| 1.3. | On how innovation policy can boost job quality | 129 |
| 2. | Limitations and future research | 130 |
| 3. | Policy implications | 133 |
| | Bibliography | 138 |
| | Nederlandse samenvatting | 154 |
| | Curriculum vitae | 155 |

ACKNOWLEDGEMENTS

What a journey! It had cry and laughter, boredom and adventure, weakness and strength, and lots of WOW! I've learned so much, on all kinds of teachings, and in interaction with great people that I was lucky to have along my journey. So many that it is not possible to name everyone involved in creating this doctoral dissertation. It could go as far as reaching Buddha. Because, guess what: everything is related! I want, however, to acknowledge some great people that were particularly present during these last five years.

My mom, Maria Isabel Antunes, is the strongest and most loving woman I know. She has the power to communicate with anyone, from all languages. She only uses her motivation, a few Portuguese words just for the beauty, and if you're lucky, a superb traditional Portuguese dish. My 2x-cousin and buddy 4ever, Gabriela Farinha, the sharpest woman I know. Her unconditional support and always wise advice, besides helping me get rid of my troubles, also taught me to do the same: to be there for my beloved ones. My Dutch handsome prince, Jorrit Kuiper, the most loving partner I could ever dream of. He has enriched my journey with lots of care and fun. I have learned greatly from his practicality, as an exemplar Dutch, and his marvellous cooking, way beyond Dutch standards! My brother, João Farinha Fernandes, the most brilliant economist. He is solid like a rock, strong like a lion, protective as a big brother. I'm deeply thankful for his support so many times along the journey. The Pandilha, a group of fun loyal loving friends that live in my heart and to whom I can reach at any time, always. Pedro, Maria, Muriel, Hang, Tatiana, Paulo, Gabriel, Tommaso, Raquel, Margarida, Ana, Juan, Eduardo have witnessed my best and my worst, and are still with me every single day! My new family, António, Marcella, and Tjeerd, their contagious joy and genuine interest in my crazy ideas greatly helped them flourish into this doctoral dissertation.

My supervisor Ron Boschma, who's wisdom goes far beyond what you can grasp from his extensive list of remarkable publications. From him, I got the best inspiration to develop the diversified set of skills that makes an excellent scholar (not to be assumed I actually got them!). His precious teachings and support, as supervisor and a friend, accompanied me under the hardest and sweetest times. With my co-supervisors Pierre-Alexandre Balland and Andrea Morrison, I could never feel alone in the lonely odyssey of a PhD. They rooted for me, gave me camaraderie, openness, patience, and prepared me for the

biggest challenges. I got to experience the fascinating world of MIT MediaLab, as well as the medical top-frontier technology in Boston. The first was the perfect place to study the future of jobs, while the latter, the result of an unfortunate bike accident. Still, I was very lucky to have César Hidalgo's outstanding support during all times. Their teachings, I'll treasure my whole life.

I'm also greatly thankful to all my co-authors, especially Ricardo Mamede, who introduced me to the world of policy evaluation and shared with me the open call to start this PhD in Utrecht University. Also, to Shade Shutters, with whom I have had exciting discussions and, in a happy coincidence, very much aligned research interests since the very beginning of my PhD. Among many other brilliant scholars that I had the joy of meeting, they will certainly continue inspiring my work. And of course, to my favourite illustrators, Tiago Albuquerque and Adriano Lameira, who brilliantly depicted the essence of this research work in a beautiful thesis cover.

Along the way, I'm fortunate to have made great friends with, but not limited to, Diana, Flávio, Annick, Severin, Martina, Isabella, Moritz, Philippa, Julie, Mary, Christian, Cristian, Carla, Márcia, Wolf, Mathilde, Alex, Paula, Andrea, Lakshmi, Thomas, Diego, Vivian, Sergio, Markus, Mathew, Stef, Nicola, Tom, Joan, Dario, Marielle, Karin, Mathias, Summer, Mirjam, Nynke, Karlijn, Hannah and Paulien. They gave me shelter during my PhD trips and showed me the vibrant cities of Utrecht, Amsterdam, Copenhagen, Toulouse, Chicago, Boston, and San Diego. We have shared great times together, including bootcamps, hiking, scuba diving, sailing, artistic dancing, and YO – yoga in the office.

I could not have gotten here without all of you. To all my superheroes, a big thank you for having inspired me immensely, in so many ways. Please remind me to offer you a Dutch beer or a Portuguese glass of wine next time we meet. Um bem haja!

ABSTRACT

Jobs disappear, jobs are created, work changes, in a geographically uneven way. Some places cope better with adverse events that threaten its labour force into unemployment. Some places end up having better jobs. Why? Labour dynamics are particular of each place, influenced by a myriad of local factors, such as the type of jobs in each city. But jobs do not stand alone. Workers interact with and influence each other within and across firms. This ends up reflecting in the structure of labour of each city, region, or country. At each point in time, the structure sets the opportunities and boundaries for labour dynamics to unravel.

This doctoral thesis investigates the way jobs relate to each other – relatedness – and how it shapes the evolving geography of jobs. Several bodies of literature – Evolutionary Economic Geography (EEG), Labour Economics, Urban Scaling, Innovation Studies, Regional Policy – come together to better understand how relatedness shapes labour dynamics in different spatial contexts.

First, using employment data of six industrialized countries, we find that bigger cities have more than proportionally higher levels of relatedness between jobs, than smaller cities – relatedness self-reinforces as cities grow. Second, relatedness promotes job diversification and prevents exit of job specializations in USA cities – “magnet effects” – in three distinct ways: agglomeration of jobs that are complementary, and/or similar, and/or synergic. Third, the impacts of automation spread through the structure of relatedness – “diffusion effects”. More concretely, being complementary, but not similar, to local high-risk-jobs increases employment grow of a job in a USA city. Finally, policy can explore the potential of relatedness effects to reach desired outcomes. For instance, by stimulating the structure of relatedness around highly specialized jobs, which tend to be denser in innovative sectors. Accordingly, we found EU policy business incentives to innovation to have increased job quality in Portuguese firms.

These findings may help design policy instruments that neutralize the negative effects of automation, while promoting the positive impacts. For instance, identifying which jobs in which cities might be at higher risk, given jobs’ technical feasibility of automation, but also how high-risk-jobs spread automation impacts to other jobs in each city. Workers will be in greater need of social support and training programs especially where their similarities to high-risk-jobs in the city

out rule the complementarities. Moreover, place-based policy instruments can target the above-mentioned relatedness effects to foster employment with job quality in lagging regions. I hope this thesis inspires future policy-science collaboration to nurture a future with good jobs for all.

Keywords:

Labour dynamics, jobs, cities, relatedness, agglomeration, automation, policy, innovation, job-quality.

CHAPTER 1

Introduction



1. MOTIVATION

Jobs move, jobs change. And some places end up having better jobs than other places. There were times when a doctor would also be the assisting nurse, and driver to patients' homes. Nowadays, a doctor alone can hardly treat a single patient. Instead, together with other specialized workers in an hospital, doctors can treat patients more promptly and efficiently than ever. But not everywhere. Hospitals of large cities tend to offer better working conditions than in countryside, with better salaries and cutting-edge technology, attracting the best doctors. Conversely, cheaper labour in developing countries has attracted medical equipment suppliers to move factories there, destroying manufacturing jobs in their homeland. Those jobs might now return home, but to be executed by modern robots instead. What determines the evolving geography of jobs?

New technologies have always pushed the evolution of work. They create new types of jobs, but they can also make existing jobs obsolete. At early 20th century, farms employed 40% of USA labour force. Few decades later, mechanization brought farms' employment share down to 2% and expanded the services sector (Autor, 2015). More recently, new digital technologies, such as Artificial Intelligence and Modern Robotics, change the nature of work again. Recent studies have tried to assess the impacts of automation on jobs. Yet, their results diverge radically, from 47% of USA jobs at risk of automation (Frey & Osborne, 2013) to 9% in OECD countries (Arntz, Gregory, & Zierahn, 2016), among many other estimations (Berriman et al., 2017; Manyika et al., 2017). A common caveat is to assess risk solely based on which tasks/jobs are suitable to automation (Brynjolfsson, Mitchell, & Rock, 2018). However, two workers from different cities with the same type of job may have different risk of automation (Nedelkoska and Quintini, 2018).

If we want to understand labour dynamics under automation, we cannot look at jobs in isolation. The way jobs are related to each other – relatedness – also condition labour dynamics. Organizing work into jobs implies that workers are brought together to create value for the firm, the client, the state, etc. Together, jobs form complex input-output chains, specialized labour pools, vibrant economic niches, i.e., a labour system. The performance of each job affects neighbouring ones, ultimately shaping the local structure of jobs. This has been acknowledged by the recent literature of Evolutionary Economic Geography (EEG) that has investigated the effects of relatedness on labour dynamics in

regions (Boschma, 2017). Empirical studies show that regions tend to diversify into new job specializations (Alabdulkareem et al., 2018; Muneeppeerakul, Lobo, Shuttters, Gómez-Liévano, & Qubbaj, 2013) that are related to pre-existing ones. And besides favouring entry, relatedness also seems to prevent the exit of existing job specializations.

However, there are still a number of gaps that still need to be investigated in the relatedness literature in connection to labour dynamics. This thesis addresses four of them

Which cities have higher levels of relatedness? The Urban Scaling literature has shown that large human agglomerations, such as big cities, show disproportionately higher concentration of economic activities, levels of labour division, and economic complexity (Bettencourt, 2013; Hidalgo, C., Klinger, Barabási, & Hausmann, 2007), than smaller ones. They also show economic growth (Frenken, Van Oort, & Verburg, 2007; Hidalgo, C. et al., 2007), employment levels, wages, local prices, and workers' welfare (Beaudry, Green, & Sand, 2012; Delgado, Porter, & Stern, 2014; Moretti & Thulin, 2013). But it is still not clear how relatedness scales with city size in terms of number of specializations.

What type(s) of relatedness matters for the entry/exit of jobs in cities? Relatedness seems to affect local labour dynamics in many ways. For instance, relatedness tends to act like a "magnet", pulling capabilities together in a city, region, or country (Hidalgo et al., 2018) – "magnet effects". It does so in distinct ways (Jacobs, 1969; Marshall, 1920; Porter, 1990). However, existing empirical studies rarely specify which force(s) of agglomeration is being captured when measuring relatedness. Different kinds of relatedness might show different labour dynamics. By distinguishing different kinds of relatedness, we open the black box of relatedness (Boschma, 2017; Tanner, 2014).

How will relatedness between jobs influence the impacts of automation on employment in cities? As outlined before, highly disruptive technologies reshuffle the demand for labour (Autor, 2015). Jobs get destroyed (e.g. taxi drivers), changed (e.g., self-driving car mechanics), and newly created (e.g., drone managers, or augmented reality architects). The resulting impacts are geographically uneven (Brynjolfsson et al., 2018; Nedelkoska & Quintini, 2018). Yet, we still know little on how geography influences labour dynamics under automation. For sure, it depends on the portfolio of jobs in each region: the

more jobs are at risk in a city, the more a city will be negatively affected by automation. But jobs do not stand alone. Jobs interact with each other, forming relatedness ties that accommodate flows of knowledge, but also of impacts between them – “diffusion effects”. For instance, the past Great Recession, initially created by a financial shock in USA in 2007, ended up spreading to the whole world economy. We have little to no knowledge on how the “diffusion effects” of relatedness might condition the local impacts of automation.

How to incorporate relatedness in regional policy to foster job quality in firms’ workforce? The EEG literature has provided crucial knowledge on what determines the evolution of regional economies (Boschma, 2015). There are a few studies that have recently incorporated relatedness in a smart specialization policy framework to help policy define directions for future regional diversification (Balland, Boschma, Crespo, & Rigby, 2019). However, it has not yet been connected to job quality. Policy can make use of local labour dynamics to reach desired outcomes in the labour force. For instance, specialized jobs in innovative industries tend to have higher productivity and labour demand, which gives them bargaining power for better working conditions (Clark, 2005; Harter, Schmidt, & Keyes, 2004; Moretti & Thulin, 2013). Such jobs usually have high employment multiplier effects (Moretti, 2012), i.e., high density of relatedness in the local economy through which they benefit other jobs. This begs the question how relatedness can be incorporated in regional policy to enhance job quality in local firms.

2. RESEARCH QUESTIONS

This thesis aims to tackle these four research gaps, one by one, in four empirical chapters. Doing so, this thesis contributes to our knowledge on how relatedness affects labour dynamics and the evolution of work in regions, and how relatedness can be integrated in regional policy framework to secure job quality in regions. The main research question that bounds this doctoral research is the following:

RQ: How does relatedness between jobs shape the geography of work?

Based on the four literature gaps discussed before, we address the main research question in four specific research questions, each of them covered in one empirical chapter.

RQ1: Does relatedness scale with city size?

Chapter 2 addresses RQ1 and employs an exponential model that tests whether city relatedness density (number of links in the network) scales with city diversity (city size, i.e., number of job specializations), against the null hypothesis of proportional growth. We use employment data at the city-job level in 2013 for six countries (United States, Canada, Sweden, Portugal, Australia, and Germany). We define a 'Job Space' which is a network representation of jobs in which nodes represent job specializations of a city, and links show the degree of relatedness between them, measured as geographical relatedness, following Hidalgo et al. (2007). Our results indicate that relatedness grows superlinearly with city diversity in all six countries. More concretely, the estimated scaling exponents range from 1.17 in Swedish cities to 2.35 in USA cities. This means that bigger cities have more than proportionally higher levels of relatedness between existing job specializations than smaller cities.

RQ2: Which dimension(s) of relatedness drive the entry and exit of jobs in cities?

Chapter 3 (RQ2) takes a dynamic and deeper approach into the "magnet effects" of relatedness. It uses employment data (city-job, industry-job, and task-job levels of analysis) from US BLS and O*NET (2005 to 2016) to unpack relatedness into three orthogonal dimensions – complementarity, similarity, and synergy. We transform the "Job Space" into a multi-layer network with three types of links, one for each dimension of relatedness. Complementarity is measured as how often two job specializations co-occur in the same input-output chain. Similarity is measured as how much their tasks overlap. Local synergy is derived from geographical relatedness. We test their impact on a job's probability of entering and exiting a city's portfolio of job specializations. Relatedness shows a positive effect on entry of new job specializations and a negative effect on exit of existing job specializations for all three relatedness dimensions, yet in different magnitudes. Local synergy shows far stronger coefficients.

RQ3: What is the impact of automation through relatedness on job growth in cities?

Chapter 4 (RQ3) investigates which relatedness links accommodate the diffusion of impacts from automation. Recent literature indicates that recently automated jobs provide productivity gains to complementary neighbouring jobs, while making it easier to automate similar jobs (Autor, 2015; Nedelkoska & Quintini, 2018). If so, the way jobs are related to each other in a city also determines

automation impacts, besides a job's share of automatable tasks and a city's particular portfolio of jobs. To test this, the "Job Space is extended once more. This time, it accounts for different types of jobs, with regard to their technical risk of automation (data from Atkinson (2017), BLS and O*NET, 2005 to 2016). A simple transformation of a commonly used relatedness measure can capture jobs' level of relatedness to high-risk-jobs in a city. This can also be computed for specific types of relatedness, such as complementarity and similarity. Results confirm the existence of "diffusion effects". Being complementary to local high-risk jobs, but not similar to them, seems to have increased employment at the job-city level, between 2007 and 2016.

The final specific research question addresses the capacity of regional policy to, indirectly, through the local structure of relatedness, foster job quality in firms' workforce. It follows below:

RQ4: What is the impact of policy incentives for business through relatedness on job quality?

Finally, Chapter 5 (RQ4) evaluates the capacity of innovation policy to increase firms' job quality, indirectly, by making use of the structure of relatedness in each region. The policy instrument under analysis is EU business incentives aimed to boost firms' competitiveness. Most selected firms belong to innovative sectors. They are expected to acquire the necessary resources that ensure the success of the project and boost the firm's future performance. This includes physical and human capital, but also the formation/strengthening of specialized teams and collaboration networks, Firms may need to hire highly specialized workers, as well as their complementary workers, beyond the project duration. And must offer attractive working conditions (specialized workers tend to have higher labour demand). After the project, the firm can be expected to have improved its levels of job quality. To Test this, we use Portuguese linked-employer-employee data (2000 to 2009) and EU supported projects data and apply Counterfactual Impact Evaluation methods (dynamic matching, difference-in-differences, and block-randomized pretest-posttest design). We find significant positive impacts on job quality. More concretely, three years after treatment, each treated firm seems to have generated, on average, additional 5 standard-working-time jobs, 3 skilled jobs, 2 permanent-contract jobs, and 17.8% higher remuneration (+2.20€/hour), in comparison with non-treated firms. This shows how the local structure of capabilities in innovative sectors can be stimulated to boost employment and job quality.

3. OUTLINE OF THIS THESIS

This thesis combines different bodies of literatures (EEG, Labour Economics, Innovation Studies, Urban Scaling, and Regional Policy) to provide empirical evidence on how relatedness affects labour dynamics and shapes the geography of work. Each of the four empirical chapters (Chapter 2 to Chapter 5) is an independent scientific article, published or submitted for publication. Three of them resulted from collaborative research, and one is single-authored. The full references are mentioned in the footnotes below.

Chapter 2 investigates whether relatedness scales with city size in six countries (United States, Canada, Sweden, Portugal, Australia, and Germany)¹. Chapter 3 unpacks relatedness in three dimensions – similarity, complementarity, and synergy, to test their impact on jobs’ probability of entry and exit in a city’s portfolio of job specializations². Chapter 4 provides empirical evidence on which relatedness links accommodate the diffusion of impacts from automation in US cities³. Chapter 5 evaluates the impacts of policy incentives to innovation in Portuguese firms on the level of job quality⁴.

Finally, Chapter 6 concludes by summarizing the research contributions presented in this thesis. It also discusses limitations and suggests new research avenues that will deepen our knowledge on local labour dynamics. It will also discuss policy implications of our findings, and how policy can use them to support a sustainable transition of skills in the workforce towards good jobs for all.

1 Shutters, S., Lobo, J., Muneeppeerakul, R., Strumsky, D., Mellander, C., Brachert, M, Farinha, T., and Bettencourt, L. (2018) Urban occupational structures as information networks: The effect on network density of increasing number of occupations. *PLoS ONE* 13(5): e0196915. <https://doi.org/10.1371/journal.pone.0196915>

2 Farinha, T., Balland, P.-A., Morrison, A., and Boschma, R. (2019) What drives the geography of jobs in the US? Unpacking relatedness, *Industry and Innovation*, 26:9, 988-1022, DOI: 10.1080/13662716.2019.1591940

3 Farinha, T. (2020) “Impacts from automation diffuse locally: a novel approach to estimate jobs risk in us cities”, *Papers in Evolutionary Economic Geography*, Section of Economic Geography, Utrecht University

4 Bondonio, D., Farinha, T., Mamede, R. (currently under review). “EU incentives to business increased Job Quality in Portuguese Firms: Evidence from Linked Employer-Employee Data and Natural Experiment Conditions of EU Cohesion Policy in Portugal”, *Evaluation Review Special Issue*, *Sage Journals* (manuscript submitted and invited to revise and resubmit)

CHAPTER 2

Urban occupational structures as information networks: the effect on network density of increasing number of occupations

Contribution:

Data curation (data for Portugal) and writing (original draft, review & editing)

Coauthors:

Shade T. Shutters, José Lobo, Rachata Muneeppeerakul, Deborah Strumsky, Charlotta Mellander, Matthias Brachert, Luis M. A. Bettencourt

Published as:

Shade Shutters, José Lobo, Rachata Muneeppeerakul, Deborah Strumsky, Charlotta Mellander, Matthias Brachert, Teresa Farinha, and Luís Bettencourt (2018) Urban occupational structures as information networks: The effect on network density of increasing number of occupations. PLoS ONE 13 (5): e0196915. <https://doi.org/10.1371/journal.pone.0196915>

1. INTRODUCTION

Urban economies are intricate webs of exchange, linking specialized production units and manifesting divisions of labour and knowledge flows (Acemoglu et al., 2012; Atalay et al., 2011; Balland and Rigby, 2017; Fujiwara and Aoyama, 2010; Hausmann and Hidalgo, 2011; Hidalgo and Hausmann, 2009). The specific goods and services such units can provide, and how well they provide them, are largely determined by the technologies, skills, and tacit knowledge integrated in the process of value creation. The interconnections among these technologies and skills form an economic structure – a structure that enables some developmental pathways while foreclosing others. A city’s transition from one type of economy to another must ultimately alter its current underlying economic structure by breaking down some parts and building up others. Thus, it is critical to develop a comprehensive understanding of the properties and dynamics of the structures underlying urban economies.

The connections in an economic network represent acts of exchange involving the flow of capital, the transportation of goods, the movement of people or, importantly, the transmission of information (Bettencourt, 2014). It is the exchange and flow of information, mediated by economic markets as well as by other channels, that makes it possible to coordinate activities, generate complementarities, and self-organize production and consumption (Hayek, 1945; Arrow, 1974). These lines of reasoning underpin the expectation that the productivity of an economy (national or urban) should increase as its level of connectivity among its constituent units increases. Under “network effects” the value of a product or service is dependent on the number of others using it implying increasing returns to network size (Shapiro and Varian, 1998), where size is the number of individual users. A similar theme runs through other networked processes. “Metcalfe’s Law” (Hendler and Golbeck, 2008), “Reed’s Law” (Reed, 2001), and “Beckstrom’s Law” (Beckstrom, 2008; Wikipedia, 2015), all posit that the utility or value of a network increases faster than linearly with network size, where size is the number of nodes in the network. The sharing, matching, and learning mechanisms theorized as the underpinnings of urban agglomeration economies also represent instances in which larger, and more connected, networks of economic agents generate positive externalities (Duranton and Puga, 2004).

How can the structure of an urban economy be conceptualized and analysed? The degree to which a city can change its economic structure is determined in part by the city's current pool of technologies and skills (Grossman and Helpman, 1993; Lucas, 1993; Morgan and Rosenberg, 1983; Stokey, 1988, 1991). Because labour occupations are defined on the basis of skills and manipulation of technologies (Lobo et al., 2014), the occupations prevalent in a city are a direct indicator of the city's current skills and technological capabilities and are thus almost ideal observational units with which to define the city's economic structure. The structure of an urban economy can be thought of as a set of occupations and the interconnections between those occupations. In other words, the structure of an urban economy is manifested by its network of occupations. Recent work has applied this network perspective to explore how cities transform their economies by altering their occupational network structure (Muneepeerakul et al., 2013; Shalters et al., 2015, 2016). As with any network, a crucial feature of an urban occupational network is the nature and density of the links, or interdependencies, between occupations.

The regularities exhibited by the relationship among the number and diversity of occupations and population size in urban areas has long been studied by urban economics, economic geography and regional science (Berry, 1968; Christaller and Baskin, 1966; Dicken and Lloyd, 1990; Isard, 1956; Krugman, 1996; Lösch, 1954; Mori et al., 2008) and have been recently revisited under a complexity science perspective (Bettencourt et al., 2014; Pumain et al., 2011; Youn et al., 2016). The common empirical thrust of all these investigations is that larger urban areas (with respect to population or workforce size) sustain a larger number of distinct occupations than smaller sized ones. Here we examine how urban scale, measured as the number of distinct occupations, affects the connectivity among these occupations. What we seek to elucidate here is whether, in an information-rich milieu characterized by greater diversity of skills, the likely intensity of interconnectivity among these occupations increases systematically.

Availing ourselves of detailed urban occupational data for six well-established urban systems – those of the United States, Canada, Sweden, Portugal, Australia, and Germany, representing the most advanced production technologies – we construct urban occupational networks. Across those six countries we then compare the density (see definition in Materials & Methods) of a given city's occupational network to its network size, where the network size is equal to

the number of occupations or nodes in the network. These networks are constructed such that each node is a distinct occupation present in a given city and the links between nodes reflect the degree to which two occupations are interdependent (Muneepeerakul et al., 2013). The occupational networks we construct are weighted networks, meaning that we do not merely indicate the existence of a link between occupations, but quantify its intensity (i.e. weight).

The scaling relationship between urban occupations and urban size is of interest not only because of the socioeconomic importance of urban economies. Urban occupational networks are informational networks and their study provides insights into how such networks behave as the number of distinct nodes increases. It has become a common expectation from studying human agglomerations that increasing the scale of such agglomerations (whether they be hunter-gatherer groups, the first market economies, or modern cities) should facilitate a division of labour and the generation of new knowledge through the combination of existing ideas (Glaeser, 2011; Henrich, 2015; Jones and Romer, 2010). Studying urban occupational networks grant us the opportunity to quantify how the connectivity of an information network changes with the scale of the network.

2. MODELING AND ESTIMATION FRAMEWORKS

The general advantages of connectivity in networked informational systems are premised on the division and coordination of labour and knowledge (Bettencourt, 2014). An urban occupational network reflects the division of labor that defines a particular urban economy, and the links connecting occupations reflect specific solutions to the coordination problem inherent in the production of goods and services. Some of the connections among occupations are necessary complementarities without which specialized occupations cannot fulfil specific tasks. But other types of connections among occupations might reflect interactions formed in order to realize new tasks or produce novel goods or services. Economic innovation might result from the opportunities presented by interactions among a growing set of occupations (Douglas, 2015). Other apparent complementarities may result simply from a deeper division of knowledge in larger networks across organizations and places.

How should one expect the density of connections in urban occupational network to scale with the number of occupations? Suppose that urban

occupation networks are indeed not simply about satisfying infrastructural needs or meeting input needs, but also about exchange of information and integration of knowledge. The information flow in such a network can be thought of as an irreversible exchange so that the generation and transportation costs associated with connectivity can in turn be understood as dissipative in nature—as it is transmitted and processed the information gets transformed—and dependent on the intensity of the exchange. Even the most basic of social information exchanges, that between two individuals talking, entails a cost: the generation of information is not energy-free, nor is its transmission even if by direct oral means. Assuming that there is a minimal cost (thermodynamic and pecuniary) which must be borne to generate and transmit information, the various process involved in such generation and transmission may have a general expected cost per connection that is independent of the system size (Bettencourt, 2014). Here “connection” refers to physical infrastructure through which information is transmitted between two agents.

It can also be reasonably posited that the average connectivity cost per node is proportional to the number of connections (representing information infrastructure) and therefore to the size of the network. Considerations of network economics and agglomeration economics lead to the expectation that the productivity of any one activity (or occupation) should be proportional to average socioeconomic connectivity (Hidalgo, 2015). Whether it gets cheaper or more expensive to add a connection partly depends on technological and regulatory considerations. But if the benefits of network connectivity outdistance the cost of establishing connectivity, then increasing network scale should lead to increasing connectivity with the increase being greater than proportional.

We adopt power-law function to represent the relationship between the generalized density of an urban occupational network D and the size of the network measured by the number of distinct occupations N :

$$D_i = \alpha N_i^\beta \quad (1)$$

with α a prefactor capturing the effects of technology and institutional arrangements on the relationship between network size and connectivity, and the subscript i identifying time in a city. Note that D is the generalized definition of network density, which applies to weighted networks as well as unweighted networks, and is defined further in the Methods section. The choice of a power-

law function assumes that the effect on connectivity of increasing network size is not additive but multiplicative which is to say that the increase in connectivity is driven by the interaction of many factors observationally summarized in an increase in network size (Coffey, 1979). The value of β can be estimated by transforming Eq (1) into a liner equation and regressing the natural logarithm of the measure of network connectivity on the logarithm of network size.

The value of the exponent β (an elasticity) determines how the connectivity of an urban occupational network varies with network size. Are urban occupation networks primarily about infrastructural and input complementarities or do they also represent the flow of information leading to new economic niches? If the former is the case, then the value of β should be approximately one, while if the increase in connectivity is driven by the flow of information and the greater scale-dependent opportunities available for creating new economic opportunities then β should be greater than one.

3. RESULTS

Using employment data from six industrialized countries, we first created an occupational network for every metropolitan area in each of those countries. We then measured both the size and the density, or mean link weight, of each occupational network. Our findings reveal, in all countries, a superlinear relationship between the size of a city's occupational network and the density of that network (Fig 1), with the scaling exponent ranging from a high of 2.35 for U.S. cities to a low of 1.17 for Swedish cities (Table 1).

In all cases, the estimated scaling exponents are distinguishable from their trivial values (e.g. $\beta = 1$ or 0), in the absence of agglomeration effects, at 95% level of confidence (Tables 1 and 2). In all cases the superlinear relationship was significant ($p < 0.00001$), with R-square values ranging from 0.80 to 0.93.

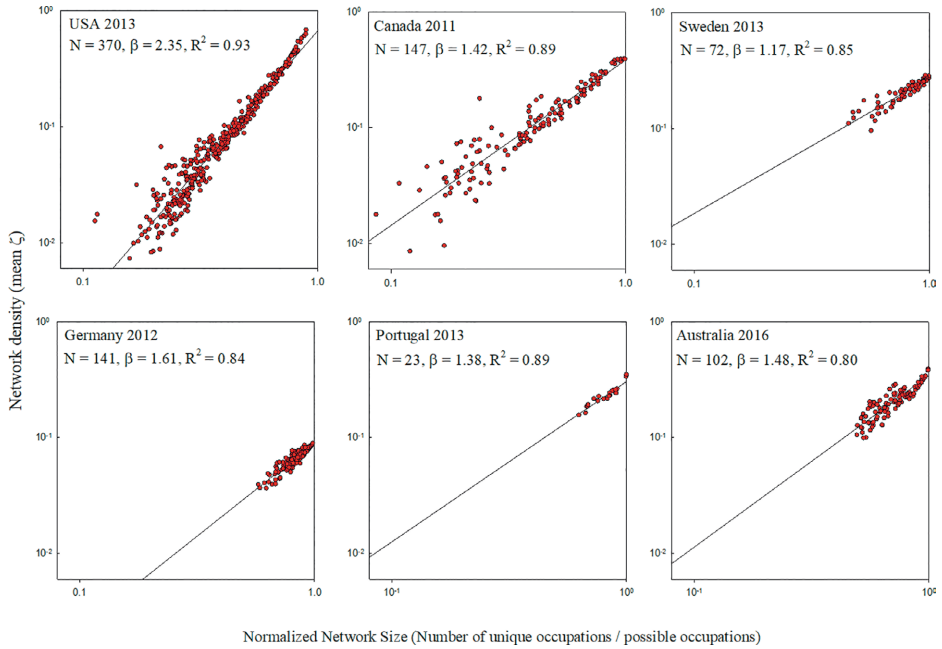


Fig 1. Network density versus network size. Among cities in the six countries studied, the density of a city's occupational network increases superlinearly with the network's size, measured as the number of unique occupations within the city. The exponent of a power law function for each country is given as β . Note that, for comparability, network size has been normalized by maximum possible size.

Table 1. Regression results (model $D = \alpha N^\beta$ where D = network density and N = number of unique occupations).

| Country (Year) | No. Cities | No. Occupations (N) | Exponent (β) | 95% C.I. | R2 | p |
|------------------|------------|-------------------------|----------------------|-------------|------|---|
| USA (2013) | 370 | 812 | 2.35 | ± 0.065 | 0.93 | * |
| Canada (2011) | 147 | 500 | 1.43 | ± 0.083 | 0.89 | * |
| Germany (2012) | 141 | 700 | 1.61 | ± 0.105 | 0.84 | * |
| Australia (2016) | 101 | 475 | 1.48 | ± 0.149 | 0.80 | * |
| Sweden (2013) | 72 | 355 | 1.17 | ± 0.116 | 0.85 | * |
| Portugal (2013) | 23 | 641 | 1.38 | ± 0.223 | 0.89 | * |

*—less than 0.0001

Table 2. Supplemental regression results (model $D = \alpha N^\beta$ where D = network density and N = number of unique occupations). See Figs 3 and 4.

| Country (Year) | No. Cities | No. Occupations (N) | Exponent (β) | 95% C.I. | R2 | p |
|----------------|------------|-------------------------|----------------------|-------------|------|---|
| USA (2013) | 370 | 812 | 2.35 | ± 0.065 | 0.93 | * |
| | 370 | 455 | 2.36 | ± 0.072 | 0.92 | * |
| | 370 | 107 | 2.69 | ± 0.135 | 0.81 | * |
| Germany (2012) | 258 | 700 | 1.60 | ± 0.105 | 0.78 | * |
| | 141 | 700 | 1.61 | ± 0.116 | 0.84 | * |
| | 96 | 700 | 1.54 | ± 0.260 | 0.59 | * |

*—less than 0.0001

These results are based on a p-test level of confidence based on the value of the variance of the coefficient obtained by standard regression. Other methods have been proposed to estimate this variance – and associated level of confidence in measures of spatial sorting – based on null models of urns, for a small number of different types. Note however that a random assignment of types to locations of various sizes is different from this situation where locations may be otherwise similar but display different type compositions. In our case, statistical significance associated with the difference of exponents from proportional scaling provides us with the appropriate test.

Further, we find that the standard deviation of link values also increases with network size (Fig 2), albeit linearly.

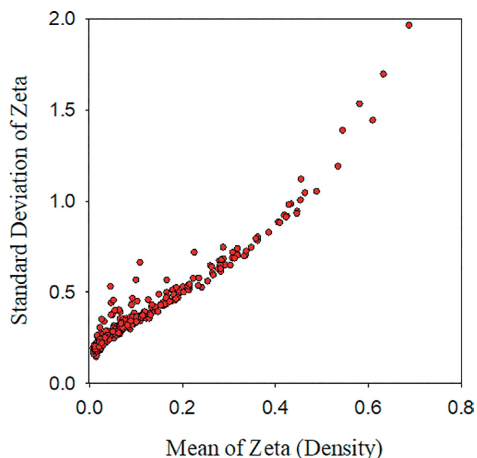


Fig 2. Network density vs. standard deviation of zeta for U.S. metropolitan areas. Increasing density (mean zeta) is correlated with increasing standard deviation of zeta driven by the appearance of rare and highly interdependent pairs of occupations.

Because each country disaggregates total employment into different numbers of occupations, it is important to assess the impact of granularity of occupation classifications on our results. We do this by isolating the effects of different employment aggregation schemes within a single country. Here we use US data, where employment data is aggregated at several different hierarchical levels. While US employment data is typically tallied at the 6-digit occupational code, we aggregated employment additionally at the 5-digit code level and the 4-digit code level. We then recalculated the network size and density for all US cities under these alternative aggregation schemes. Results showed a consistent super-linear relationship regardless of employment aggregation level and showed no qualitative differences in scaling exponent (Fig 3).

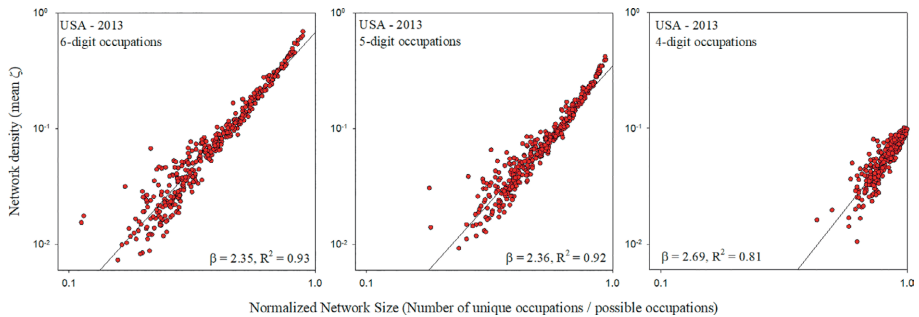


Fig 3. Occupational network size vs. density for U.S. cities at different employment aggregation levels. When occupational network size is compared to its density, the resulting scaling exponents differed little when 2013 U.S. employment is aggregated at the 6-digit, 5-digit, or 4-digit occupation code. The number of distinct occupations in each case are 812 (6-digit), 455 (5-digit), and 107 (4-digit). Note that, for comparability, network size has been normalized by maximum possible size.

Finally, we examined how the relationship between a city's occupational network size and density is affected by the way in which a city is spatially defined. Each country has its own method of defining the spatial boundaries of its metropolitan areas, and so it is prudent to understand how the method of spatial demarcation affects our results. To isolate the effects of spatial delineation methodology, we use employment data from Germany, which has three alternative schemes for spatially defining its metropolitan areas. While our initial analysis used an aggregation method that results in 141 German metropolitan areas, we additionally analyzed employment data using two other aggregation methods that result in 96 urban units and 258 urban units, respectively. Keeping the German occupational classifications constant, we recalculated network size and density using these alternative spatial definitions.

Results revealed a consistent superlinear relationship regardless of how metropolitan areas were defined, while showing no qualitative differences in scaling exponent (Fig 4).

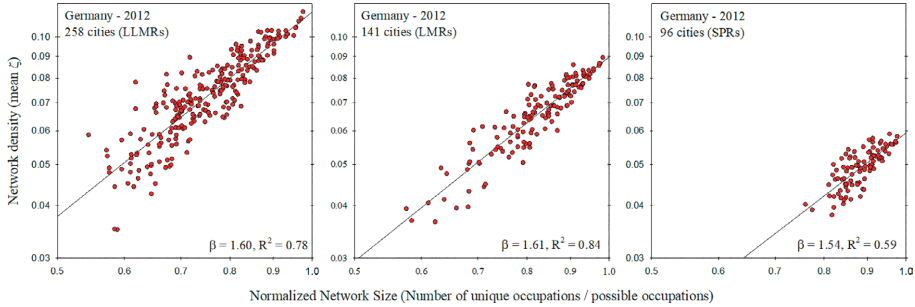


Fig 4. Occupational network size vs. density for German cities at different spatial aggregation levels. When the size of an occupational network is compared to its density, the resulting scaling exponents differed little when German employment is geographically aggregated into 258 LLMRs, 141 LMRs, or 96 SPRs. Note that, for comparability, network size has been normalized by maximum possible size.

4. DISCUSSION

Superlinear scaling of network connectivity with network size

Most straightforwardly, the results presented here indicate that as the number of distinct occupation present in an urban area increase, the linkages among these occupations increase as well (on average). Obviously, some of the connections among occupations result from complementarities as most occupations do not produce end products or services by themselves. But because the increase is greater than proportional, it is evidence that these connections are not infrastructural in nature, as is the case with other types of urban networks, nor is the formation of the linkages animated by economies of scale (Bettencourt, 2013). The superlinear scaling exhibited by urban occupational networks suggests instead increasing returns due to information aggregation (Bettencourt, 2014).

The increase in network density is largely driven by two phenomena. The first is the appearance of rare, but highly interdependent pairs of occupations in larger networks. These highly specialized occupations tend to appear only in larger cities, both in terms of population and network size, that are able accommodate the prerequisite and complimentary occupations needed for

these rare occupations. The second driver of increasing network density is that the proportion of occupations with low or negative link values with other occupations tend to decrease as a city's network size increases. This is likely due to cities becoming more integrated, interdependent, and economically efficient as they increase in size.

Differences among national urban occupational networks

The superlinear scaling of density versus size is exhibited by urban occupational networks corresponding to urban systems of widely differing vintage, history, socioeconomic development and technological capabilities. One can be confident that the superlinearity is therefore not an artefact. Nevertheless, a prominent feature of the result presented in Table 1 is the difference in scaling exponents among national urban systems. One possible reason for the highest scaling exponent occurring in the U.S. is that the U.S. occupational reporting scheme has the highest granularity, recognizing 812 unique occupations, while Germany recognizes 700, Canada 500, Australia 475, Sweden 355, and Portugal 641. However, results of our analysis on multiple aggregation levels of U.S. data (Fig 3) revealed little difference in scaling exponent across various levels of occupational aggregation. This suggests that differences in scaling exponent between countries is likely due to factors other than simply the number of unique occupations that each country recognizes (see Table 2 for full regression details).

Another possible reason for the difference in scaling exponent among countries is that there exists no standard method among countries for defining metropolitan areas. Even within a given country there are typically multiple hierarchical levels of spatial aggregation or alternative schemes of urban delineation. Yet, our results using these three spatial aggregation methods for Germany, presented in Fig 4, reveal little difference in scaling exponents. This suggests that differences in scaling exponent between countries may not be due simply to differences in how countries spatially define their metropolitan areas (see Table 2 for full regression details).

Yet another possibility for differences among countries' scaling exponents is number and distribution of cities by size, where size is measured as the number of total employees. While the U.S. has 23 cities with at least 1 million employees, the smallest country in our study, Sweden, has only one. In addition, neither Canada nor Sweden have cities that compare in magnitude to the largest

U.S. cities. While Sweden's largest labour market has 1.3 million employees (Stockholm), Germany's largest has 2.5 million (Berlin), and Canada's largest has 2.8 million (Toronto), the largest U.S. labour market has 8.3 million employees (New York) and three other U.S. labour markets have 3 million or more. Thus, the scaling exponent of the U.S. may be influenced by the incomparable size and quantity of its largest cities.

5. MATERIALS AND METHODS

5.1. Data

Our analysis is based on employment statistics that aggregate the number of workers in each occupation in each metropolitan area of an entire country. In the current analysis, we use such datasets for six countries—the U.S., Canada, Germany, Sweden, Australia, and Portugal. While there are nuances to the way that each country defines its occupations, in general occupations are defined based partly what work is performed and partly on the skills and training needed to perform the work (Florida, 2012; Florida et al., 2008; Moretti, 2012).

The spatial units of analysis for the U.S. urban system are its Metropolitan Statistical Areas (MSAs) for which employment data is compiled by the U.S. Bureau of Labor Statistics (BLS). MSAs consist of a core county, or counties, containing a city with a population of at least 50,000, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties. MSAs are unified labour markets and encompass geographical areas of high economic cohesion (Glaeser et al., 1995; US Census Bureau, 2015). The BLS included 380 MSAs in its 2013 employment data. Of those eight were excluded because they are in Puerto Rico, which has a unique economic environment, and two were excluded because, although they are treated as metropolitan areas by the BLS, they are classified as micropolitan areas by the U.S. Census Bureau. Together the remaining 370 U.S. metropolitan areas account for nearly 85% of U.S. population and over 90% of U.S. economic output. We use data from the BLS's 2013 Occupational Employment Survey, which includes the estimated number of people employed in each of 812 distinct occupations for each MSA (US Bureau of Labor Statistics, 2014).

For the analysis of Canada's urban system, the spatial units of analysis are the 32 Census Metropolitan Areas, which have a population greater than 100,000, and the 115 Census Agglomerations, which have a population between

10,000 and 100,000 (Statistics Canada, 2015). These 147 units are defined by a high level of economic cohesion and are typically geographically contiguous. Employment data was extracted from Canada's 2011 National Household Survey (NHS), which was compiled by Statistics Canada in conjunction with Canada's quinquennial census. For each urban area, the survey collects the number of workers employed in each of 500 occupations.

For Germany the units of analysis are 141 Labour Market Regions (LMRs) as defined by Kosfeld and Werner (Kosfeld & Werner, 2012). LMRs consist of one or more of the 402 German districts. LMRs are characterized as essentially independent economic areas with close commuter links within the regional labour markets. To analyse the effects of alternate spatial definitions of metropolitan areas we apply two additional functional demarcations of regions available in Germany. First, we use 258 Local Labour Market Regions (LLMRs) delineated by the German Federal Government for use by business development programs. Second, we make use of the 96 Spatial Planning Regions (SPR) used by the Federal Government for regional planning. LLMRs and SPRs are also based on accessibility and interdependence criteria such as the catchment areas and commuting flows but capture different geospatial definitions. Employment data is provided by the German Federal Employment Agency, which collects information on all employees subject to German social insurance contributions (including health, pension, long-term care, and unemployment insurance funds) as reported by employers. German employment data come from year 2012 and includes the number of workers in each of 700 occupations.

For Sweden the units of analysis are the 72 Functional Analysis Regions (FAs). Swedish FAs are delineated by the Swedish Agency for Growth Policy Analysis (Tillväxtanalys) and consist of one or more of the country's 290 municipalities exhibiting a high level of commuting workers across municipal boundaries. Like U.S. MSAs, this transboundary movement of labour is taken to indicate a high level of economic cohesion (Roto and Claessen, 2012). Many Swedish FAs consist of only one or two municipalities, especially in the northern part of the country where population density is low, while the Stockholm FA – the largest – is an agglomeration of 28 municipalities.

Employment data were extracted from Statistics Sweden's 2012 microdata, which covers all individuals in the country's Labour Force Survey and tabulates

the number of employees in each of 355 occupations across the country's municipalities.

For Portugal we use 23 metropolitan areas defined by Nomenclature of Territorial Units for Statistics level 3 (NUTS3). These areas are contiguous and stable regions that reflect urban socio-economic and administrative geographical unity. We exclude from the analysis two Atlantic archipelagos, the Autonomous Regions of Açores and Madeira, due to their unique social-economic contexts. NUTS3 correspond to *Áreas Metropolitanas e Comunidades Intermunicipais*, which are formed for regional planning purposes by the Portuguese Local Public Administration, aggregating the 308 Portuguese municipalities according to their geospatial interdependencies (labour mobility, urban economy and public administration links, etc). We use metropolitan level employment data for 641 occupations provided by the 2012 *Quadros de Pessoal*, the Portuguese Linked-Employer-Employee-Data. *Quadros de Pessoal* is collected annually by the Portuguese Ministry of Employment and covers all establishments' employees subject to Portuguese social insurance contributions.

For Australia we use 101 Significant Urban Areas (SUA) as defined by the Australian Bureau of Statistics' Australian Statistical Geography Standard. These units represent towns and cities with 10,000 or more residents in either a single or a cluster of urban centres. The 2016 Australian census reports employment for SUAs in each of 475 occupations using the most recent Australian and New Zealand Standard Classification of Occupations (ANZSCO) at the 4-digit occupation codes level.

5.2. Occupational interdependence

Two aggregate metrics are calculated for each city's occupational interdependence network, the network's size and its average connectivity (or density). In these networks, nodes are occupations and the weight of a link between any two nodes is a function of how often those two occupations are co-located in the same city. All urban economies within a national system have certain economic activities in common, namely those that address the needs of individuals and households and that satisfy demands common across a society. But what distinguishes urban economies from each other are those activities in which each city specializes. In calculating the occupational interdependencies that define an urban economic structure we focus on those occupations in which

a city specializes, that is, those occupations that define a city's comparative human capital advantage.

A city is specialized in an occupation if the proportion of the city's labour force engaged in that occupation is greater than the same proportion nationally. Thus, specialization can be stated in terms of the widely-used location quotient (LQ), which for occupation i in MSA m is defined as:

$$LQ_{m,i} = \frac{\left(\frac{x_{m,i}}{\sum_i x_{m,i}} \right)}{\frac{\sum_m x_{m,i}}{\sum_m \sum_i x_{m,i}}} \quad (2)$$

where $x_{i,m}$ is the number of workers employed in occupation i in city m . City m is specialized in occupation i if its location quotient $LQ_{m,i} > 1$. Thus, for each country in our study we derive an $M \times O$ specialization matrix where O is the number of recognized occupations in a country, M is the number of metropolitan areas in a country (for which employment data exists), and element $a_{m,i} = 1$ if $LQ_{m,i} > 1$ and 0 otherwise.

How can one infer from the presence of specialized occupations in an MSA that their co- location is not merely accidental but indicative of interactions through which complementarities are realized and information flows? We employ conditional probability: specifically, in this context if the presence of one specialized occupation in an MSA is statistically affected by the presence of another specialized occupation, one would expect conditional probabilities to differ from marginal ones. The co-location patterns of specialized occupations among all cities to define the interdependence between any two occupations i and j , $\zeta_{i,j}$, as:

$$\zeta_{i,j} = \frac{P(LQ_{m,i} > 1, LQ_{m,j} > 1)}{P(LQ_{m',i} > 1)P(LQ_{m'',j} > 1)} - 1 \quad (3)$$

where m , m' , and m'' denote a randomly selected city [21]. This metric measures how an MSA's specialization in one occupation may enhance or hinder its specialization in another. The emphasis on "may" acknowledges that—as is the case for many statistical analyses—without additional information or experiments, our analysis cannot imply direct causality; at best, it identifies structural relationships and points to potential places where one may search

for such causality. Thus, z has the characteristic of being positive when two occupations co-occur in a city more frequently than expected by chance, and of being negative when they co-occur less frequently. Note that, because our networks are undirected, the interdependence is symmetric so that $\zeta_{ij} = \zeta_{ji}$. The calculation of ζ is specific to each nation's occupational classification system but applies to every city within that nation.

We treat the occupational interdependencies to be the weights that link every pair of occupations in an occupational interdependence network. Network weights are a quantification of the intensity of the link between any two nodes in a network. Weights could indicate the magnitude of flows between nodes, the frequency of interaction, the strength of a relationship, etc. In our occupational networks, weights quantify the intensity of co-occurrence for any two occupations.

Because of heterogeneity among occupational relationships, occupations are not uniformly distributed within a country's occupational interdependence network. Instead, a country's network typically contains a denser core of highly interdependent occupations and a periphery of occupations that tend to be weakly or negatively interdependent with others. Two occupations i and j tend to be, on average, close to each other in a country's occupation network if ζ_{ij} is positive and farther apart if ζ_{ij} is negative. Thus, occupations that appear closer in a network also tend to co-exist within a given city more frequently.

Having specified an occupation network specific to each country, we can then locate specific cities within a country's network space. To understand what is meant by a city's location in a network, it is important to note that no city in our study has employees in every possible occupation. Instead, each city contains a subset of all possible occupations and when this subset is mapped to nodes within the full occupational network, it defines a subnetwork representative of the city, which we equate here with the city's location within the full network.

Note that for Germany, we used employment data at the aggregation level of the 141 LMRs to calculate the country's zeta values. Furthermore, while 10 MSAs are excluded from the US comparative analysis (8 because they were in Puerto Rico, 2 because they are generally recognized as micropolitan statistical areas), employment for those 10 were included in the basis for calculation

US zeta values to ensure that interdependencies were based on the most comprehensive data possible.

5.3. Density and size of urban occupational networks

Given these weighted occupational networks for metropolitan areas in our countries of interest, we seek to determine the relationship between each network's average connectivity and its size. Here we equate a network's average connectivity with the network's density. The traditional metric of network density, which applies to unweighted networks, is simply the number of links in a network divided by the number of possible links (Wasserman & Faust, 1994). Because our networks are weighted, we utilize a subsequently developed generalized definition of network density, which is the sum of all weights divided by the number of possible links (Guimei Liu et al., 2009; Tokuyama, 2008). Thus, the density of a weighted network is synonymous with the average of all weights in that network.

An unweighted and undirected network of N nodes can be represented as a symmetric $N \times N$ matrix in which element $a_{ij} = a_{ji} = 1$ if and only if a link exists between nodes i and j . Other-wise $a_{ij} = a_{ji} = 0$. Typically, an arbitrary threshold of interaction strength between two nodes is used to determine the existence of a link (e.g. $a_{ij} = 1$), and once determined, all links are thus equivalent. Such networks are amenable to wider range of analytical tools and are typically easier to analyse. However, in collapsing interaction strength to a binary determination, important information regarding the network and the system it governs is lost (Fagiolo et al., 2010). Thus, weighted networks exist as an alternative representation of interaction systems in which the elements of the network's adjacency matrix may be other than 0 or 1. Instead, each element holds a weight w , or a value representative of the strength of interaction. In an undirected network weights are symmetric so that $a_{ij} = a_{ji} = w_{ij}$.

For a given city m , the generalized density D_m of its occupational network can be calculated as:

$$D_m = \frac{2}{N_m(N_m - 1)} \sum_{i>j}^{N_m} \zeta_{i,j} \quad (4)$$

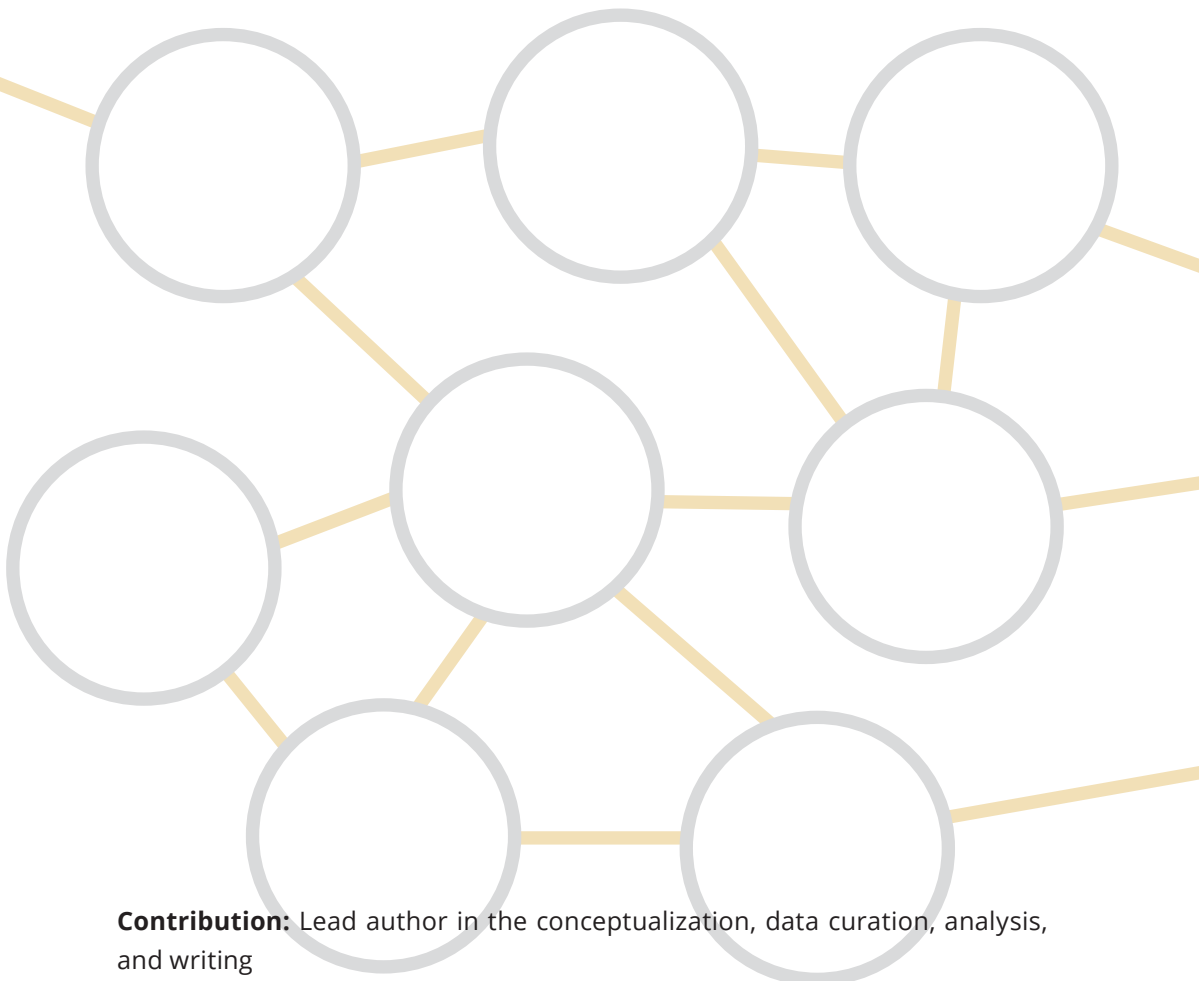
where N_m = the number of nodes (e.g., unique occupations) in city m and ζ_{ij} is the interdependence (e.g., weight) between occupations i and j , both of which must

be present in m . Because our occupational networks have an interdependence value for every pair of occupations, they are complete networks (e.g. every node is linked to every other node) and their density is therefore the mean z across all links in a given city. Note that when equation 4 is applied to an unweighted network, so that $\zeta_{i,j} = 1$ when a link exists between nodes i and j and $\zeta_{i,j} = 0$ when there is no link, $D \in [0, 1]$ and is simply the number of links present divided by the number of links possible.

A noteworthy aspect of our networks is the existence of negative weights, which is being increasingly addressed in analyses of networks (Everett & Borgatti, 2014; Labianca, 2014; Szell & Thurner, 2010). By permitting negative weights, we incorporate into our analysis those instances where two occupations interact negatively. That is, only one of the pair tends to exist in a city, suggesting a form of competitive exclusion or similar interference.

CHAPTER 3

What drives the geography of jobs in the US: Unpacking relatedness



Contribution: Lead author in the conceptualization, data curation, analysis, and writing

Coauthors: Pierre-Alexandre Balland, Andrea Morrison, and Ron Boschma

Published as: Teresa Farinha, Pierre-Alexandre Balland, Andrea Morrison & Ron Boschma (2019) What drives the geography of jobs in the US? Unpacking relatedness, Industry and Innovation, DOI: 10.1080/13662716.2019.1591940

1. INTRODUCTION

The 2008 crisis has led to unprecedented job losses and the destruction of human capital in many regions worldwide. On top of that, technological change, automation, and offshoring of jobs are leaving their marks (Autor, 2010; Mehta, 2014; Moretti, 2012; Rodriguez & A. Jayadev, 2010). Though such shocks and global trends affect all regional economies, they tend to do so in varying degrees (Shutters et al., 2015). This has initiated a recent interest of scholars to study systematically the evolution of occupational structures in regions over time.

Muneepeerakul et al. (2013) was the first study assessing how relatedness affects entry and exit of occupations in US metropolitan regions (see also Brachert, 2016; Shutters et al., 2016). These studies follow a recent body of literature on regional diversification that shows that regions tend to diversify into new industries (Boschma et al., 2013; Essletzbichler, 2015; He et al., 2015; Neffke et al., 2011) or new technologies (Kogler et al., 2013; Rigby, 2015) that are closely related to their pre-existing capabilities. What these studies on regional diversification have not unraveled so far are the mechanisms through which industries, technologies or occupations may be related. In fact, there is still little understanding of the sources of relatedness that impact on regional diversification (Boschma, 2017; Tanner, 2014).

The main objective of this paper is to unpack the mechanisms through which the entry and exit of jobs in cities take place. While previous papers looked at the effect of geographical density only, we argue that co-location of jobs tells little about the forces that make jobs co-occur in the same city: new local jobs may be related to local jobs because they share similar skills, provide complementary tasks, or both, or because they benefit from each other's co-location. We make a distinction between three mechanisms: (1) occupations can be related because they incorporate a similar set of skills of high relevance for each job; (2) occupations may be complementary in the process of producing a good or service; and (3) occupations may jointly benefit from synergies in cities. There is no study yet that has investigated the importance of each of these three mechanisms in the evolution of the geography of jobs.

We use a network approach to unpack the relatedness concept into three dimensions and develop a measure for each of them. We test the impact of each relatedness dimension on the dynamics of the occupational structure of 389

Metropolitan Statistical Areas in the US from 2005 to 2016, more specifically, on the probability of occupations entering and exiting the employment structure of cities. Our paper confirms the results found in other studies that cities enter new jobs related to ones already existing in that cities, and exit jobs unrelated to their jobs portfolio. Moreover, we found that all three relatedness dimensions have a significant effect but they seem to prevent exit of jobs in cities more than promoting entry of jobs in cities. Local synergy density shows the largest effect on both entry and exit of jobs in cities.

The structure of the paper is as follows. In Section 2, we present the concept of relatedness as developed in Evolutionary Economic Geography, and we explain how we unpack relatedness into three dimensions. Section 3 presents the data, our measures for each relatedness dimension, and the network representation of the occupational structure. Section 4 presents the study on how job relatedness, in its different dimensions, has influenced the entry and exit of occupational specializations in US cities. Section 5 discusses the results and concludes.

2. REGIONAL DIVERSIFICATION IN JOBS: THREE MECHANISMS

In Evolutionary Economic Geography, history is key to understand the economic evolution of regions (Boschma & Frenken, 2006; Martin & Sunley, 2006). Past structures set opportunities but also boundaries to future development. A large body of empirical studies shows that diversification occurs in regions mainly by making use of and recombining pre-existing regional capabilities: in other words, it is subject to path-dependency (Boschma, 2017). Moreover, regions localized in the dense parts of the 'product space' (i.e. having many products related to each other) have also more diversification options and higher economic growth rates (Frenken et al., 2007; Hidalgo et al., 2007; Hidalgo & Hausmann, 2009).

These studies tend to look at diversification in terms of new products (Hidalgo et al., 2007), new industries (Neffke et al., 2011) or new technologies (e.g. Balland et al., 2019; Kogler et al., 2013; Petralia et al., 2017; Rigby, 2015). However, industry, product, and technology classifications capture some but not all capabilities in regions (Kline & Moretti, 2013; Markusen, 2004). This point was made by Thompson & Thompson (1985, 1987) who made a strong claim in favour of an occupational functional approach to understand the changing spatial division of labour in which advanced regions focus on high value-added activities and

jobs (design, marketing, R&D) while off-shoring labour-intensive (and low-skilled) jobs to places where labour costs are comparatively low (Barbour & Markusen, 2007; Gereffi & Korzeniewicz, 1994; Markusen, 2001; Markusen & Schrock, 2006; Renski et al., 2007). This growing separation of functions within the same industry implies that regions with similar industrial specialisations can reflect very different underlying capabilities in terms of knowledge and skills (Markusen et al., 2008). Technology classifications do not cover all capabilities in regions either because they tend to capture scientific and technical skills. Shifting away from industries and technologies to jobs reveal what regions *do* with their skills, as opposed to what regions *make* as the outcome of their activity (Feser, 2003; Thompson, 1985). This change of perspective is important as growth opportunities in knowledge-based economies are considered to depend on the accumulation of human rather than physical capital (Moretti, 2012). And last but not least, an occupational approach can cover better service industries than the industry/technology approach.

Muneepeerakul et al. (2013), Brachert (2016), and Shutters et al. (2016) were the first to acknowledge the relevance of the occupational structure to analyse regional evolution. These studies provide a network representation of the structure of interdependent job classes in US cities, called *occupational space*. They show that co-located occupational specializations can interact positively or negatively with each other; and that the balance between these interactions determines productivity, wealth, and possible development paths of urban economies. Hasan et al., (2015) found that interdependencies between jobs (either as task overlap or task coordination) tend to protect jobs. On the other hand, regarding the whole job structure, they found that interdependence (ties between jobs) makes a job vulnerable to the exit of other jobs in that job's cluster, decreasing its survival chance.

However, these studies on job diversification in cities have not looked at the types of mechanisms through which related diversification unfolds. This means we have to unpack the broad notion of relatedness, as advocated by some scholars (Boschma, 2017; Breschi et al., 2003; Tanner, 2014). Inspired by Duranton & Puga (2003), we distinguish between three mechanisms or channels through which agglomeration externalities may be exploited, and we make an explicit connection to job dynamics.

The first mechanism refers to *similarity* of skills between jobs. This applies when a certain set of skills can be used to perform more than one type of task or

job activity: job classes that have those skills are similar (but not identical) and substitutable to a considerable degree. This has a close resemblance with the notion of skill-relatedness introduced by Neffke & Henning (2013). The second mechanism refers to *complementarity* of skills between jobs. Here, skills in different job classes are required to produce a certain good or service within a value chain, like a doctor and a nurse in a hospital provide complementary skills to cure illnesses. In modern societies, as products/services complexity increases, the amount of interdependent tasks increases within each value chain. We will capture this skill complementarity by looking at the co-occurrence of job classes in economic activities. The third mechanism is associated with *local synergy* effects between different jobs when the co-location of two jobs (e.g. a businessperson and a taxi driver) benefit each other, while not being similar or complementary in skills with one another. These local synergies may arise due to common natural endowments, demand-driven interdependencies of jobs, or amenities (Florida, 2012; Moretti, 2012). This latter dimension also covers local multipliers in which high-skilled jobs provide benefits for low-skilled jobs (Kline & Moretti, 2013; Moretti, 2013). We will capture local synergy by identifying the geographical co-occurrence of job classes, after having it filtered from the other two dimensions.

There is no study yet that has investigated the importance of each of these three mechanisms in the evolution of the geography of jobs. We examine which of the mechanisms can explain best the entry of new jobs and the exit of existing jobs in 389 US cities from 2005 to 2016.

3. OCCUPATIONAL DATA AND NETWORK ANALYSIS

3.1. Occupational data

The main source is employment data provided by the Bureau of Labor Statistics of the U.S. Department of Labor (BLS)⁵. It contains several workers statistics, such as total employment and mean hourly wage by job class (approximately 800 categories at the 6-digit level) by industry (NAICS) and by U.S. Metropolitan Statistical Areas (MSA). The Standard Occupational Classification (SOC) System groups similar jobs into job classes (OCC) based on the work performed, skills, education, training, and credentials required to carry out specific work tasks. Some OCC are found in just one or two industries, others in a large number of industries. NAICS is a production-oriented classification that groups

⁵ publicly available at <http://www.bls.gov/oes>

establishments into industries based on their prime activity. MSAs represent unified labour markets (Muneepeerakul et al., 2013). Each MSA contains a core urban area of at least 50,000 population in one or more core counties, including adjacent counties with a high degree of social and economic integration with the urban core. MSAs account for nearly 85% of U.S. population and 90% of U.S. economic output (US Census Bureau, 2015).

To account for classification schemes revisions and assure a comparable multi-year analysis, we use data from 2005 to 2016 and exclude from our analysis the MSAs (eight MSA and five NECTA) and the OCCs that came into existence after 2005, and the “All Other” type of OCC which is not available in the O*NET data⁶. After cleaning data, we end up with statistics on number of people employed in each year-OCC-MSA (12 years, 733 OCC, and 389 MSA).

After that, we cross the BLS employment data with occupational content classification from the Occupational Information Network (O*NET). O*NET provides a detailed classification of occupational contents – occupational requirements and worker attributes for each job class⁷. O*NET attributes to each job class the correspondent workers’ capabilities, according to the O*NET classification schemes. After testing their typology and employment data distributions, we chose the Intermediate Work Activities (IWA)⁸ classification scheme that represents all skill specifications needed to perform each job class⁹, and is, therefore, better suited to compute our measure of *job similarity*. The result is a dataset with job requirement weight for each OCC-Skills (same 733 OCC, 332 Work Activities).

Because many unified product value chains bring together different NAICS classifications, we cross the BLS employment data with an industry classification defined and made available by BLS, the Industry Sectoring Plan¹⁰. This industry

6 “All Other” titles represent job classes with a wide range of characteristics, which do not fit into one of the detailed O*NET-SOC occupations.

7 publicly available at <https://www.onetonline.org>

8 O*NET provides classification schemes for Work Activities at three levels of aggregation (41 Generalized Work Activities; 332 Intermediate Work Activities; and finally, 2070 Detailed Work Activities). Intermediate Work Activities is the level of aggregation that provides us better network analysis conditions (enough categories, and that are not too common and not too rare across job classes)

9 Here we refer to skills in its broad sense, equivalent to the concept of regional capabilities, commonly used in the evolutionary economic geography literature. It corresponds not to O*NET classification schemes for skills (which refers to a much stricter sense of skills), but to O*NET definition of workers’ competencies (it includes classification schemes for skills in the stricter sense, knowledge, abilities, experience and training, etc.).

10 BLS aggregates NAICS (4 digit level) into the industry sectors, further used in BLS’s employment projections (https://www.bls.gov/emp/ep_data_input_output_matrix.htm)

classification groups together the narrowly defined U.S. industry codes (NAICS) that are related in terms of inter-industry linkages (input-output measures) into industry sectors, or more simply referred as clusters. In other words, we aggregate the BLS employment-OCC-NAICS data into an employment-OCC-cluster dataset for the last year of the period under consideration (same 733 OCC, 179 industry clusters, for the year 2016)¹¹. The result is an industry cluster's labour demand dataset, from which we compute our *job complementarity* measure.

After cleaning and merging data, we compute the geographical measure of relatedness (co-location-based measure), and our measures for complementarity and similarity dimensions of relatedness. We obtain a bipartite dataframe with three variables of relatedness for each possible pair of job classes in each year. We use this data in the network analysis and further transform it into a new dataset to be used in the regression analysis. For ease of interpretation, we will use the terms "job class", "city", "industry", and "skills" when referring to OCC, MSAs, Industry Sectoring Plan categories, and Work Activities, respectively.

3.2. Unpacking relatedness

In line with the network-based framework of Hidalgo et al. (2007) and Muneeppeerakul et al. (2013), we build a network of job classes and relatedness between them – the Job Space – to represent the U.S. labour market structure. The Job Space will have three types of links based on three measures of relatedness: a geographical, a complementarity and a similarity measure. From those three measures of relatedness, we will deduce the fourth one for the local synergies dimension of relatedness – the pairs of job classes that are poorly complementary, poorly similar, but most frequently co-located, due to local synergies.

Geographical relatedness of jobs

First, we identify job classes in which U.S. cities specialize in. We use the location quotient (LQ) of job class j in city c , based on the number of employees (x)

¹¹ Due to classifications correspondence constrains, we exclude the "Private households" sector (not available in BLS employment data) and further pull together a few industry sectors, ending up with 179 sectors instead of 186. More specifically, we aggregate into one the "Crop production", "Animal production and aquaculture", "Forestry" (including "Support Activities for Forestry"), and "Fishing, hunting and trapping". We also aggregate into one the governmental sectors (which corresponds to the 2digits NAICS 92 – Public Administration).

engaged in job class j , in city c , in relation with the total number of employees engaged in job class j in the country:

$$LQ_{c,j} = \frac{\left(\frac{x_{c,j}}{\sum_j x_{c,j}} \right)}{\frac{\sum_c x_{c,j}}{\sum_c \sum_j x_{c,j}}}$$

A LQ higher than one means that the proportion of the labour force engaged in that job class is “overrepresented” in that city. As a result, we get a binary jobs-cities matrix ($N \times M$ matrix). Then, we compute the geographical measure of relatedness between each pair of job classes, based on their co-occurrences as specializations in cities, for each year during the 2005-2013 period. More concretely, we use a conditional-probability-based measure developed by Van Eck & Waltman (2009) and reformulated by (Steijn, 2018). This results in a symmetric $N \times N$ job classes matrix, in which each cell (i, j) contains the geographical measure of relatedness (*GeoRel*) between job class i and job class j , i.e., the probability of a city c being specialized in job class i given that it is also specialized in job class j , as follows:

$$GeoRel(C_{ij}, S_i, S_j, T) = C_{ij} / (m * ((S_i/T) * S_j / (T - S_i) + (S_j/T) * (S_i / (T - S_j)))$$

where C_{ij} , S_i , and S_j are, respectively, the number of co-occurrences of i and j , the number of occurrences of job class i and the number of occurrences of job class j , as occupational specializations in cities. T is the sum of all cities occupational specializations, and m is the total number of co-occurrences. The geographical measure of relatedness indicates the probability of two job classes being together in the same city. *GeoRel* is lower bounded by zero (job classes i and j are never together as specializations in same city) but not upper bounded. A *GeoRel* higher than 1 means that two job classes co-locate in the same city more often than by chance.

Although commonly used as an outcome-based measure of relatedness, co-location of job classes does not inform us about the type(s) of relatedness between two jobs. In order to empirically unpack the dimensions of relatedness for each pair of job classes, we create other two measures of relatedness: jobs similarity and jobs complementarity.

Jobs similarity

Based on BLS job classes and O*NET's Work Activities classification scheme, we compute jobs similarity as the frequency of co-occurrences of jobs classes in work activities classes. More specifically, in line with Hasan et al (2015), we first construct a $1 \times W$ vector for each job class, with W being the number of O*NET IWA categories, and join them to form a binary jobs-IWA matrix ($N \times W$ matrix). Then, we apply conditional probabilities for computing jobs similarity measure of relatedness (equivalent to the *GeoRel* equation, the jobs co-location measure, but based on the jobs-IWA matrix instead). In result, we get a symmetric $N \times N$ job classes matrix in which each cell (i, j) contains the skills similarity between job classes i and j . In other words, skills similarity represents, therefore, job classes' co-occurrences in IWA as the main occupational destination of such skills (e.g., Work Activity w is a highly required skill, more than average in regional labour markets, for both job class a and b).

Jobs complementarity

Based on industry clusters' labour demand, we compute complementarity by looking at which pairs of job classes are jointly required in the same value chain(s). We determine how often two job classes co-occur in the same industry cluster. We first compute each industry cluster's LQ in each job class, i.e., each cluster employment shares in each job class, compared to the average employment shares of all clusters (same LQ equation we used for jobs co-location measure, but based on the jobs-cluster matrix). Then, we apply conditional probabilities for measuring jobs complementarity (equivalent to *GeoRel* equation but based on the jobs-cluster matrix). So, we construct a symmetric $N \times N$ job classes matrix in which each cell (i, j) contains the jobs complementarity index between job classes i and j .

Jobs local synergies

From the three measures of geographical relatedness, complementarity and similarity, we can derive the local synergies dimension of relatedness. Pure geographical relatedness confounds the different forces that make jobs co-occur in the same city. Indeed, jobs may co-locate for reasons of complementarity or similarity, so we cannot tell for sure if local synergies do operate or not. However, local synergies are notoriously difficult to identify. They refer to strong agglomerative forces, but not of the complementarity and the similarity kind. Because some pairs of complementary and/or similar job classes may also have a tendency to co-locate, we need to control for that. We argue that if

two job classes have high geographical relatedness but low skills similarity and low industry complementarity, we assume these two job classes show local synergies. So, we deduce the presence of local synergies by identifying pairs of job classes that are most probable to co-locate in cities but do neither show a high degree of jobs complementary nor high jobs similarity.

Figure 1 presents the top 50 pairs of related job classes, that is, the 50 highest links of relatedness between occupational specializations in US cities. Some pairs of jobs, such as “roof bolters mining” and “mine cutting and channeling machine operators”, show to be highly related simultaneously due to similarity, complementarity, and co-location. Other pairs of jobs are highly related mainly due to sharing similar skills, as is the case of “lawyers” and “paralegals and legal assistants”, while pairs of jobs like “political scientists” and “industrial-organizational psychologists” are related due to local synergies.

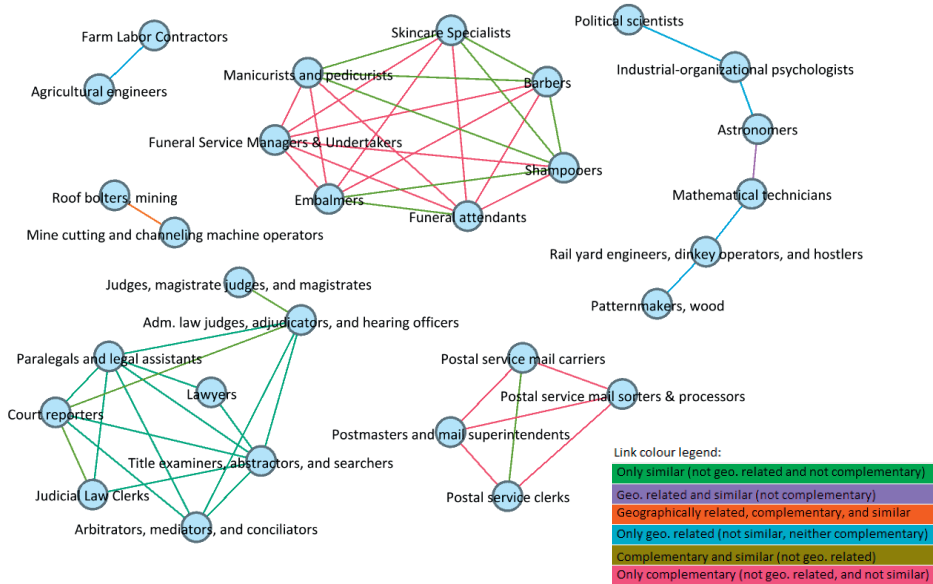


Figure 1. Top 50 Pairs of Related Job Classes

3.3. The Job Space - a descriptive analysis

We use the three relatedness measures of geographical relatedness, complementarity, and similarity across jobs to build the Job Space. It is a network-based representation of the regional occupational structure in which each node stands for a job class, and links between nodes represent jobs' relatedness. Because the three relatedness measures may overlap for each pair

of job classes (e.g. pairs of job classes may be simultaneously complementary and similar), we graphically illustrate the Job Space in three distinct layers. Figure 2 shows the Job Space in 2016. A first layer shows job classes (nodes) and geographical relatedness between them (links). In the second layer, we keep the position of the nodes from the first layer (for comparability purposes) and show jobs similarity links instead. We repeat this procedure for the third layer showing jobs complementarity. We use the *Minimum spanning tree* network representation algorithm in order to offer a visualization in which all job classes are included and the network is connected with the minimum links possible, i.e., $N-1$ links.

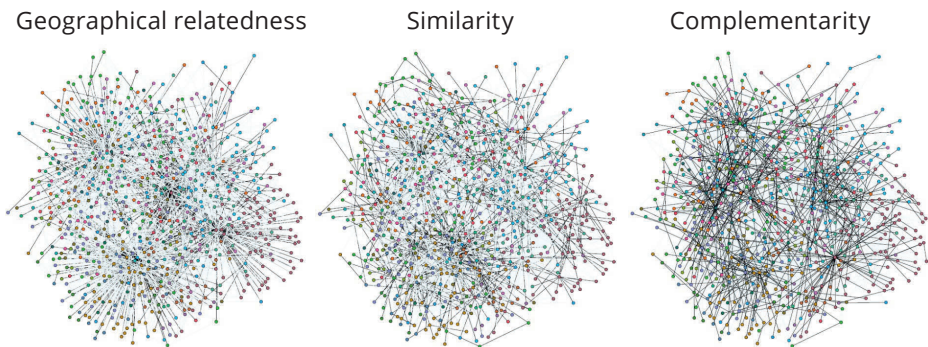


Figure 2. The Job Space in Three Layers. Geographical relatedness Similarity Complementarity

4. RELATEDNESS DIMENSIONS AND THE RENEWAL OF THE JOB-SPACE

Once the job-space is built, we use econometric tools in order to analyse how jobs relatedness affects the renewal of the employment structure of US cities and, in particular, how different dimensions of jobs relatedness (similarity, complementarity, or local synergies) may differently affect that evolution. Starting from 2005, we track yearly changes in the employment structure of each city until 2016 and applying linear probability models to estimate how jobs relatedness affects the entry and exit of job classes in US cities.

4.1. Variables and descriptives

We first construct two dummy variables, *Entry* and *Exit*. *Entry* is conventionally computed as equal to one if a job class did not belong to the occupational specialization portfolio of city c in time $t-1$, and enters in time t . And *Exit* is equal

to one if a job class did belong to the occupational specialization portfolio of city c in time $t-1$, but exits in time t :

$$\begin{aligned} Entry_{c,i,t} &= 1, \text{ if } LQ_{c,i,t} > 1 \text{ and } LQ_{c,i,t-1} \leq 1 \\ Exit_{c,i,t} &= 1, \text{ if } LQ_{c,i,t} \leq 1 \text{ and } LQ_{c,i,t-1} > 1 \end{aligned}$$

LQ ranks cities level of specialization in relation to the average level of specialization of all regions in a year. This means that the position in the ranking of a city may vary from one year to another, not due to changes in that city's level of specialization but to changes in other cities' level of specialization that affect the average level of specialization of an economy. So, a job class could change from being a city specialization $t-1$ but not anymore in t , just because the ranking of specialization of that job class increased overall in the average economy, not because the share of employees in that city decreased. To exclude such "false" changes in computing Entry and Exit, we made a slight adjustment to the LQ in t ¹². We track the evolution of an occupational specialization in the city in relation to the pre-existing structure of the city, from $t-1$ to t , independent of the evolution of the economy's average specialization level, which we fix at $t-1$ when computing LQ in t , as follows:

$$\begin{aligned} Entry_{c,i,t} &= 1, \text{ if } LQ_{c,i,t,t-1} > 1 \text{ and } LQ_{c,i,t-1,t-1} \leq 1 \\ Exit_{c,i,t} &= 1, \text{ if } LQ_{c,i,t,t-1} \leq 1 \text{ and } LQ_{c,i,t-1,t-1} > 1 \end{aligned}$$

which translates into:

$$\begin{aligned} Entry_{c,i,t} &= 1, \text{ if } \frac{\left(\frac{x_{c,i}}{\sum_i x_{c,i}}\right)_t}{\left(\frac{\sum_c x_{c,i}}{\sum_c \sum_i x_{c,i}}\right)_{t-1}} > 1 \text{ and } \frac{\left(\frac{x_{c,i}}{\sum_i x_{c,i}}\right)_{t-1}}{\left(\frac{\sum_c x_{c,i}}{\sum_c \sum_i x_{c,i}}\right)_{t-1}} \leq 1 \\ Exit_{c,i,t} &= 1, \text{ if } \frac{\left(\frac{x_{c,i}}{\sum_i x_{c,i}}\right)_t}{\left(\frac{\sum_c x_{c,i}}{\sum_c \sum_i x_{c,i}}\right)_{t-1}} \leq 1 \text{ and } \frac{\left(\frac{x_{c,i}}{\sum_i x_{c,i}}\right)_{t-1}}{\left(\frac{\sum_c x_{c,i}}{\sum_c \sum_i x_{c,i}}\right)_{t-1}} > 1 \end{aligned}$$

We must account for other variables that may influence *Entry* and *Exit* of cities' occupational specializations. In our econometric analysis, we use three-way-fixed effects models, with fixed effects for job classes (θ_j), cities (δ_c), and years

¹² For robustness purposes, we also computed Entry and Exit in its traditional form and run the same models in our analysis. The econometric results are very similar, with coefficients changing only slightly and keeping its statistical and economic significance.

(α), accounting for unobservable and invariant specific economic context. In addition, we use six control variables. Because a bigger and/or more diversified city is more prone to attract new jobs, we compute, for each city in each year, its total employment (log) (*City total employment*) and its number of occupational specializations (*City diversity*). To account for short-term employment growth (especially for years during the crisis), we compute yearly employment growth for cities (*City employment growth*) and for job classes (*Job employment growth*). Moreover, given global employment trends – like jobs involving many skills having higher labour demand (Moretti, 2012) – we also account for labour demand in each job class. We compute the total employment for each job class (*Job total employment*) and the number of cities in which a job class is an occupational specialization (*Job ubiquity*), as a measure of how common/systemic each job class is. As a proxy for the level of complexity a job class in terms of education, training and experience, we also construct an interaction term between the rarity of a job class in cities ($1/\textit{Job ubiquity}$) and the number of job specializations a city has (*City diversity*).

Following Hidalgo et al. (2007) and Boschma (2015), we compute geographical relatedness density (*GeoRelatedness Density*) for each job class j in city c in time t , which represents the relatedness of a new job class specialization to the set of job classes the city is already specialized in, in a given year. This density measure is derived from the relatedness of job class j to all other job classes i in which the city is specialized in, divided by the sum of relatedness of job class j to all other job classes i in country at time t :

$$\textit{GeoRelatedness Density}_{j,c,t} = \frac{\sum_{i \in c, i \neq j} \textit{GeoRel}_{j,i}}{\sum_{i \neq j} \textit{GeoRel}_{j,i}} * 100$$

Likewise, we compute density measures for similarity and for complementarity for each job class j in city c . *Similarity Density* represents the relatedness of a new job class specialization to the set of job classes the city is already specialized in, in terms of having similar skills. *Complementarity Density* represents the relatedness of a new job class specialization to the set of job classes the city is already specialized in, in terms of having complementary skills within the same industry cluster(s).

As explained before, we consider two jobs related in terms of local synergy when they are frequently co-located as city specializations, but showing low complementary and similarity. To calculate *Local Synergies Density*, we first

regress *GeoRelatedness Density* on *Similarity Density* and *Complementarity Density*, using a three way fixed effects model, as follows:

$$\begin{aligned} & \text{GeoRelatedness Density}_{j,c,t-1} = \\ = & \beta_1 \text{Similarity Density}_{j,c,t-1} + \beta_2 \text{Complementarity Density}_{j,c,t-1} + \theta_j + \delta_c + \alpha_t + \varepsilon_{j,c,t} \end{aligned}$$

We then save the residuals of the regression, $\varepsilon_{j,c,t}$, for computing our *Local Synergies Density* measure. It represents the relatedness of a new occupational specialization to the set of job classes the city is already specialized in, not in terms of having similar skills or complementary skills with existing job classes, but in terms of sharing the same location.

The panel data includes 11 years (from 2006 to 2016) and 733 job classes in 389 MSA. All our independent variables are lagged by one period (t-1), to avoid potential endogeneity problems. All our relatedness density variables are centred around the mean for purposes of coefficients' interpretation. Table 1 below shows some descriptive statistics.

Table 1. Descriptive statistics

| Statistic | N | Mean | St. Dev. | Min | Max |
|-------------------------|-----------|-----------|-----------|-------|-----------|
| Entry | 2,372,990 | 0.1 | 0.2 | 0 | 1 |
| Exit | 763,517 | 0.2 | 0.4 | 0 | 1 |
| Geo-Relatedness Density | 3,136,507 | 22.6 | 9.7 | 1.0 | 100.0 |
| Complementarity Density | 3,136,507 | 24.4 | 12.2 | 0.0 | 99.8 |
| Similarity Density | 3,136,507 | 24.3 | 9.7 | 0.0 | 93.3 |
| Local Synergies Density | 3,136,507 | 0.0 | 2.2 | -26.2 | 80.4 |
| City total employment | 3,136,507 | 271,105.4 | 508,783.7 | 6,900 | 6,367,200 |
| Job total employment | 3,136,507 | 143,874.5 | 339,269.6 | 30 | 4,041,050 |
| City employment growth | 3,136,507 | 0.01 | 0.1 | -0.5 | 2.3 |
| Job employment growth | 3,136,507 | 0.04 | 0.4 | -1.0 | 25.0 |
| City diversification | 3,136,507 | 312.5 | 128.9 | 50 | 654 |
| Job ubiquity | 3,136,507 | 165.8 | 128.1 | 1 | 389 |

*Number of observations = 733 OCC * 389 MSA * 11 years (from 2006 to 2016)

Table 2 presents the correlations between our four measures of relatedness density, including the interaction term between complementarity density and similarity density.

Table 2. Correlation between Measures of Relatedness Density

| Relatedness Density Variables | Geo-Relatedness Density | Complemen- tarity Density | Similarity Density | Local Synergies Density | Complem. D. * Similarity D |
|--------------------------------------|--------------------------------|----------------------------------|---------------------------|--------------------------------|-----------------------------------|
| Geo-Relatedness Density | 1,00 | 0.71 | 0.83 | 0.22 | 0.79 |
| Complementarity Density | | 1,00 | 0.76 | 0.00 | 0.92 |
| Similarity Density | | | 1,00 | -0.00 | 0.90 |
| Local Synergies Density | | | | 1,00 | -0.00 |
| Complem. D. * Similarity D | | | | | 1,00 |

4.2. Entry and Exit Models - Only Geographical Relatedness measure

In the first econometric model, we simply regress Entry of a new occupational specialization in a city on geographical relatedness density (plus controls and fixed effects), as follows:

$$\begin{aligned}
 Entry_{j,c,t} = & \beta_1 GeoRelatedness\ Density_{j,c,t-1} + \\
 & + \beta_2 \ln(City\ Total\ Emp.)_{c,t-1} + \beta_3 \ln(Job\ Total\ Emp.)_{j,t-1} + \\
 & + \beta_4 City\ Emp.\ Growth_{c,t-1} + \beta_5 Job\ Emp.\ Growth_{j,t-1} + \\
 & + \beta_6 City\ Diversification_{c,t-1} + \beta_7 Job\ Ubiquity_{j,t-1} + \\
 & + \beta_8 City\ Div./\ Job\ Ubiquity_{c,j,t-1} + \theta_j + \delta_c + \alpha_t + \varepsilon_{j,c,t}
 \end{aligned}$$

with, θ , δ and α being fixed effects, respectively for job classes, cities, and years (ε is the error term). In the second model, we simply repeat the first model, but for Exit of an occupational specialization in a city instead:

$$\begin{aligned}
 Exit_{j,c,t} = & \beta_1 GeoRelatedness\ Density_{j,c,t-1} + \\
 & + \beta_2 \ln(City\ Total\ Emp.)_{c,t-1} + \beta_3 \ln(Job\ Total\ Emp.)_{j,t-1} + \\
 & + \beta_4 City\ Emp.\ Growth_{c,t-1} + \beta_5 Job\ Emp.\ Growth_{j,t-1} + \\
 & + \beta_6 City\ Diversification_{c,t-1} + \beta_7 Job\ Ubiquity_{j,t-1} + \\
 & + \beta_8 City\ Div./\ Job\ Ubiquity_{c,j,t-1} + \theta_j + \delta_c + \alpha_t + \varepsilon_{j,c,t}
 \end{aligned}$$

Table 3. Entry and Exit Models – Only Geographical Measure of Relatedness

| | <i>Dependent variable (2006 – 2016):</i> | |
|--------------------------|--|-------------------------|
| | Entry (=1) | Exit (=1) |
| | (1) | (2) |
| Geo. Relatedness Density | 0.015*** (0.0001) | -0.024*** (0.0002) |
| City Total Emp. (ln) | 0.080*** (0.004) | -0.134*** (0.011) |
| Job Total Emp. (ln) | 0.004*** (0.001) | -0.078*** (0.004) |
| City Emp. Growth | -0.004 (0.003) | 0.040*** (0.009) |
| Job Emp. Growth | -0.002*** (0.0004) | -0.008 (0.005) |
| City Diversification | -0.001*** (0.00002) | 0.001*** (0.0001) |
| Job Ubiquity | -0.0002*** (0.00001) | 0.0003*** (0.00004) |
| City Div. / Job Ubiquity | -0.00000*** (0.000) | 0.00000*** (0.00000) |
| City fixed effects | Yes | Yes |
| Job class fixed effects | Yes | Yes |
| Year fixed effects | Yes | Yes |
| Observations | 2,372,990 | 763,517 |
| R ² | 0.075 | 0.068 |
| Adjusted R ² | 0.074 | 0.067 |

Note: The dependent variable entry (exit) equals one if a city c gains (loses) a relative occupational advantage in a given job class j in year t , and equals zero otherwise. All the independent variables are mean-centred and lagged by one year. Period under analysis ranges from 2005 to 2016 ($t-1 = [2005 - 2015]$). Coefficients are statistically significant at the $\hat{p}<0.1$, $*p<0.05$, $**p<0.01$, $***p<0.001$ level. Standard errors in parentheses.

The results presented in Table 3 show a statistically and economically significant impact of geographical relatedness density on both Entry and Exit. It shows a positive coefficient of 0.015 in the entry model, meaning that when *GeoRelatednes Density* increases by 10 percentage points, the probability of entry of a new job specialization in the city increases by 15%. Regarding Exit, the results show a negative impact of relatedness on the probability of a job class exiting a city's portfolio of occupational specializations. When *GeoRelatednes Density* increases by 10 percentage points, the probability of exit of a job class in the city decreases by 24%. Although almost all our control variables show

to be statistically significant, the interaction term representing job complexity (*City Div. / Job Ubiquity*) seems to be economically insignificant, in both Entry and Exit models.

The results so far are in line with the recent literature showing that relatedness seems to play a role in the renewal of the employment structure of US cities, at least when referring to geographical relatedness alone. But given the different reasons for job classes to co-occur, or put differently, given the different dimensions of relatedness that can be at work in the city, we still lack understanding of which dimensions influence the evolution of the employment structure in cities? To test this, instead of geographical relatedness density, we include our density measures for similarity, complementarity and local synergies.

4.3. Entry and Exit Models - All Dimensions of Relatedness

We start by regressing *Entry* and *Exit* on each of the three dimensions of relatedness density one at a time. Then, we include them all together, plus an interaction term between *Similarity Density* and *Complementarity Density*, to account for pairs of jobs that are simultaneously similar and complementary. The complete models for *Entry* and *Exit* are as follows:

$$\begin{aligned}
 Entry_{j,c,t} = & \beta_1 Complementarity\ Density_{j,c,t-1} + \beta_2 Similarity\ Density_{j,c,t-1} + \\
 & + \beta_3 Local\ Synergies\ Density_{j,c,t-1} + \\
 & + \beta_4 Similarity\ Density * Complementarity\ Density_{j,c,t-1} + \\
 & + \beta_5 \ln(City\ Total\ Emp.)_{c,t-1} + \beta_6 \ln(Job\ Total\ Emp.)_{j,t-1} + \\
 & + \beta_7 City\ Emp.\ Growth_{c,t-1} + \beta_8 Job\ Emp.\ Growth_{j,t-1} + \\
 & + \beta_9 City\ Diversification_{c,t-1} + \beta_{10} Job\ Ubiquity_{j,t-1} + \\
 & + \beta_{11} City\ Div./\ Job\ Ubiquity_{c,j,t-1} + \theta_j + \delta_c + \alpha_t + \varepsilon_{j,c,t}
 \end{aligned}$$

$$\begin{aligned} \text{Exit}_{j,c,t} = & \beta_1 \text{Complementarity Density}_{j,c,t-1} + \beta_2 \text{Similarity Density}_{j,c,t-1} + \\ & + \beta_3 \text{Local Synergies Density}_{j,c,t-1} + \\ & + \beta_4 \text{Similarity Density} * \text{Complementarity Density}_{j,c,t-1} + \\ & + \beta_5 \ln(\text{City Total Emp.})_{c,t-1} + \beta_6 \ln(\text{Job Total Emp.})_{j,t-1} + \\ & + \beta_7 \text{City Emp. Growth}_{c,t-1} + \beta_8 \text{Job Emp. Growth}_{j,t-1} + \\ & + \beta_9 \text{City Diversification}_{c,t-1} + \beta_{10} \text{Job Ubiquity}_{j,t-1} + \\ & + \beta_{11} \text{City Div./ Job Ubiquity}_{c,j,t-1} + \theta_j + \delta_c + \alpha_t + \varepsilon_{j,c,t} \end{aligned}$$

The results in Tables 4 and 5 show that each dimension of relatedness density, either alone or jointly, has a significant effect on the probability that a city specializes in a new job class or loses an existing job class. The stronger effect on *Entry* comes from *Local Synergies Density*, where an increase of 10% is associated with a 13% increase in the probability of entry. Its effect seems to be even stronger for *Exit*, with a decrease of 20% on exit probability when *Local Synergies Density* increases by 10%.

Table 4. Entry Models – All Dimensions of Relatedness Density

| | Dependent variable: | | | |
|---|---------------------------------|-------------------------|-------------------------|-------------------------|
| | Entry (=1) 2006 – 2016 | | | |
| | (1) | (2) | (3) | (4) |
| Complementarity Density | 0.003*** (0.00002) | | | 0.003*** (0.00003) |
| Similarity Density | | 0.004*** (0.00004) | | 0.002*** (0.00005) |
| Local Synergies Density | | | 0.011*** (0.0001) | 0.013*** (0.0001) |
| City Total Emp. (ln) | 0.032*** (0.004) | 0.034*** (0.004) | 0.053*** (0.004) | 0.082*** (0.004) |
| Job Total Emp. (ln) | 0.007*** (0.001) | 0.007*** (0.001) | 0.005*** (0.001) | 0.004*** (0.001) |
| City Emp. Growth | 0.002 (0.003) | 0.001 (0.003) | -0.001 (0.003) | -0.005* (0.003) |
| Job Emp. Growth | 0.003*** (0.0004) | 0.003*** (0.0004) | -0.001 (0.0004) | -0.002*** (0.0004) |
| City Diversification | -0.0002*** (0.00002) | -0.0003*** (0.00002) | -0.001*** (0.00002) | -0.001*** (0.00002) |
| Job Ubiquity | 0.0003*** (0.00001) | 0.0003*** (0.00001) | -0.00003** (0.00001) | -0.0002*** (0.00001) |
| Complementarity Density * Similarity Density | | | | 0.00001*** (0.00000) |
| City Div. / Job Ubiquity | -0.00000*** (0.000) | -0.00000*** (0.000) | -0.00000*** (0.000) | -0.00000*** (0.000) |
| City fixed effects | Yes | Yes | Yes | Yes |
| Job class fixed effects | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes |
| Observations | 2,372,990 | 2,372,990 | 2,372,990 | 2,372,990 |
| R ² | 0.069 | 0.066 | 0.069 | 0.078 |
| Adjusted R ² | 0.068 | 0.065 | 0.068 | 0.077 |

Note: The dependent variable entry equals one if a city c gains a new relative occupational advantage in a given job class j in year t , and equals zero otherwise. All the independent variables are mean-centred and lagged by one year. Period under analysis ranges from 2005 to 2016 ($t-1 = [2005 - 2015]$). Coefficients are statistically significant at the $\ast p < 0.1$, $\ast\ast p < 0.05$, $\ast\ast\ast p < 0.01$, $\ast\ast\ast\ast p < 0.001$ level. Standard errors in parentheses.

Table 5. Exit Models – All Dimensions of Relatedness Density

| | Dependent variable: | | | |
|---|--------------------------------|-------------------------|-------------------------|-------------------------|
| | Exit (=1) 2006 – 2016 | | | |
| | (1) | (2) | (3) | (4) |
| Complementarity Density | -0.005*** (0.0001) | | | -0.005*** (0.0001) |
| Similarity Density | | -0.007*** (0.0001) | | -0.005*** (0.0001) |
| Local Synergies Density | | | -0.014*** (0.0003) | -0.019*** (0.0003) |
| City Total Emp. (ln) | -0.065*** (0.011) | -0.074*** (0.011) | -0.089*** (0.011) | -0.130*** (0.011) |
| Job Total Emp. (ln) | -0.056*** (0.004) | -0.058*** (0.004) | -0.071*** (0.004) | -0.072*** (0.004) |
| City Emp. Growth | 0.013 (0.009) | 0.016* (0.009) | 0.025*** (0.009) | 0.037*** (0.009) |
| Job Emp. Growth | -0.028*** (0.005) | -0.028*** (0.005) | -0.016*** (0.005) | -0.011** (0.005) |
| City Diversification | -0.0004*** (0.00005) | -0.0003*** (0.00005) | 0.0003*** (0.0001) | 0.001*** (0.0001) |
| Job Ubiquity | -0.001*** (0.00004) | -0.001*** (0.00004) | -0.0003*** (0.00004) | 0.0001** (0.00004) |
| Complementarity Density * Similarity Density | | | | 0.0001*** (0.00000) |
| City Div. / Job Ubiquity | 0.00000*** (0.00000) | 0.00000*** (0.00000) | 0.00000*** (0.00000) | 0.00000*** (0.00000) |
| City fixed effects | Yes | Yes | Yes | Yes |
| Job class fixed effects | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes |
| Observations | 763,517 | 763,517 | 763,517 | 763,517 |
| R ² | 0.065 | 0.061 | 0.060 | 0.073 |
| Adjusted R ² | 0.064 | 0.060 | 0.058 | 0.072 |

Finally, in table 6 below, we repeat the complete models but with standardized variables of relatedness density instead (scaled versions of our previous relatedness density variables), in order to jointly test their explanatory power on *Entry* and *Exit* and compare coefficients. We find that an increase of 1 standard deviation in *Local Synergies Density* increases the probability of entry of a new job class in the city's portfolio of job specializations by 14%, and decreases the probability of exit by 22%. An increase of 1 standard deviation in *Complementarity Density* increases the probability of entry of a new job specialization by 3%, and decreases the probability of exit by 6.5%. And when *Similarity Density* increases

1 standard deviation, the probability of entry increases by nearly 3% and exit probability decreases by 4.6%. The only finding not in line with expectation is the effect of the combination of complementarity and similarity: it shows a negative effect on entry and a positive effect on exit, although the effects are not sizable.

Table 6. Entry and Exit Models – All Dimensions of Relatedness Density (scaled)

| | Dependent variable (2006 – 2016): | |
|--|-----------------------------------|-----------------------|
| | Entry (=1) | Exit (=1) |
| | (1) | (2) |
| Complementarity Density (sc) | 0.037*** (0.0003) | -0.066*** (0.001) |
| Similarity Density (sc) | 0.024*** (0.0004) | -0.046*** (0.001) |
| Local Synergies Density (sc) | 0.137*** (0.001) | -0.214*** (0.002) |
| City Total Emp. (ln) | 0.089*** (0.004) | -0.136*** (0.011) |
| Job Total Emp. (ln) | 0.004*** (0.001) | -0.071*** (0.004) |
| City Emp. Growth | -0.006** (0.003) | 0.039*** (0.009) |
| Job Emp. Growth | -0.002*** (0.0004) | -0.008* (0.005) |
| City Diversification | -0.001*** (0.00002) | 0.002*** (0.0001) |
| Job Ubiquity | -0.0004*** (0.00001) | 0.001*** (0.00004) |
| Complem. Density (sc) * Similarity Density (sc) | -0.001*** (0.0002) | 0.009*** (0.0005) |
| City fixed effects | Yes | Yes |
| Job fixed effects | Yes | Yes |
| Year fixed effects | Yes | Yes |
| Observations | 2,372,990 | 763,517 |
| R ² | 0.076 | 0.073 |
| Adjusted R ² | 0.076 | 0.071 |

Note: The dependent variable entry (exit) equals one if a city *c* gains (loses) a new relative occupational advantage in a given job class *j* in year *t*, and equals zero otherwise. All the independent variables are mean-centred, and lagged by one year. Additionally, relatedness density variables are scaled by their sd, for comparison purposes (scaled variables denoted by "sc". Period under analysis ranges from 2005 to 2016 (t-1 = [2005 - 2015]). Coefficients are statistically significant at the $\hat{p}<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$ level. Standard errors in parentheses.

5. CONCLUDING REMARKS

While many studies have looked at regional diversification into new products (Hidalgo et al., 2007), new industries (Neffke et al., 2011) or new technologies (Kogler et al., 2013; Rigby, 2015), this paper has taken an occupational-network approach examining the evolution of job portfolio's in US cities. The paper replicates the result found in other studies on the evolution of occupational structures in cities (Brachert, 2016; Muneeppeerakul et al., 2013; Shutters et al., 2016) that cities enter new jobs related to existing ones in cities, and exit existing jobs unrelated to their job portfolio's. What is new about this paper is that we have unpacked three mechanisms through which the entry and exit of jobs in cities takes place. While previous papers looked at the effect of geographical relatedness only, we unravel three mechanisms through which the effect of geographical relatedness might work because co-location of jobs does not tell us much about the forces that make jobs co-occur in the same city: new local jobs may be related to existing local jobs because they share similar skills or provide complementary tasks, or both, or because they benefit from being co-located.

First, we constructed a job space that represents a network of interdependent job classes that includes the three dimensions through which jobs may be related to each other. In doing so, we can unravel links between pairs of jobs in terms of being similar, being complementary, being both similar and complementary, or in terms of sharing local synergies. Second, we investigated the importance of each of these three job relatedness dimensions for the evolution of jobs in 389 US cities for the period 2005-2016. For this purpose, we introduced a new methodological approach to distinguish between the three relatedness effects.

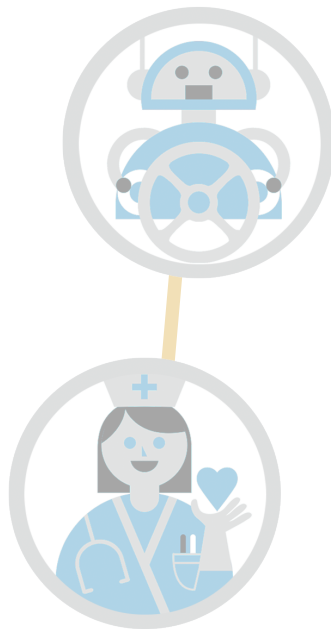
The main finding is that all three relatedness dimensions (similarity, complementarity and local synergies) increase the chances of entry of a new job in a city on the one hand, and decrease the probability of disappearance of an existing job in a city on the other hand. Moreover, we found the negative effect of relatedness on exits of jobs to be stronger than the positive effect of relatedness on entry of jobs: all three relatedness dimensions seem to prevent exit of jobs in cities more than promoting entry of jobs in cities. The local synergy density effect shows the largest effect on both entry and exit: this outcome suggests that the stronger local synergies across job classes are, the greater the effect on diversification and the harder to dislocate existing job classes. The complementarity density effect reflects the tendency of an increasing division

of labour in cities which brings higher levels of interdependence between job specializations (Shutters et al., 2018) where each worker's productivity depends on whether or not she has access to co-workers with specialized skills and know-how that complement her own (Neffke, 2017). The similarity density effect found is in line with the tendency of firms and people to cluster geographically to benefit from a pool of labour with related skills (Neffke & Henning, 2013). Similarity seems to prevent exit and promote entry of jobs in cities, but not in combination with complementarity.

Although this paper provides an important step to unpack relatedness, it is still far from comprehensive (Boschma, 2017). First, while we have started to unravel the geographical density effect (controlling for similarity and complementarity), there is a need to investigate what the local synergies dimension consists of. Second, we need more studies in other countries to shed more systematic light on the importance of the different relatedness dimensions. Third, the three dimensions of relatedness density might play different roles depending on the level of knowledge complexity of activities (Balland & Rigby, 2017) and should be employed and tested in studies on regional diversification into new products, industries or technologies, besides new jobs. Finally, we have to make an effort to include institutions in this framework, because regional diversification might also be affected by institutional requirements that different jobs, industries or technologies have in common (Boschma & Capone, 2015).

CHAPTER 4

Impacts from automation diffuse locally: a novel approach to estimate jobs risk in US cities



Contribution: Single-author

Published as: Teresa Farinha (currently under review) Impacts from automation diffuse locally: a novel approach to estimate jobs risk in US cities

1. INTRODUCTION

The first quantitative study on *“How Susceptible Are Jobs to Computerisation”* (Frey & Osborne, 2013) put the world in jitters by estimating that 47% of US jobs were at risk of automation in the next few decades. It also sparked strong critics and, since then, many more studies have been added to the Future of Work literature, for different countries, levels of analysis, in less and more conservative approaches, including extensive regional comparative studies (Arntz et al., 2017; Bechichi, Grundke et al., 2018; Brynjolfsson et al., 2018; Lund et al., 2019; Manyika et al., 2017; Nedelkoska & Quintini, 2018; Roux, 2018). And yet, rather than converging, their estimations vary profoundly in numbers and forms of disruption (Winick, 2018).

A common caveat in these studies is that they only consider the technical feasibility of automating a task against the share of non-automatable tasks within a job (or within the portfolio of jobs in a city or country). But two workers with the same type of jobs (job-class) in different cities might have a different risk of having their jobs automated, given the local specificities that shape the diffusion of technology (Bessen et al., 2019; Brynjolfsson et al., 2018). Two main effects occur. First, once it starts automating jobs in a city, a new technology can more easily adapt to automate similar neighbouring jobs (Bechichi et al., 2018; Manyika et al., 2017; Nedelkoska et al., 2018). Second, besides substituting jobs, new technologies may also complement jobs, raising their productivity and labour demand (Autor, 2015). For instance, AI procedures may more efficiently assist health professionals with medical records and diagnosis. Here too, impacts seem to diffuse locally, this time positive, to where more complementary jobs can collaborate with new robots.

This, how local capabilities are related to each other, and how it affects their evolution, has been robustly investigated in the Evolutionary Economic Geography (EEG) literature (Boschma & Frenken, 2006; Hidalgo et al., 2018). For instance, relatedness between jobs seems to favour employment growth and job diversification in a city, preventing the exit and facilitating the entry of new job specializations (Alabdulkareem et al., 2018; Muneeppeerakul et al., 2013; Neffke et al., 2018). This “magnet” effect seems particularly strong for the relatedness dimension of local synergies (specialized amenities and knowledge spillovers), but also for complementarities (input-output relationships) and similarities between jobs (Farinha et al., 2019). However, no study has shown how relatedness might, in particular, support the spread of automation impacts from one job to another, i.e., the “diffusion” effects of relatedness (Jun et al., 2019; Morrison et al., 2013).

This paper aims to address this gap in both literatures, the Future of Work and the EEG. It goes beyond the technical risk of automation to investigate the relatedness links through which impacts diffuse from automatable jobs to neighbouring jobs in a city, putting them in higher jeopardy or safety. Two pathways were found, similarity to high-risk jobs and complementarity to high-risk jobs. I test their impact on employment growth within job classes in US cities, from 2007 to 2016.

The structure of the paper is as follows. Section 2 presents the determinants of technology adoption and the role of local network dynamics. Section 3 describes the data, and Section 4 presents the results. Finally, Section 5 discusses the findings in this paper and implications for policy.

2. WHAT DETERMINES A JOB'S RISK OF AUTOMATION?

2.1. Technical feasibility

The execution of each task requires a certain set of skills, some of which offering better applications of technology than others (Autor et al., 2003; Brynjolfsson et al., 2018). This shapes the technical feasibility of automating tasks within a job. The ones that rely the most on automatable tasks are at higher technical risk of becoming automated (Frey & Osborne, 2013).

From Robotics to Artificial Intelligence (AI) and beyond, the new technologies evolve rapidly, and so its bottlenecks (Autor, 2014; Perrault et al., 2019). Initially confined to codified knowledge, computer algorithms could only automate routine and manual tasks, usually associated with low skills (Acemoglu & Autor, 2011; Autor & Dorn, 2013; Autor et al., 2003). Later on, it could also take over cognitive and non-standardized tasks, collaborate with other robots (M2M), even “learn” through experience and surpass humans in image and speech recognition (Brynjolfsson & Mitchell, 2017; Klinger et al., 2018; McAfee & Brynjolfsson, 2016). Now, complex cognitive tasks are easy for robots, which can transform all sorts of problems into prediction ones, drastically reshaping labour in both manufacturing and services (Agrawal et al., 2018; Decker et al., 2017). Conversely, tasks such as gardening, caring for others, negotiating, usually require tacit skills of creativity, social and emotional intelligence, and cognitive flexibility¹³, which are very difficult to codify in whatever form of language or require greater amounts of computation

13 Certain skills have low demand if isolated, yet high if combined with other skills (e.g., offer surplus of STEM skills in academia versus shortage in policy, where they require social skills too (Benzell et al., 2019; Xue & Larson, 2015).

(Bradberry, 2017; J. Davies, 2019; Decker et al., 2017; WEF, 2016; World Bank, 2019). In sum, what the new technologies can and cannot automate at each point in time¹⁴ shapes the future demand for labour.

However, although a necessary condition, technical feasibility is not sufficient to generate adoption of a new technology. Firms, institutions, and society at general, take time to adopt technology (Brynjolfsson, Rock, & Syverson, 2017), in some places more than others (Bresnahan & Greenstein, 1996). That is why, for instance, cashiers in the busy city of New York can be expected to get automated sooner than cashiers in the winery region of Napa Valley. In other words, geography conditions the diffusion of technology and, therefore, the impacts of automation.

2.2. Local feasibility

A myriad of local idiosyncrasies conditions the reach of AI, its relative costs to labour, and ultimately firms' choice for AI adoption (Craglia et al., 2018; Nedelkoska & Quintini, 2018). When a new technology has the potential to increase total factors productivity and profits, firms evaluate the relative costs of AI versus labour, considering all the reorganization necessary to accommodate the new technology (Brynjolfsson et al., 2018). For instance, in implementation, maintenance, training, and displacement, which highly depend on the bargaining power of labour unions and regulations on dismissal and working conditions (Harris & Krueger, 2015; Kochan, 2016; Nedelkoska & Quintini, 2018; Wisskirchen et al., 2017). Also, AI is "fed" by big data, and data access is heavily conditioned by technical and legal issues (ownership, cybersecurity, etc.).

In sum, from technical feasibility to actual automation of tasks, the firm must also assure the local feasibility (economic, legal, etc.) of a new technology (Manyika et al., 2017). Depending on the specific local context, some firms and institutions (governments, universities, etc.) are more prepared to adjust than others. In result, new technologies diffuse quicker in some cities and slower in others, as each place has its own portfolio of capabilities (workers, firms, institutions, etc.) and intricate web of interactions and connectivity between them (relatedness).

We can only account for this by analysing the local structure of capabilities, which has been extensively investigated in the Evolutionary Economic Geography

¹⁴ Initially bounded by Polanyi's paradox i.e., the fact that "we know more than we can tell" (Polanyi, 1966), now the most natural to humans seems to be the most difficult for robots (Moravec, 1988)

(EEG) literature. For cities, regions, and countries, at several levels of analysis, being industries, products, jobs, or knowledge (Boschma & Frenken, 2006; Hidalgo et al., 2018), relatedness has robustly shown to affect both individual performance and the evolution of the local structure. Particularly, relatedness between jobs seems to favour employment growth and job diversification in a city (Alabdulkareem et al., 2018; Farinha et al., 2019; Muneeppeerakul et al., 2013; Neffke et al., 2018; Shutters et al., 2018).

Moreover, these “magnet effects” of relatedness – also referred as forces of agglomeration (Marshall, 1920) – may operate in three distinct ways. Local capabilities may co-locate because they share similar skills (forming labour market pools), complementary skills (input-output chains), or local synergies (specialized amenities and knowledge spillovers). Each of these three dimensions of relatedness has its own way of pulling capabilities together, with local synergy showing a particularly strong “magnet effect” in US cities (Farinha et al., 2019). And, although orthogonal to each other, they may occur in simultaneous. For instance, both similarity and complementarity make the relationship between lawyers and paralegals.

But relatedness might have an additional role, still poorly investigated, yet particularly relevant under technological transitions or any serious threat to labour systems. Relatedness seems to channel the spread of impacts between local capabilities – “diffusion effects”. For instance, the Great Recession of 2008 was much caused by cascading impacts beyond the initial real estate bubble. Negative impacts diffused mainly through complementarities to products and services for which the final demand was contracting the most (Dolfman et al., 2018; Goodman & Mance, 2011). For instance, through input-output linkages in the car industry, from manufacturing to insurance. While jobs of low recessionary risk (e.g., doctors and nurses) could anchor the labour demand of their complementary jobs (e.g., medical equipment technicians)¹⁵.

Also, the relatedness links through which impacts diffuse in a city differ from one type of event to another. In the case of disruptive technologies, rather than massively contracting the labour demand, they reshuffle the allocation of production factors (Acemoglu & Restrepo, 2019). Two main effects weight

¹⁵ See Appendix 2 for how relatedness to jobs of high recessionary risk may have affected all other jobs in US cities.

against each other in firm's choices between capital and labour (Acemoglu & Restrepo, 2019; Lordan & Neumark, 2018) – displacement costs and productivity gains.

Displacement costs

Firms choose to automate tasks where the alternative choice of labour would be more expensive (Feng & Graetz, 2015). As low skill jobs usually display low labour cost-benefit ratio (especially in the initial development states of the technology), they might have less probability of being substituted by modern service robots than its technical feasibility would tell (Decker et al., 2017).

But technology gets cheaper the more it gets implemented by firms and diffuses in the local economy, as later adopters benefit and learn from the pioneers' adoption process (Manyika et al., 2017). And it diffuses quicker (shorter period from technical feasibility to actual implementation) the more similar tasks are available to the technology, within and across firms in the city. In other words, the more similarities exist between the jobs being automated and their neighbouring jobs, the easier to adapt the new technology to automate the latter too.

In result, automation tends to concentrate among jobs that share a similar set of automatable skills, despite having routine or non-routine, standard or cognitive tasks (Nedelkoska et al., 2018). Workers at the core of such labour pools face longer adaptation paths towards non-automatable jobs and higher probability of unemployment (Alabdulkareem et al., 2018). Conversely, similarities to low-risk jobs might facilitate labour flows to less automatable jobs.

Productivity gains

Automation has the power not only to substitute human labour, but also to augment it, wherever human-computer collaboration can be exploited (Brynjolfsson et al., 2018; Licklider, 1960; Sankar, 2012). In many jobs, only particular tasks can be rendered by robots in higher quality, in which human-computer collaboration may by far overcome the results of AI alone. For instance, certain delicate medical operations might be better performed by robots than humans, assisting the surgeon in its overall job, but hardly substituting all her/his tasks.

Technology also transforms and creates tasks within jobs. AI models and algorithms are not always interpretable or explainable, which creates new

possibilities for their coordination with humans (Autor, 2014; Lin, 2011). Also, AI may create new tasks meant exclusively for robots, thus, definitely not substituting labour, possibly expanding it. For instance, the human capacity to land on Mars, or to dive towards the bottom of the Mariana Trench, is limited compared to a robot (Decker et al., 2017) and considerably eases the work of scientists. Finally, AI gradually leaves well-defined environments (like factories) and enables workers with no skills in information technology to control new robots and AI systems (Decker et al., 2017). A recent study (Merritt, 2018) found that AI have not only substituted but also changed traditional office jobs in Mexico.

In result, new technologies tend to bring higher productivity, earnings, and labour demand for the jobs that are complementary to recently automated ones (Autor, 2015; Decker et al., 2017; Griliches, 1969; Kremer, 1993). Accordingly, studies (Bessen et al., 2020; Dahlin, 2019; Graetz & Michaels, 2018) have found a positive impact of modern robots in labour productivity growth¹⁶.

In sum, the demand for labour tends to increase in jobs complemented, and not substituted, by the new technologies, initiating adjustments in the labour supply and adaptation of the workforce (Bessen et al., 2019). Moreover, each city has its own portfolio of jobs and structure of similarities and complementarities between them, making technology to spread unevenly, like water choosing the best path to penetrate the soil. In each city, as a result of those two opposite effects, displacement costs and productivity gains, automation impacts diffuse through the existing structure of relatedness, “selecting” which workers lose their jobs to automation and which workers benefit from it in terms of productivity gains and labour demand.

For each worker in each city, this means that, besides the technical risk of automating her/his job, the overall risk also depends on how similar or complementary she/he is to co-workers in that city, given their risk to automation. At the end, the stronger complementarities, and weaker similarities, a job has with neighbouring high-risk jobs, the better it is expected to perform in terms of productivity and labour demand¹⁷. The opposite would rather increase chances of unemployment.

¹⁶ Also, a stronger positive impact for high skill jobs. This is expected in certain industries like manufacturing, where capital tends to be more complementary to high skills (Acemoglu et al., 2020; Griliches, 1969). While in personal care, for instance, low skill jobs tend to have low automation risk (Atkinson, 2017; Nedelkoska & Quintini, 2018).

¹⁷ Or conversely, being similar, and not complementary, to low-risk jobs would also have a positive impact.

Therefore, in this paper, I test the following two hypotheses. Employment growth is (*H1*) higher for job-classes that are more complementary to neighbouring high-risk jobs (high technical risk of automation), and (*H2*) lower for job-classes that are more similar to neighbouring high-risk jobs.

3. EMPLOYMENT DATA

In order to test *H1* and *H2*, first, I need to identify which jobs have high technical risk of automation. Second, the similarities and complementarities between jobs in each city. Third, an employment performance indicator, and relevant control variables. This requires a considerable amount of data from distinct sources that is uniquely available for the USA, as follows.

The Bureau of Labour Statistics (BLS) provides yearly employment statistics for around 800 detailed job-classes (7digit OCC), within 22 job families (2digit OCC), 400 industries (NAICS), and 400 Metropolitan Statistical Areas (MSA), which represent unified labour systems (US Census Bureau, 2020). Also, the Industry Sectoring Plan (ISP) cluster classification, which unifies product value chains based on inter-industry linkages. This paper uses employment data at the OCC-MSA and OCC-NAICS levels of analysis, the latter which I aggregate into an OCC-ISP dataset.

The Occupational Information Network (O*NET) provides extensively detailed data on the work scope of each job-class. This paper uses two variables. The Intermediate Work Activities (IWA) describes how much each task is required in each job-class (% of importance), in an optimal level of analysis that allows network computation while providing enough detail to reveal the underlying skills in each task. The Job Zone captures the level of required skills for each job-class (low=1, high=5). It covers academic degree, experience, on-the-job training, and certifications.

Finally, I use Atkinson (2017)'s index of automation risk (high=1, low=5)¹⁸, provided by the Information Technology and Innovation Foundation (ITIF). It estimates the technical risk of automation in each job-class by combining BLS employment data¹⁹ with experts evaluation on the possibility of a job being

18 E.g., cashiers and credit analysts have high risk, while actors, dentists, firefighters, hairdressers have low risk.

19 Data series for 2014-2024, available in <https://data.bls.gov/projections/occupationProj>

radically altered by new technologies given its work scope (Atkinson, 2017). It directly covers all job-classes, aligns well with BLS employment projections (which also accounts for technological change) and, as expected, is weakly correlated with educational background ($=-0.4$), since recent developments of AI can substitute high-skill jobs too.

Since OCC and MSA classification schemes had major revisions before 2005 and after 2016, I restrict the period of analysis within those years. Also, I drop the OCCs and MSAs created/ceased during that period, and the “All Other” type of classes without match in O*NET data. The final data includes 733 OCC, 389 MSA, 179 ISP, 332 IWA, 5 Job Zones, 5 ITIF Automation Risk categories, and 12 years. For ease of interpretation, from here on, I will refer to those as “job”, “city”, “cluster”, “task”, “JobSkills”, and “JobAutRisk”, respectively. And the three datasets of OCC-MSA, OCC-ISP, and OCC- IWA, as “job-city”, “job-cluster”, and “job-task” datasets. Next, I transform the data to build the variables of interest, as presented in the next subsections.

3.1. The Job Space Under Automation

In order to capture the structure of jobs in cities, scholars have built network representations of the workforce (Alabdulkareem et al., 2018; Farinha et al., 2019; Muneeppeerakul et al., 2013; Neffke et al., 2018; Shutters et al., 2018)²⁰. In this paper, such network should also show which jobs are similar or complementary to other existing jobs of high (or low) technical risk of automation.

Therefore, I build a network representation of the USA workforce under automation, where nodes are job-classes, links are the level of relatedness between them, with two types of links for the relevant dimensions of relatedness (complementarity and similarity), and three types of nodes for technical risk of automation (high, medium, or low). This way, although the network is not directed (relatedness matrices are symmetric), from the perspective of each node (source node), both its links and the destination of its links (target node)²¹ are captured as having high, medium, or low technical risk of automation.

20 Recent applications of network visualization tools have resulted in remarkable online interactive platforms for exploring skills, professions, country and city profiles. To name a few, the Observatory of Economic Complexity (<https://oec.world>), the Data USA (<https://datausa.io/>), and the Skillscape (<http://skillscape.mit.edu>).

21 Each link, directed or undirected, connects two nodes, commonly referred as the source node and the target node.

Figure 1 below allows the visualization of this network – *Job Space Under Automation*. The nodes' shapes represent the three levels of technical risk of automation. And their colours, their job families. Finally, links' colours are dimensions of relatedness (when two job-classes are simultaneously similar and complementary to each other, the stronger is displayed).

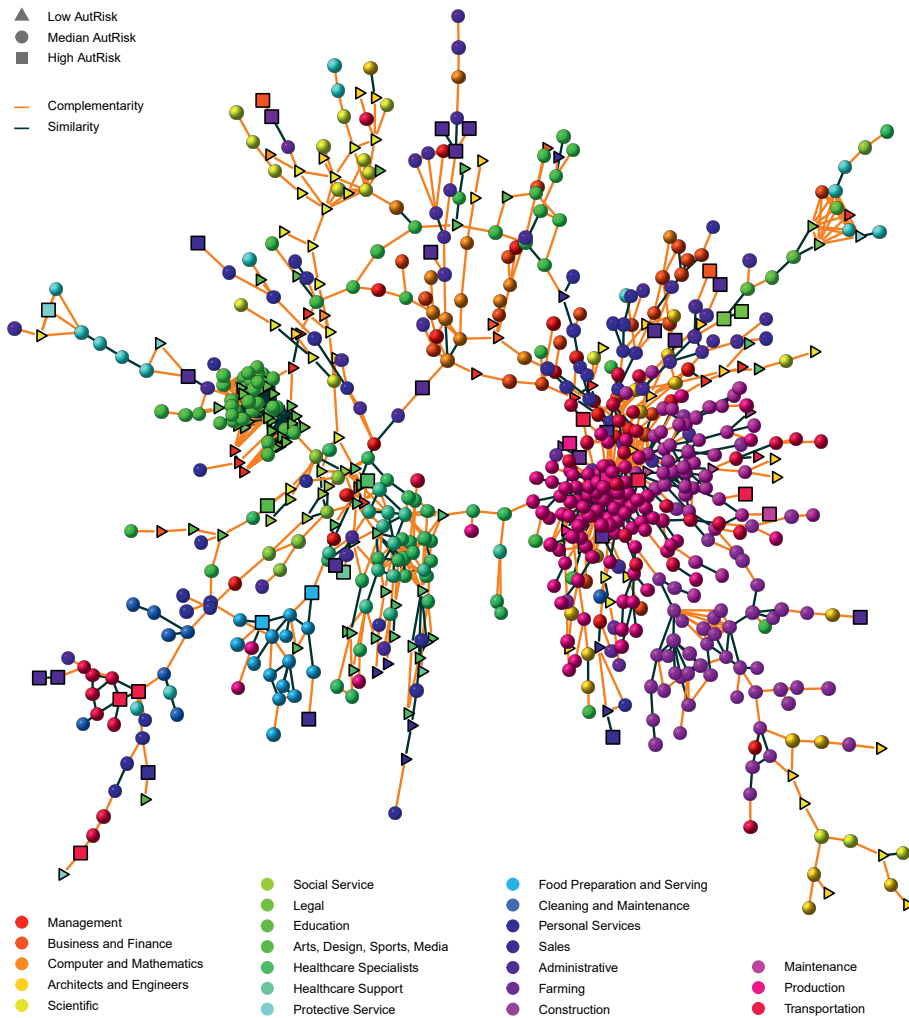


Figure 1. The Job Space Under Automation

For instance, farm inspectors (see examples in Figure 2 below) have low technical risk of becoming automated. Moreover, they are similar and also, although less, complementary to farm sorters, which have high automation risk. Thus,

according to the literature discussed above, while farm inspectors can benefit with the automatization of colleague sorters, the latter can use its similarities to the former to adapt skills and become agricultural inspectors, which is expected to be in high demand in the near future. A different example, taxi drivers and bus drivers are not particularly complementary, they rather share similar tasks (that require navigation skills, etc.). Although the former has high risk of automation, bus drivers require some other skills that are bottlenecks of AI (associated with social trust and safety), making them less susceptible to automation. Yet, in cities where automated cars are already well established, the technology can more easily expand to automate bus drivers too.

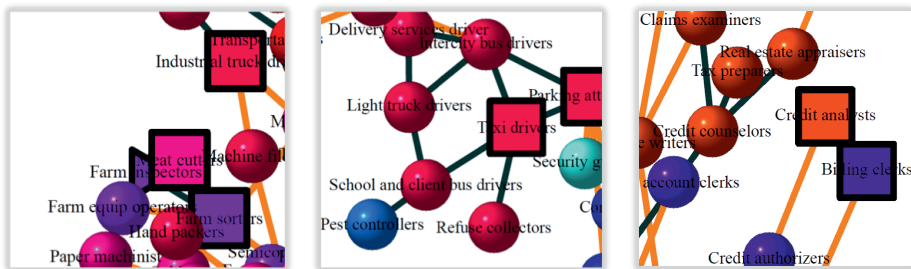


Figure 2. Examples (“zoom in” sections of the Job Space Under Automaton)

At the city level, the *Job Space Under Automation* assumes the same network structure of nodes and links, but only displaying the nodes that are job specializations of that city in a given year (nodes’ colour turn grey if that job is not a specialization of the city). Moreover, for a comparison between cities, one can build a multi-layer network, choosing layers of the *Job Space Under Automation* for specific cities, years, or relatedness dimensions.

Figure 3 below allows a separate visualization of complementarities and similarities (which shows when two nodes are simultaneously similar and complementary), for both Boston and for Napa Valley, in 2016. Boston shows multiple job specializations in Education, Health Care, and Sciences (left side of the network), but also in Management, Finance, and Computer jobs (centre of the network). Whereas Napa Valley is much less diversified, with job specializations revolving around the wine industry (right side of the network), including food engineering, restaurants, and leisure. Such different job portfolios between Boston and Napa Valley must translate into different diffusion paths for the impacts of automation. For example, credit analysts (also in Figure 2 above) is a specialization of Boston but not of Napa Valley. Therefore, effects of automating

credit analysts should reflect on credit analysts' employment levels, but also on jobs related to them, such as credit authorizers, which are more strongly represented in Boston than in Napa Valley.

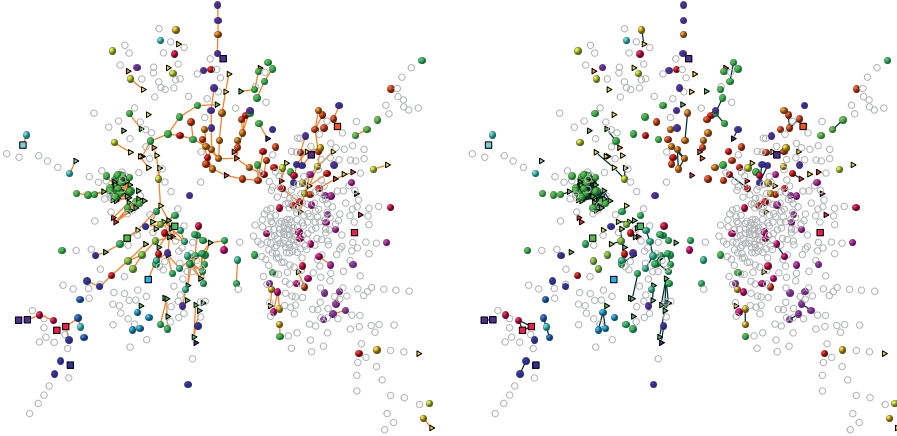


Figure 3. Network layers – complementarities and similarities in Boston, 2016

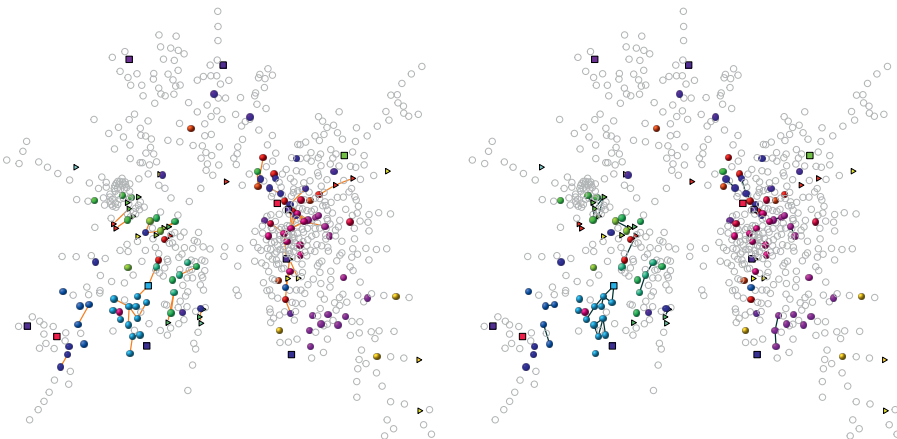


Figure 4. Network layers – complementarities and similarities in Napa Valley, 2016

3.2. Relatedness to high-risk-jobs

From the *Job Space under Automation* of US cities, I “extract” the variables of interest for testing hypothesis *H1* and *H2*, as the level of relatedness of each job (source node) in each city to other existing jobs (target nodes) that have high technical risk of automation. It is computed for both similarities and complementarities, as they are expected to have opposite effects on

employment growth. It follows a sequence of computation phases, described below.

Phase 1 computes each city's portfolio of job specializations that have high risk of automation, based on ITIF's index of jobs' automation risk (Atkinson, 2017). First, I transform the BLS data on the number of workers at the job-city level, into a $C \times P$ matrix, where C is the number of cities, and P is the number of job-classes. Second, I compute the commonly used location quotient (LQ) for each cell (c, i) of the matrix, as follows:

$$LQ_{c,i} = \frac{\left(\frac{x_{c,i}}{\sum_j x_{c,i}} \right)}{\frac{\sum_c x_{c,i}}{\sum_c \sum_j x_{c,i}}}$$

$LQ_{c,i}$ describes how much specialized a city c is in job i , in relation to the national employment levels of that job i . If higher than one, the job i is "over-represented" in city c (otherwise "sub-represented"). Third, I transform the above into a binary matrix, where =1 means the job i is "overrepresented" in city c and has $JobAutRisk=1$ (=0 otherwise).

Phase 2 computes the matrices for similarity and complementarity dimensions of relatedness (as in Farinha et al., 2019). First, I use O*NET data on tasks to build a $1 \times W$ vector for each job i , where W is the number of task classes (as in Hasan et al., 2015). Second, I join them to form a job-task matrix, $W \times P$. Third, I compute a location quotient for tasks in jobs, $LQ_{w,i}$, in each cell (w, i) , and transform it into binary, where =1 means the task w is of crucial importance for job i (=0 otherwise). Forth, I build a symmetric $P \times P$ matrix, in which each cell (i, j) contains similarity between jobs i and j , measured as the probability that a specific task is crucial for job i given that is also crucial for job j (co-occurrences measure, as in Eck & Waltman, 2009). Finally, I repeat Phase 2 steps for *Complementarity*, this time departing from the job-cluster data, to measure how often two jobs are jointly required in the same value chain. *Similarity* and *Complementarity* are lower bounded by zero (no task/cluster is relevant for job i and j) and upper bounded by one (all tasks/clusters that are relevant in job i are also relevant in job j , and vice versa).

Phase 3 computes the final variables for relatedness density of each job in each city considering only the target nodes that have high risk of automation. For

this, I combine the two P×P relatedness matrices from Phase 1 and 2 (*Similarity* and *Complementarity*) into two P×C relatedness density matrices (as in Balland et al., 2019) as follows:

$$RelatednessToHighRisk_{i,c} = \frac{\sum_{i \neq j, j \in c, JobAutRisk_j=1} Relatedness_{i,j}}{\sum_{i \neq j} Relatedness_{i,j}} * 100$$

In result, the *Similarity* density matrix contains, in each cell (i, j), similarity density of job i to all other jobs j that exist in the city and have high technical risk of automation (note how a job in a city can have a low relatedness density to high-risk-jobs even having an overall high relatedness density, in which case, its relatedness to existing jobs would concern mostly medium and low risk jobs). The *Complementarity* density matrix contains complementarity values instead. Finally, I transform these matrices into the following two relatedness variables:

- (i) Similarity of a job to neighbouring high-risk jobs (*SimilarToHighRisk*)
- (ii) Complementarity of a job to neighbouring high-risk jobs (*ComplementaryToHighRisk*)

4. IMPACT OF AUTOMATION ON EMPLOYMENT GROWTH IN US CITIES

4.1. Variables and descriptives

Although impossible to pin-point its start, the first wave of automation under the current technological transition has obviously arrived and has been affecting jobs at speeding rates in the last few years (Brynjolfsson & McAfee, 2014; Rao & Verweij, 2018). Therefore, the local net impacts from automation (replacement needs and job losses, higher productivity and labour demand, transformations within the job) must reflect on employment growth at the job-city level. This is, of course, controlling for other factors that affect this performance indicator under the same period. Such as the past Great Recession, which heavily affected employment from 2008 to 2014, when it finally reached pre-crisis levels (Dolfmanm et al., 2018; NBER, 2010).

Therefore, in order to test how relatedness to high-risk-jobs have recently affected local employment performance, while “jumping” the past recessionary shock, the dependent variable in this analysis is employment growth of a job in a city from 2007 to 2016 ($Growth_{i,c}$). The independent variables of interest

(*SimilarToHighRisk* and *ComplementaryToHighRisk*) are in levels of 2007. And the control variables include factors and trends since 2005, as follows below.

As previously discussed, relatedness between jobs seems to promote employment growth and diversification in cities (Farinha et al., 2019; Muneeppeerakul et al., 2013; Shuttters et al., 2018). Therefore, although this paper focus on the “diffusion” effects of relatedness, the “magnet” effects should still be controlled for. Accordingly, I compute the geographical relatedness of a job to all job specializations of a city in 2007 (*GeoRelated*). As an overall measure of relatedness, it combines all dimensions of relatedness that bring jobs together in a city (Farinha et al., 2019).

City size (*CitySize*) and job size (*JobSize*) in terms of employment levels are also included, in log levels of 2007, as both are expected to affect employment growth (Chen et al., 2019; Frank et al., 2018). Moreover, I control for the local dominance of a job in a city, with a dummy variable =1 if job i was a job specialization of city c in 2007, and 0 otherwise (*RCA*). I also account for major labour demand trends prior to 2007 that might have conditioned the effects from both the Great Recession and automation in subsequent years. More concretely, I add employment growth from 2005 to 2007 at the city level (*CityEmpTrend*), at the job level (*JobEmpTrend*), and, for local trends, at the job-city level (*EmpTrend*). Also, the Great Recession had its own diffusion dynamics, particularly strong between 2008 and 2010, with some jobs in certain cities recovering quicker than others, thus affecting the geography of jobs in the US (Beyers, 2013) beyond the impacts of automation. Therefore, I control for a job’s local resilience capacity during the worse years of the crisis, i.e., between 2007 and 2010, using two variables. Employment growth (*EmpResilienceGR*) accounts for the capacity to maintain employment levels regardless of ongoing adaptation and structural changes. And geographical relatedness growth (*StructuralResilienceGR*), as the Great Recession may also have affected a job’s level of relatedness in a city, which has been shown to strenghten its future capacity grow in terms of employment (Muneeppeerakul et al., 2013).

Table 1. Descriptive statistics

| | N | Mean | St. Dev. | Min | Pctl(25) | Pctl(75) | Max |
|------------------------|---------|-----------|-----------|------|----------|----------|-----------|
| SimilarToHighAR | 125,183 | 1.9 | 1.2 | 0.0 | 1.1 | 2.3 | 18.4 |
| ComplementaryToHighAR | 125,183 | 2.0 | 1.4 | 0.0 | 1.1 | 2.7 | 17.0 |
| GeoRelated | 125,183 | 35.7 | 8.8 | 4.7 | 29.6 | 42.0 | 100.0 |
| RCA | 125,183 | 0.5 | 0.5 | 0 | 0 | 1 | 1 |
| EmpTrend | 125,161 | 0.1 | 0.5 | -1.0 | -0.2 | 0.2 | 9.8 |
| EmpResilienceGR | 122,451 | -0.02 | 0.6 | -1.0 | -0.3 | 0.2 | 9.8 |
| StructuralResilienceGR | 125,183 | 0.01 | 0.1 | -0.6 | -0.04 | 0.1 | 1.3 |
| CitySize | 125,183 | 426,125.1 | 661,648.0 | 7915 | 66,250 | 507,790 | 4,814,970 |
| CityEmpTrend | 125,183 | 0.02 | 0.04 | -0.2 | -0.005 | 0.04 | 0.2 |
| JobSize | 125,183 | 287,570.5 | 467,646.2 | 90 | 50,500 | 294,660 | 3,843,040 |
| JobEmpTrend | 125,183 | 0.03 | 0.1 | -0.6 | -0.02 | 0.1 | 0.9 |
| JobComplexity | 125,183 | 19.2 | 13.6 | 0.0 | 7.8 | 27.7 | 100.0 |
| JobAutRisk | 125,183 | 3.2 | 1.2 | 1 | 2 | 4 | 5 |
| JobZone | 125,183 | 3.0 | 1.1 | 1 | 2 | 4 | 5 |

Finally, I control for specific characteristics of jobs that are associated with current labour demand trends (in 2007 levels). First, automation impacts are, of course, and to a certain extent, conditioned by the technical risk of automating tasks within a job (*JobAutRisk*). Second, there seems to be a non-linear (U-shaped) relationship between level of skills within a job (*JobSkills*) and automation impacts, commonly referred as skills polarization (Autor & Dorn, 2013; Goos et al., 2014; Jaimovich and Siu, 2012). Third, jobs that rely the most on complex skills currently show an increasing demand for labour (Davies & Maré, 2019; Moretti, 2012). Therefore, I add jobs' complexity (*JobComplexity*), computed with the method of reflexions (Hidalgo & Hausmann, 2009) adapted to jobs in cities (Farinha et al., 2019), where a highly complex job tends to be found in few cities (low Job ubiquity) that are very diverse (high City diversity). The final data includes 733 job-classes in 389 cities. Table 1 below presents variable's descriptive statistics.

As expected, *JobAutRisk* is positively yet weakly correlated with employment growth at the job-class level, possibly due to local dynamics that also affect technology adoption and employment levels, as this paper aims to demonstrate (see correlation matrix in Appendix 1).

4.2. Job-City employment growth under automation

In this section, I test the hypothesis *H1* and *H2* by regressing $Growth_{j,c}$ on *SimilarToHighRisk* and *ComplementaryToHighRisk*, plus controls, as follows:

$$\begin{aligned}
 Growth_{j,c} = & \beta_1 SimilarToHighRisk_{j,c} + \beta_2 ComplementaryToHighRisk_{j,c} + \\
 & + \beta_3 GeoRelDen_{j,c} + \beta_4 RCA_{j,c} + \beta_5 EmpTrend_{j,c} + \\
 & + \beta_6 EmpResilienceGR_{j,c} + \beta_7 StructuralResilienceGR_{j,c} + \\
 & + \beta_8 \ln(JobSize)_j + \beta_9 JobEmpTrend_j + \beta_{10} JobComplexity_j + \\
 & + \beta_{11} dyJobSkills2_j + \beta_{12} dyJobSkills3_j + \beta_{13} dyJobSkills4_j + \beta_{14} dyJobSkills5_j + \\
 & + \beta_{15} dyJobAutRisk2_j + \beta_{16} dyJobAutRisk3_j + \beta_{17} dyJobAutRisk4_j + \beta_{18} dyJobAutRisk5_j + \\
 & + \delta_c + \varepsilon_{j,c}
 \end{aligned}$$

where ε is the error term, and δ fixed effects for cities, which comprises all invariant factors that characterise each city economic context. In such model specification, observable variables at the city level, such as cities' 2007 employment levels (*CitySize*) and growth trends (*CityGrowth*), are controlled for as fixed effects, rather than as independent variables. Whereas observable

variables at the job level (*JobAutRisk*, *JobSkills*, *JobComplexity*, etc) are independent variables, so to allow its isolated analysis. Alternative model specifications are presented in Appendix 3. Finally, the categorical variables *JobAutRisk* and *JobSkills* are included as dummies (from 2 to 5), in reference to *JobAutRisk*=1 (high-risk) and *JobSkills*=1 (low-skills), respectively.

Table 2 below presents the results for the econometric analysis in four model specifications. Model (1) contains only control variables, all of them statistically significant, except the dummies *dyJobSkills5* and *JobAutRisk2*. Together with the other dummies, they seem to confirm the current trend of skills polarization referred above (U-shaped relationship between skills level and employment under automation). More concretely, while the coefficients for the *JobAutRisk* dummies (in relation to high-risk-jobs) show somewhat linear effects on employment growth (starting from no significant differences between high and medium-high risk, to major differences between high and low risk), the coefficients for *JobSkills* (in relation to low-skills) are non-linear, with negative and stronger coefficients for medium-skills. Also, note how the national and local employment trends before the recession (*JobEmpTrend* and *EmpTrend*), and the local growth and structural change during recession (*EmpResilienceGR* and *GeoRelResilienceGR*) do not exclude each other in terms of statistical significance. Their strong coefficients seem to confirm distinct growth dynamics between national and local, before and during the Great Recession. Finally, geographical relatedness (*GeoRel*) is statistically significant in all five models. Independently of which variables of interest enter the model, “magnet” effects keep relevant along with “diffusion” effects of relatedness. Models (2) and (3) add the variables of interest, one at time. Model (4) adds them together. Model (5) has scaled variables of interest instead, for comparison of coefficients.

Table 2. The impact of automation in city-job employment growth

| <i>Dependent variable:</i> | | | | | |
|----------------------------|--------------------------------|----------------------|----------------------|----------------------|----------------------|
| | EmpGrowth 2007-2016 (%) | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| SimilarToHighAR | | 0.0003 (0.002) | | -0.013*** (0.002) | -0.014*** (0.003) |
| ComplementaryToHighAR | | | 0.031*** (0.002) | 0.034*** (0.002) | 0.043*** (0.002) |
| GeoRelated | 0.017*** (0.001) | 0.017*** (0.001) | 0.017*** (0.001) | 0.017*** (0.001) | 0.203*** (0.007) |
| RCA | -0.189*** (0.005) | -0.189*** (0.005) | -0.193*** (0.005) | -0.193*** (0.005) | -0.193*** (0.005) |
| EmpTrend | -0.110*** (0.004) | -0.110*** (0.004) | -0.109*** (0.004) | -0.109*** (0.004) | -0.109*** (0.004) |
| EmpResilienceGR | 0.525*** (0.004) | 0.525*** (0.004) | 0.522*** (0.004) | 0.523*** (0.004) | 0.523*** (0.004) |
| GeoRelResilienceGR | 0.599*** (0.043) | 0.599*** (0.043) | 0.604*** (0.043) | 0.598*** (0.043) | 0.598*** (0.043) |
| ln(JobSize) | 0.036*** (0.002) | 0.036*** (0.002) | 0.028*** (0.002) | 0.029*** (0.002) | 0.029*** (0.002) |
| JobEmpTrend | 0.931*** (0.025) | 0.931*** (0.025) | 0.925*** (0.025) | 0.926*** (0.025) | 0.926*** (0.025) |
| JobComplexity | 0.003*** (0.0003) | 0.003*** (0.0003) | 0.002*** (0.0003) | 0.002*** (0.0003) | 0.002*** (0.0003) |
| dyJobSkills2 | -0.085*** (0.009) | -0.084*** (0.009) | -0.067*** (0.009) | -0.073*** (0.009) | -0.073*** (0.009) |
| dyJobSkills3 | -0.069*** (0.009) | -0.069*** (0.009) | -0.054*** (0.009) | -0.060*** (0.009) | -0.060*** (0.009) |

Table 2. The impact of automation in city-job employment growth (Continued)

| <i>Dependent variable:</i> | | | | | |
|--------------------------------|----------------------|----------------------|---------------------|---------------------|---------------------|
| EmpGrowth 2007-2016 (%) | | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| dyJobSkills4 | (0.010) -0.043*** | (0.010) -0.043*** | (0.010) -0.015 | (0.010) -0.025** | (0.010) -0.025** |
| dyJobSkills5 | (0.010) -0.012 | (0.011) -0.012 | (0.011) 0.016 | (0.011) 0.006 | (0.011) 0.006 |
| dyJobAutRisk2 | (0.012) 0.013 | (0.012) 0.013 | (0.012) 0.020** | (0.012) 0.017* | (0.012) 0.017* |
| dyJobAutRisk3 | (0.009) 0.072*** | (0.009) 0.072*** | (0.009) 0.088*** | (0.009) 0.080*** | (0.009) 0.080*** |
| dyJobAutRisk4 | (0.009) 0.149*** | (0.010) 0.149*** | (0.009) 0.164*** | (0.010) 0.152*** | (0.010) 0.152*** |
| dyJobAutRisk5 | (0.009) 0.171*** | (0.009) 0.171*** | (0.009) 0.191*** | (0.009) 0.179*** | (0.009) 0.179*** |
| Observations | (0.011) 122,356 | (0.011) 122,356 | (0.011) 122,356 | (0.011) 122,356 | (0.011) 122,356 |
| R ² | 0.267 | 0.267 | 0.269 | 0.269 | 0.269 |
| Adjusted R ² | 0.265 | 0.265 | 0.267 | 0.267 | 0.267 |
| Residual Std. Error | 0.721 (df = 121951) | 0.721 (df = 121950) | 0.720 (df = 121950) | 0.720 (df = 121949) | 0.720 (df = 121949) |

Note: *p<0.1; **p<0.05; ***p<0.01

In the main Model (4), *SimilarToHighRisk* shows a statistically significant and negative coefficient (-0.014), which seems to confirm that, when the similarity of a job to local high-risk jobs increases by, say, 10 percentage points, its local employment growth decreases (-14%). And the reverse for *ComplementaryToHighRisk*, which shows a statistically significant, positive, and even stronger coefficient (0.034), thus, increasing employment (34%). This is independent of having low or high technical risk of automation, as dummies for *JobAutRisk* are included. An illustrative example taken from the data, surgeons (*JobAutRisk*=5) were less *SimilarToHighRisk* and more *ComplementaryToHighRisk* in New York than in Portland in 2007. In line with the above results, surgeons had greater employment growth in New York (0.25) than in Portland, where it even decreased (-0.75), perhaps due to less replacement needs. In Model (5), we see that effects of productivity gains seem to be more than three times stronger than substitution effects.

Moreover, *SimilarToHighRisk* and *ComplementaryToHighRisk* show to be weaker or even losing statistical significance when including one and not the other (Models 2 and 3). This is also expected, since these two diffusers of impacts have opposite effects on employment and yet may occur in simultaneous between two jobs for which the automation impacts are mixed. For instance, paralegals (number 7 of previous Figure 1) are complementary and similar to the high-risk job of law examiners (8). Therefore, when AI automates tasks within the latter job, the former may have its productivity increased while becoming easier to automate too. Finally, note how the variables of interest seem to affect employment at the job-city level even when controlling for *RCA* and city fixed effects (i.e., beyond cities specific portfolio of job specializations, commonly used in previous studies to extrapolate jobs' risk to cities' risk of automation). Which seems to confirm that, indeed, the "diffusion" effects of relatedness should be accounted for when estimating impacts from automation, both for jobs and for cities.

4.3. Robustness analysis

Some methodological considerations were taken when building the main model (4). For instance, regarding the chosen index of technical risk of automation (*JobAutRisk*). Although from Frey & Osborne (2013) to Atkinson and ITIF (2017) many indexes have been built, criticised, and improved, they keep somewhat redundant in its essence, i.e., based on experts' opinion on the bottlenecks and technical potential of AI. As this paper goes on step further to capture

local diffusion of impacts, results should be robust to which index of technical risk it departs from. Accordingly, robustness checks using alternative indexes for *JobAutRisk* are strongly aligned with the main results. See Appendix 4 for robustness analysis with Frey & Osborne (2013)'s index.

I run additional robustness analysis for alternative model specifications regarding fixed effects (Appendix 3), period of analysis (Appendix 2, where I further add the particular “diffusion” effects of the Great Recession, i.e., relatedness to jobs of high-recession-risk), and stratified results for different skill levels and automation risk (Appendix 5). All these robustness exercises confirm the main analysis in this paper. In particular, the negative effects of *SimilarToHighRisk* become even stronger within low-skill and high-risk groups (as discussed before, skills of high risk tend to concentrate in labour pools of low-skills jobs, leaving them not only with longer adaptation paths towards low-risk skills, but also more susceptible to diffusion of technology through similarities). The positive effects of *ComplementaryToHighRisk* also become stronger, but for high-skills and low-risk instead (lower concentration of similarities to high-risk-jobs, thus negative effects spread less, and positive ones spread more given higher potential for human-computer-collaboration).

Still, as discussed before, labour pools with high concentration of low skills (or high skills) may contain some heterogeneity regarding automation risk. And jobs' levels of relatedness to high-risk jobs also differ among cities, as each city has its own structure of jobs. For instance, taxi drivers in Napa Valley and taxi drivers in Boston must have different adaptation paths towards less automatable skills (as for paralegals, or credit analysts, or any other job-class). This means that, based on the technical risk of automation and the local relatedness to high-risk jobs, opportunities can be found to prevent job losses and increase productivity of each job-class in each city.

5. CONCLUDING REMARKS

In this paper, I empirically investigate how impacts from automation spread between jobs in US cities. The results show that job losses/gains can be exacerbated by the local employment structure, beyond what jobs' technical risk of automation alone would determine. More concretely, controlling for the latter, a job in a city seems to show (i) higher employment growth when having

more complementarities to local high-risk jobs, and (ii) lower employment growth when having more similarities to local high-risk jobs.

The results agree with existing case studies and theory explaining how automation can benefit jobs that are complementary, and not similar, to neighbouring high-risk jobs (Autor, 2015). Moreover, it confirms that, under major economic events, impacts may spread unevenly between cities, as some existing relatedness links are more prone to transmit effects than others. Thus, beyond how much related to local capabilities, it matters to which type of local capabilities one is related to.

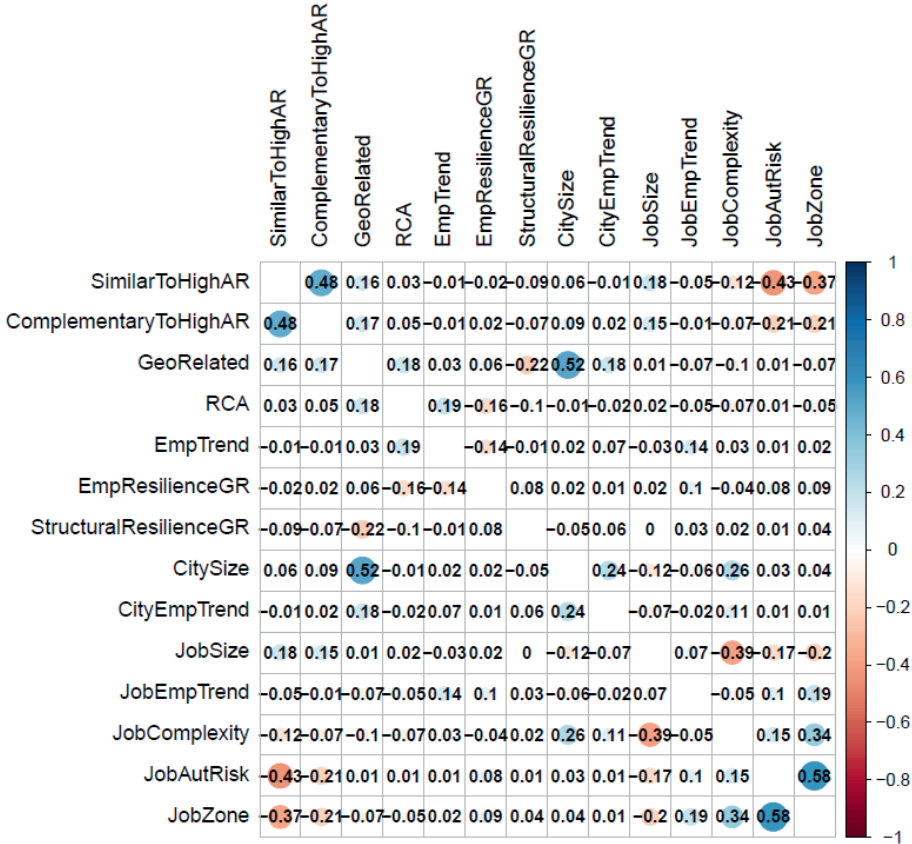
This paper also opens new research avenues that need further investigation. First, diffusion effects should be incorporated and tested in a new automation risk index that goes beyond technical risk. As each city has its own labour structure, such would result in place-based rankings. Second, to which extent the results hold for different economic contexts, such as developing countries, and for more granular levels of analysis. Other levels of analysis, such as job-industry-city, firm level, or case studies, can more directly capture technology adoption in firms and reveal specific industrial dynamics. A third avenue of research regards the interaction with other strong events that have distinct diffusion mechanisms, as they might either exacerbate or smooth the diffusion of impacts from automation. For instance, the current global health pandemic, Covid19, is pushing many jobs to adopt new digital technologies while destroying the ones that cannot go online (Lu, 2020), including jobs of low automation risk, such as hairdressers, dentists, etc. Forth, this paper asks for a new theoretical formulation of how labour systems evolve under technological change.

Finally, we might be able to prevent great damage from automation by redesigning jobs and reengineering business processes to meet symbiotic relationships between technology and labour (Brynjolfsson & McAfee, 2014; Brynjolfsson et al., 2018; Rio-Chanona et al., 2019; Nedelkoska & Quintini, 2018). As shown in this paper, such “right” linkages would be local complementarities to high-risk-jobs, whereas the “harmful” linkages would be local similarities to high-risk-jobs. Therefore, for cities with a large labour pool of high-risk skills, jobs outside that labour pool that collaborate with jobs within it will likely benefit from automation. But workers within are in greater jeopardy (than in cities where such labour pool is smaller) and urge to adapt skills and/or find a symbiotic relationship with the new technologies. Regional policies can use the

linkages between local capabilities (Balland et al., 2019) to neutralise negative effects and promote the spread of positive ones. Especially needed where similarities to high-risk-jobs out rule complementarities to high-risk-jobs. As a demonstration, based on the results of this paper, I provide a novel (and very preliminary) index of automation that accounts for the local diffusion of impacts in each US city (available in tfarina.wixsite.com/tfarinha).

APPENDIX 1

Correlation Matrix



APPENDIX 2

Diffusion effects of the great recession

The current technological revolution has been operating at least since the 90s, i.e., substituting human tasks per robots before, during, and after the last recessionary shock (Acemoglu & Restrepo, 2017). More recently, the Great Recession started spreading in the economy and labour markets in 2008 (Xue & Larson, 2015), pushed by the American financial shock of 2007. It massively destroyed jobs until 2010 (the worse year of the crisis) and only in 2014 the economy seemed to have returned to pre-crisis employment levels. These two (totally different) events are contemporary to each other during the period

2008-2014, and their spreading effects might have overlapped each other²². For instance, between 2007-2010, cashiers and chief executives experienced similar employment decrease while having opposite technical risk, *JobAutRisk*=1 and *JobAutRisk*=5 respectively. Or, cashiers and court clerks have the same high automation risk (#1) and yet they performed opposite in terms of employment growth distribution (cashiers grew at -6%, court clerks at 16%). Accordingly, the correlation between national employment growth during the past recession (2007 to 2010) and the Automation Risk is rather low (0.11).

Model (1) in Table 8 below repeats the analysis in main model except that employment growth between 2007 and 2010 (*EmpResilienceGR*) is now the dependent variable (Cappelli et al., 2018). Results show that, under the Great Recession, the recessionary employment dynamics seem to have outperformed the effects of automation. Not only the “diffusion” effects of automation loose strenght, *SimilarToHighRisk* coefficient even becomes positive. Moreover, while jobs of low *JobAutRisk* seem to have been less harmed, in medium-risk employment growth was worse than in high-risk. This is expected as some jobs of high automation risk have rigid labour demand, which seems to rule employment growth during a major recession. Model (2), in same table, repeats the analysis, while adding the particular “diffusion” effects of the Great Recession, i.e., similarity and complementarity to jobs most harmed, in terms of employment, under the recession (*SimilarToHighRiskGR* and *ComplementaryToHighRiskGR*). Again, results seem to confirm how the “diffusion” effects identified in the main model seem to be specific of the current wave of automation, and less effective under the past Great Recession, when the recession’s “diffusion” effects must have outcome the automation ones (being similar or complementary to jobs more severely affected by the recession had negative impacts on employment under the past recession).

22 The past recession might have intensified a change of skills in the workforce (Schumpeterian creative destruction process), with automation selectively helping the recovery of industries better aligned with the skills of the future and labour saving through substitution per automated solutions (Autor, 2015). Or it might have delayed the technological transition (Charles et al., 2016) where, for instance, the automation negative effects on manufacturing would have appeared earlier without the large and temporary increases in housing demand pre-crisis.

Table 4. Models for city-job employment growth under the Great Recession

| | <i>Dependent variable:</i> | |
|---------------------------|--------------------------------|----------------------|
| | EmpGrowth 2007-2010 (%) | |
| | (1) | (2) |
| SimilarToHighRisk | 0.005** (0.002) | -0.002 (0.002) |
| ComplementaryToHighRisk | 0.016*** (0.001) | 0.012*** (0.001) |
| SimilarToHighRiskGR | | -0.013*** (0.001) |
| ComplementaryToHighRiskGR | | -0.009*** (0.001) |
| GeoRelated | 0.019*** (0.001) | 0.019*** (0.001) |
| RCA | -0.219*** (0.004) | -0.214*** (0.004) |
| EmpTrend | -0.149*** (0.003) | -0.149*** (0.003) |
| GeoRelatedTrend | 0.302*** (0.033) | 0.296*** (0.033) |
| ln(JobSize) | 0.015*** (0.002) | 0.011*** (0.002) |
| JobEmpTrend | 0.601*** (0.020) | 0.596*** (0.020) |
| JobComplexity | 0.002*** (0.0002) | 0.002*** (0.0002) |
| dyJobSkills2 | -0.032*** (0.007) | -0.029*** (0.007) |
| dyJobSkills3 | 0.019** (0.008) | -0.001 (0.008) |
| dyJobSkills4 | 0.065*** (0.009) | 0.031*** (0.009) |
| dyJobSkills5 | 0.117*** (0.010) | 0.068*** (0.010) |
| dyJobAutRisk2 | -0.029*** (0.007) | -0.011 (0.007) |
| dyJobAutRisk3 | -0.021*** (0.008) | -0.006 (0.008) |
| dyJobAutRisk4 | 0.046*** (0.008) | 0.044*** (0.008) |
| dyJobAutRisk5 | 0.058*** (0.009) | 0.054*** (0.009) |
| Observations | 122,429 | 122,429 |
| R ² | 0.091 | 0.094 |
| Adjusted R ² | 0.088 | 0.091 |
| Residual Std. Error | 0.582 (df = 122023) | 0.581 (df = 122021) |

Note: *p<0.1; **p<0.05; ***p<0.01

APPENDIX 3

Alternative model specifications regarding fixed effects

For robustness analysis, I test alternative model specifications regarding fixed effects. In model (1) of Table 9 below, I run a one-way fixed effects for jobs. Since local economy dynamics affecting employment are vast, observable variables at the city level must be added. But they are also usually highly correlated, which can compromise the analysis if jointly added to the model. Therefore, I include only *CitySize* (ln) and *CitySizeGrowth* as regressors, as these are not highly correlated to each other, and are highly correlated to other city characteristics left out (such as city complexity, etc.). In model (2) of the same table, I run the a two-way-fixed effects version, for cities and for jobs. In result, all invariant specific variables for cities and for jobs become *NA* in the model (“the matrix is either rank-deficient or indefinite” compromising the analysis), and we loose further variability (and statistical significance of *SimilarToHighRisk*) yet an increase in the R^2 . Finally, in model (3), I test the model specification of no fixed effects, adding all previous variables together in the regression. Results seem to indicate the importance of city fixed effects, which are mostly unobservable, while keeping the variables of interest at the job-city and job level.

Table 5. Models with alternative model specification regarding fixed effects

| | <i>Dependent variable:</i> | | |
|-----------------------|--------------------------------|----------------------|----------------------|
| | EmpGrowth 2007-2016 (%) | | |
| | (1) | (2) | (3) |
| SimilarToHighAR | -0.013*** (0.003) | 0.003 (0.004) | -0.021*** (0.002) |
| ComplementaryToHighAR | 0.015*** (0.003) | 0.019*** (0.003) | 0.040*** (0.002) |
| GeoRelated | 0.077*** (0.005) | 0.242*** (0.008) | 0.074*** (0.005) |
| RCA | -0.175*** (0.004) | -0.204*** (0.004) | -0.169*** (0.005) |
| EmpTrend | -0.125*** (0.004) | -0.127*** (0.004) | -0.108*** (0.004) |
| EmpResilienceGR | 0.485*** (0.004) | 0.473*** (0.003) | 0.535*** (0.004) |
| GeoRelResilienceGR | 0.084*** (0.027) | 0.136*** (0.052) | 0.252*** (0.027) |
| ln(CitySize) | -0.002 | | -0.004 |

Table 5. Models with alternative model specification regarding fixed effects (Continued)

| <i>Dependent variable:</i> | | | |
|--------------------------------|---------------------|---------------------|---------------------|
| EmpGrowth 2007-2016 (%) | | | |
| | (1) | (2) | (3) |
| | (0.003) | | (0.003) |
| CitySizeGrowth | 1.025*** | | 0.981*** |
| | (0.052) | | (0.055) |
| ln(JobSize) | | | 0.032*** |
| | | | (0.002) |
| JobEmpTrend | | | 0.916*** |
| | | | (0.025) |
| JobComplexity | | | -0.0001 |
| | | | (0.0003) |
| dyJobSkills2 | | | -0.075*** |
| | | | (0.009) |
| dyJobSkills3 | | | -0.062*** |
| | | | (0.010) |
| dyJobSkills4 | | | -0.039*** |
| | | | (0.011) |
| dyJobSkills5 | | | 0.003 |
| | | | (0.012) |
| dyJobAutRisk2 | | | 0.015* |
| | | | (0.009) |
| dyJobAutRisk3 | | | 0.071*** |
| | | | (0.010) |
| dyJobAutRisk4 | | | 0.149*** |
| | | | (0.010) |
| dyJobAutRisk5 | | | 0.172*** |
| | | | (0.011) |
| Constant | | | -0.310*** |
| | | | (0.047) |
| Observations | 122,356 | 122,356 | 122,356 |
| R ² | 0.329 | 0.360 | 0.239 |
| Adjusted R ² | 0.325 | 0.354 | 0.238 |
| Residual Std. Error | 0.691 (df = 121614) | 0.676 (df = 121228) | 0.734 (df = 122335) |

Note: *p<0.1; **p<0.05; ***p<0.01

APPENDIX 4

Alternative index for technical risk of automation – Frey and Osborne (2013)

Table 4 below shows the results for the alternative analysis using the Frey & Osborne (2013) index for professions' technical risk of automation (*JobAutRiskF&O*), transformed into five dummies (equivalent to *JobAutRisk* categories from high to low risk, i.e. from 1 to 5). The results are strongly aligned with the main results in this paper. Both indexes capture the technical feasibility to automate tasks within jobs and arrive at similar rankings for professions' automation risk. More importantly, as this paper aims to demonstrate, both indexes are insufficient to evaluate the overall impacts from automation, as it further depends on local spread of impacts, which can be captured by jobs' levels of relatedness to local high-risk-jobs.

Table 6. Models with F&O index (alternative index of jobs' technical risk of automation)

| | <i>Dependent variable: Job-City Employment Growth 2007-2016 (%)</i> | | | |
|-------------------------|---|----------------------|----------------------|----------------------|
| | (1) | (2) | (3) | (4) |
| SimilarToHighRisk | | -0.014*** (0.002) | | -0.019*** (0.002) |
| ComplementaryToHighRisk | | | 0.006*** (0.001) | 0.011*** (0.002) |
| GeoRelated | 0.019*** (0.001) | 0.019*** (0.001) | 0.019*** (0.001) | 0.019*** (0.001) |
| RCA | -0.180*** (0.005) | -0.180*** (0.005) | -0.181*** (0.005) | -0.181*** (0.005) |
| EmpTrend | -0.113*** (0.004) | -0.114*** (0.004) | -0.113*** (0.004) | -0.114*** (0.004) |
| EmpResilienceGR | 0.524*** (0.004) | 0.524*** (0.004) | 0.524*** (0.004) | 0.524*** (0.004) |
| StructuralResilienceGR | 0.747*** (0.045) | 0.737*** (0.045) | 0.754*** (0.045) | 0.746*** (0.045) |
| ln(JobSize) | 0.028*** (0.002) | 0.031*** (0.002) | 0.027*** (0.002) | 0.029*** (0.002) |
| JobEmpTrend | 0.987*** (0.026) | 0.983*** (0.026) | 0.989*** (0.026) | 0.987*** (0.026) |
| JobComplexity | 0.003*** (0.0003) | 0.003*** (0.0003) | 0.002*** (0.0003) | 0.003*** (0.0003) |
| dyJobSkills2 | -0.077*** (0.009) | -0.080*** (0.009) | -0.078*** (0.009) | -0.085*** (0.009) |

Table 6. Models with F&O index (alternative index of jobs' technical risk of automation) (Continued)

| | <i>Dependent variable: Job-City Employment Growth 2007-2016 (%)</i> | | | |
|-------------------------|---|------------------------|------------------------|------------------------|
| | (1) | (2) | (3) | (4) |
| dyJobSkills3 | -0.031*** (0.010) | -0.030*** (0.010) | -0.033*** (0.010) | -0.035*** (0.010) |
| dyJobSkills4 | -0.023** (0.011) | -0.022** (0.011) | -0.026** (0.011) | -0.028*** (0.011) |
| dyJobSkills5 | 0.020 (0.013) | 0.017 (0.013) | 0.020 (0.013) | 0.016 (0.013) |
| dyJobAutRiskF&O2 | 0.063*** (0.009) | 0.054*** (0.010) | 0.069*** (0.010) | 0.063*** (0.010) |
| dyJobAutRiskF&O3 | 0.084*** (0.010) | 0.067*** (0.010) | 0.092*** (0.010) | 0.076*** (0.010) |
| dyJobAutRiskF&O4 | 0.120*** (0.010) | 0.101*** (0.010) | 0.129*** (0.010) | 0.112*** (0.010) |
| dyJobAutRiskF&O5 | 0.174*** (0.011) | 0.151*** (0.011) | 0.183*** (0.011) | 0.161*** (0.011) |
| Observations | 116,876 | 116,876 | 116,876 | 116,876 |
| R ² | 0.270 | 0.270 | 0.270 | 0.271 |
| Adjusted R ² | 0.267 | 0.268 | 0.268 | 0.268 |
| Residual Std. Error | 0.708 (df = 116471) | 0.708 (df = 116470) | 0.708 (df = 116470) | 0.708 (df = 116469) |

Note: *p<0.1; **p<0.05; ***p<0.01

APPENDIX 5

Results for stratified dependent variable

I repeat the analysis of the main model (relatedness variables are scaled to allow comparison), for alternative dependent variables. More specifically, the dependant variable is *EmpGrowth* 2007-2016 (%) for the specific groups: high-skills, low-skills, high-risk, and low-risk jobs (*JobAutRisk*)

Table 7. Results for stratified data

| <i>Dependent variable:</i> | | | | |
|--------------------------------|------------------------------|-----------------------------|-----------------------------|----------------------------|
| EmpGrowth 2007-2016 (%) | | | | |
| | HighSkillJobs (1) | LowSkillJobs (2) | HighRiskJobs (3) | LowRiskJobs (4) |
| SimilarToHighAR | -0.011 (0.008) | -0.025*** (0.004) | -0.012*** (0.003) | 0.008 (0.007) |
| ComplementaryToHighAR | 0.087*** (0.006) | 0.039*** (0.003) | 0.028*** (0.003) | 0.052*** (0.004) |
| GeoRelated | 0.190*** (0.014) | 0.233*** (0.012) | 0.210*** (0.013) | 0.180*** (0.011) |
| RCA | -0.226*** (0.008) | -0.166*** (0.007) | -0.164*** (0.007) | -0.224*** (0.007) |
| EmpTrend | -0.106*** (0.007) | -0.115*** (0.006) | -0.103*** (0.006) | -0.113*** (0.006) |
| EmpResilienceGR | 0.574*** (0.006) | 0.466*** (0.006) | 0.468*** (0.006) | 0.559*** (0.005) |
| GeoRelResilienceGR | 0.332*** (0.076) | 0.587*** (0.070) | 0.336*** (0.067) | 0.908*** (0.067) |
| log(JobSize) | 0.050*** (0.004) | 0.010*** (0.004) | 0.027*** (0.003) | 0.038*** (0.004) |
| JobEmpTrend | 1.161*** (0.040) | 0.370*** (0.041) | 1.234*** (0.046) | 0.614*** (0.036) |
| JobComplexity | 0.005*** (0.0005) | -0.003*** (0.0005) | 0.0005 (0.0005) | 0.004*** (0.0004) |
| JobSkills2 | | -0.096*** (0.009) | -0.069*** (0.010) | -0.160*** (0.031) |
| JobSkills3 | | | -0.051*** (0.012) | -0.169*** (0.030) |
| JobSkills4 | | | -0.046*** (0.016) | -0.126*** (0.031) |
| JobSkills5 | 0.075*** (0.009) | | -0.177*** (0.034) | -0.096*** (0.032) |

Table 7. Results for stratified data (Continued)

| <i>Dependent variable:</i> | | | | |
|--------------------------------|------------------------------|-----------------------------|-----------------------------|----------------------------|
| EmpGrowth 2007-2016 (%) | | | | |
| | HighSkillJobs (1) | LowSkillJobs (2) | HighRiskJobs (3) | LowRiskJobs (4) |
| JobAutRisk2 | -0.063 (0.045) | 0.059*** (0.010) | 0.013 (0.008) | |
| JobAutRisk3 | 0.041 (0.044) | 0.074*** (0.011) | | |
| JobAutRisk4 | 0.179*** (0.044) | 0.135*** (0.012) | | |
| JobAutRisk5 | 0.161*** (0.044) | 1.102*** (0.030) | | 0.034*** (0.007) |
| Observations | 41,103 | 47,339 | 41,899 | 57,954 |
| R ² | 0.293 | 0.265 | 0.260 | 0.268 |
| Adjusted R ² | 0.286 | 0.259 | 0.252 | 0.262 |
| Residual Std. Error | 0.745 (df = 40699) | 0.698 (df = 46935) | 0.670 (df = 41495) | 0.752 (df = 57550) |

Note: *p<0.1; **p<0.05; ***p<0.01

CHAPTER 5

EU Incentives to Business Increased Job Quality In Portuguese Firms: Evidence from Linked Employer-Employee Data and Natural Experiment Conditions



Contribution: Conceptualization, data curation, formal analysis, and writing (original draft, review & editing)

Coauthors: Daniele Bondonio and Ricardo Paes Mamede

Published as: Bondonio, D., Farinha, T., Mamede, R., (currently under review) EU incentives to business increased job quality in Portuguese Firms: Evidence from Linked Employer-Employee Data and Natural Experiment Conditions of EU Cohesion Policy in Portugal, Evaluation Review Special Issue, Sage Journals (manuscript submitted and invited to revise and resubmit)

1. INTRODUCTION

Public support to business investments, especially for Small and Medium Sized firms (SMEs), is a key mechanism through which the EU promotes socioeconomic convergence among European regions. Under the current programming period (2014-2020), the European Commission has allocated around 20% of the European Regional Development Fund (ERDF) – €57 billion – to supporting SME investments in competitiveness, research and innovation, the low-carbon economy and information and communication technologies²³. The total amount of funds dedicated to these domains has been increasing over the years, as regional policy shifts away from investment in hard infrastructure towards business support and innovation (European Commission, 2014).

Job quality – skilled, safe, fairly remunerated, work-life balanced jobs²⁴ – has been shown to increase worker motivation, customer loyalty, and productivity, to lower absenteeism and turnover rates. (Harter et al., 2004; Warr, 1999), and to spur improvements in organizational performance and profitability (Boxall & Macky, 2009; Kaufman, 2010). A number of studies (Clark, 2005; Green, 2006, 2008; Olsen et al., 2010) have shown that work intensity and work stress has increased in several countries and regions during the 1990s and 2000s, with deleterious consequences for individual well-being and social inclusion. Thus, as cited in many EU strategic documents since the Lisbon Agenda in 2000 (European Commission, 2008)²⁵, the creation of quality jobs is a necessary condition for the economic convergence of disadvantaged regions.

Moreover, as we enter a new technological paradigm with Artificial Intelligence and other highly disruptive new technologies, local labour systems must undergo a renewal of skills in the workforce in order to keep workers employed and cope with deep transformations in the workplace. So far, the new technologies seem to have increased unemployment more than job productivity (Acemoglu & Restrepo, 2019), and aggravated job polarization in terms of skills content in professions (Brynjolfsson & McAfee, 2014). Job polarization is associated

23 http://ec.europa.eu/regional_policy/en/policy/themes/sme-competitiveness/

24 OECD Job Quality framework (www.oecd.org/statistics/job-quality.htm)

25 The increased emphasis on job quality is not exclusive to the EU. Several international organisations advocate the implementation of policy action to address the deleterious impacts of low quality employment on individual well-being and social inclusion (International Labour Organization, 2014; World Bank, 2013).

with higher income inequality and also lower quality of jobs, as lower skills are dominant in the workforce and usually low paying and more precarious jobs.

All these findings and current challenges help to explain the growing concern with job quality as an important policy outcome to be considered. According to the International Labour Organization (2014) and the World Bank (2013), public policies can influence the quality of work and employment in several ways. By establishing and enforcing labor market regulations, governments can affect payments, entitlements, working conditions, and job security. Governments can also exert influence on social dialogue, enhancing the chances of achieving higher work standards through collective bargaining and agreements between trade unions and employers' representatives. Public policy interventions can also foster job quality at the firm level, either through employment subsidies, or through enterprise support mechanisms. Moreover, business support policy schemes may generate a demand-side complement of education and training policies aimed at increasing the supply of skilled labor and firm productivity.

In sum, improving job opportunities in disadvantaged regions, both quantitatively and qualitatively, is a main expected impact of regional policies (Bartik, 2012) and one of the main policy goals of the EU Cohesion Policy. Even when not directly established as the objective of a business support scheme²⁶, job quality is both a "by-product" and a "mean to the end" of boosting regional firms' competitiveness, innovation, etc. Conversely, failure to create good quality jobs is expected to jeopardize the development prospects of less-developed regions. On one hand, the expansion of low quality jobs increases the risk of outsourcing production to areas with lower wages. On the other, the limited availability of high-quality jobs fosters the outflow of human capital to the more advanced regions.

Therefore, providing EU policy makers with reliable empirical evidence on the job-quality impact of enterprise support programs implemented as cohesion policy tools assumes crucial importance. Especially considering that policy institutions worldwide are currently redirecting efforts to guide the transition of skills in the workforce towards a future with quality jobs for everyone (OECD, 2018), the capacity to evaluate its impact on job quality gets more relevant than ever.

26 The nature and complexity of the outcome variable itself, job quality, makes it hard to tackle directly in policy intervention at the firm level. Thus, more often set as a policy goal rather than a direct program objective.

But producing such empirical evidence, however, is not an easy task for a number of reasons. First, measuring job quality outcomes requires the rare availability of linked employer-employee microdata merged with complete program activity data. Second, the existence of unobserved additional sources of public support generates potentially serious treatment measurement errors and firms often self-select into the program with no sources of exogenous treatment exclusions (making it hard to control for unobservable differences between the treatment and the comparison group). Finally, it is often difficult to determine the exact place in time in which the treatment occurs and the period in which to measure the observable outcomes related to the supported investment. Due to such difficulties, to the best of our knowledge, no reliable empirical evidence has yet been produced on the job quality impact of enterprise support programs in the EU.

This paper aims at filling the gap in the literature by providing reliable counterfactual impact estimates on the following job-quality outcomes: number of jobs with standard working hours²⁷; number of jobs with high skill requirements²⁸; number of jobs with permanent contracts; and level of hourly remuneration²⁹. The paper is organized as follows.

Section 2 provides a brief survey of the related literature. Section 3 describes in greater detail the policy mechanisms under analysis, the data and some descriptive statistics. Section 4 discusses the identification problem and our preferred estimation model. Section 5 presents the main results and section 6 the sensitivity analysis. Section 7 concludes, discussing the policy implications of our study.

2. PREVIOUS LITERATURE

The literature on impact evaluation of public subsidies in employment growth at the firm-level, is rather diverse, regarding both the geographical and time scopes, and the empirical modelling strategies used in the analysis. Even if considering solely and specifically EU cohesion policy tools.

For instance, Bia & Mattei (2012) applied a *generalized propensity score matching* adjusting for differences in a set of observed pre-treatment variables to assess

27 The standard number of working hours, considered to allow worker's work-life balance, is 40 hours per week.

28 Job that requires academic degree of bachelor or higher, or high level of experience.

29 Median of worker's hourly remuneration in the firm.

the employment impact of financial aid attributed to enterprises located in the Italian region of Piedmont in the early 2000s. Bondonio and Greenbaum (2006 and 2014) implemented a parametric *conditional difference-in-differences* model to estimate the employment impact of EU funded business incentives in northern and central Italy during the late 1990s and early 2000s. Bronzini & de Blasio (2006) also adopted a *difference-in-differences* framework to analyze the effect of similar policies on firms located in southern Italian regions during the same period.

Bernini & Pellegrini (2011) used a combination of matching methods with *difference-in-differences*, to extend the analysis of EU-supported Italian policies to the mid-2000s, while Cerqua & Pellegrini (2014) focused on Southern-Italian regions to reassess the same enterprise support mechanism analyzed in Bronzini & de Blasio (2006) and Bernini & Pellegrini (2011) by means of a *regression discontinuity design*. On the same data, Bondonio and Martini (2019) enlarged the scope of the analysis to all Italian regions and used *matching estimators* and a *discontinuity approach* (using non-subsidized applicant firms as a control group), in combination with a *difference-in-difference* scheme.

In the UK, Criscuolo et al. (2012) studied the effects of the Regional Selective Assistance program from the mid-1980s to the mid-2000s, using a combination of *instrumental variables* and firm-level fixed-effects. In Finland, Koski & Pajarinen (2013) studied the effect of investment support and other subsidies granted to SMEs during the 2000s, applying both an *instrumental variable* approach and *difference-in-differences*. Bondonio and Martini (2012) investigated the combined effects of multiple sources of support to SMEs in Northern Italy with a multiple-treatment *propensity score matching* approach combined with a *difference-in-difference* scheme.

The findings of the aforementioned studies converge to the conclusion that direct support to firm-investments had some positive impact on employment growth at the firm level. And yet, no empirical investigation in such literature is offered to shed light on the possible impacts of the policy interventions on changes in the job quality of the supported firms.

In part because, although an important expected outcome and main policy goal in business incentive programs, their central objectives are usually firm performance outcomes, such as general employment, investments, sales, etc. Thus,

longitudinal outcome variables regarding job quality need to be merged to the complete program activity data. Which leads to a second major reason, i.e., the rare availability of proper data (including controls for unobserved variables and clear time cuts and location of the treatment and observable outcomes). In this paper, we were able to gather such research conditions, and contribute to this literature with empirical evidence on the impact of policy interventions on job quality.

Our analysis focuses on the direct firm-level support schemes offered by a large Operational Program (the “Program for the Modernization of the Economy”, or POE/PRIME) implemented in Portugal over the 2000s and co-funded by the ERDF (Third Community Support Framework). We merged the program activity data from POE/PRIME with the linked employer-employee data collected yearly by the Portuguese Ministry of Employment and Social Solidarity on a compulsory basis. The resulting unique database offers adequate impact identification conditions to investigate the effects of firm-support interventions co-funded by the EU on job quality because of the following three reasons.

First, the POE/PRIME program has been a very prominent EU cohesion policy intervention, with a total budget of nearly €8.3 billion, financed in approximately equal parts by private and public sources (about 70% of the latter came from the ERDF). The POE/PRIME’s direct support to firms accounts for nearly $\frac{3}{4}$ of the total public funds of the program and it was implemented through several policy schemes that represented by far the most prominent source of public aid (in terms of both the amount of funds and number of beneficiaries) for the Portuguese firms in the 2000-6 programming period. These circumstances ensure a large external validity to the analysis and avoid the treatment contamination issues that would derive from the concurring presence of other unobserved policy interventions.

Second, the linked employer-employee administrative data used in the analysis are virtually free of the many attrition bias and measurement error limitations that plague the firm-level balance sheet data that are very often used in the evaluation of firm-support programs.

Third, in our analysis we exploit a unique period of the POE/PRIME program implementation that occurred within the years 2003-6, in which the operational rules for assigning the support deviated from the customary practice in Portugal. Within such period, the program was administered through a series of

different calls for applications in which the cut-offs in the rankings of applicants were based on exogenous application scores that did not correlate with any of the pre-treatment trends of the job-quality outcomes of the firms. This circumstance yields a natural experiment scenario in which the cut-offs in the application scores determine a pseudo-experimental treatment assignment process that enhance the impact identification conditions of the analysis.

Our preferred estimation strategy is aimed to fully exploit these desirable natural experiment conditions of the 2003-6 waves of calls for applications of the program. This is done by means of a dynamic exact matching model, combined with a difference-in-differences scheme, that is adopted separately for each of the successive yearly cohorts of applicant firms. This estimation approach mimics the features of a pre-test-post-test block-randomization design in a dynamic context and it is preferred to a mere difference in difference design because of the small sample size of applicant firms, within each wave of calls for applications, that could generate some remaining unbalancing in crucial observable controls between treated and non-treated applicants.

Compared to a one-dimensional matching based on a propensity score (PS), our exact matching model is preferable because the observable control variables capture a small number of crucial firm characteristics that can have an equally-relevant impact on the future determinations of the job-quality outcomes. In these circumstances, as well established in the literature (Ho et al., 2007; King et al., 2011), PS matching would entail some admissible degree of mismatching on single controls that may lead to some bias in the impact estimates. Finally, compared to standard (static) matching approaches, our dynamic specification avoids reliance on control variables that are fixed at an initial point in time for all subsequent annual cohorts of treatments, enabling the analysis to estimate more reliably the counterfactual trend of all treated firms.

3. POE/PRIME PROGRAM, DATA AND DESCRIPTIVE STATISTICS

The POE/PRIME³⁰ program (the main Operational Program of the Third Community Support Framework in Portugal) that we exploit in our analysis is a

³⁰ The Program was originally called POE (Operational Program for the Economy) and was lately re-named PRIME (Program for the Modernisation of the Economy).

large and comprehensive policy intervention that was co-funded by EU funds in the programming period 2000-6. The program had a budget of nearly 8.3 billion €, financed in approximately equal parts by private and public sources (about 70% of the latter came from the ERDF).

The POE/PRIME's direct support to firms included several schemes targeting specific types of investment and/or types of firm. These schemes differed somehow in technical-administrative aspects such as the minimum size and type of admissible investment³¹. In practice, however, all POE/PRIME firm-level support schemes were subjected to similar sets of program application rules and all shared the same overall policy goals of enhancing the competitiveness of Portuguese SMEs by means of investments in innovation and internationalization. These circumstances greatly diminish the policy relevance of estimating heterogeneous impacts based on the different types of program schemes and are conducive to pooling together the program activity data for the sake of improving the statistical efficiency and external validity of the analysis.

Even though the POE/PRIME was not the only firm support program in Portugal during the period under analysis, it was by far the most relevant one, both in terms of public funds and number of beneficiaries. Less than 1% of the firms supported under these schemes received some kind of direct support from other programs. This scenario virtually eliminates the treatment contamination threats that may derive from the concurring presence of other unobserved public subsidies.

The POE/PRIME direct support schemes were administered through a process that sorted the applicants into a treated and a non-treated group based on an application score described in greater detail in the next section of the paper. The presence of such non-treated applicants enables us to compare beneficiaries with non-treated firms that shared the same desire to invest. This is crucial to control for many important unobservable characteristics of the firms such as the position in their investment cycle, risk aversion, and management style.

The POE/PRIME program activity data at our disposal cover all the investment projects submitted by the applicant firms, with information on: application

31 Schemes supporting investments in physical capital, however, were by far predominant, representing nearly 90% of the projects and 80% of the funds, while the remaining types of admissible investments were related mostly to R&D activities, IPR and energy efficiency.

date, total value of the investment project, the amount of the program support and the timing in which the supported investment took place (if applicable). This information allows us to sort the beneficiaries into different cohorts based on the exact dates of their application and investment project, virtually eliminating all the measurement errors that would arise when instead no data were available to determine the correct moment in which to measure the job-quality outcomes and the pre-intervention control variables.

The linked-employer-employee-data (LEED) used in the analysis stem from an administrative database maintained by the Portuguese Ministry of Employment and Social Solidarity. This database includes information related to both the characteristics of firms (total employment, industry classification, geographic location, legal status, ownership, number of plants, etc.) and their employees (sex, date of birth, educational background, professional category, type of contract, etc.). All information included in the LEED database is compulsorily submitted on a yearly basis by each and every firm operating in Portugal. For this reason, the database is virtually free of the many attrition bias and measurement error limitations that plague the employment information extracted from the balance sheet databases (e.g. the Bureau Van Dick's Amadeus) often used in the evaluation of firm-support programs.

Using the firms' National Tax Code as the common firm-identifier, we merged the LEED data with the POE/PRIME program activity data. Although the POE/PRIME program was active between 2000 and 2009, the focus of our analysis is on the years 2003-6 because of two reasons. First, for such period we can exploit the favourable natural experiment conditions that we describe in the next section. Second, for the 2003-6 years, we had the possibility to measure, both in the LEED and the program activity data, an adequate three-year observation period before and after each of the yearly cohorts of treatments.

During the 2003-6 period, a total of 1,493 firms applied for the program incentives (Table 1). Of these firms, 634 were excluded from the program incentives, based on cut-offs in the rankings of the application score, while the remaining 859 applicants were treated by the program. Sorted by year of application, 21% (315 out of 1,493) of these firms applied in the 2003, 11% and 8% of the firms in the years 2004 and 2005, respectively, while 60% of the firms applied in the year 2006.

Table 1 – Number of treated and non-treated applicant firms by year of application (t)

| <i>Year (t)</i> | <i>Treated applicants</i> | <i>Non-treated applicants</i> | <i>Total</i> |
|-----------------|---------------------------|-------------------------------|--------------|
| 2003 | 187 | 128 | 315 |
| 2004 | 58 | 104 | 162 |
| 2005 | 99 | 27 | 126 |
| 2006 | 515 | 375 | 890 |
| Total | 859 | 634 | 1,493 |

The descriptive statistics of the pre-post treatment trends of the job quality outcomes of the applicant firms are summarized in Table 2. These figures (pooled across the four different yearly cohorts of applicants) compare the changes of the outcome variables of the treated and non-treated applicants from one-year (t-1) before the application to three years (t+3) after the application. The 859 treated applicants that we consider in the analysis show a positive pre-post treatment change that ranges from +2.9 jobs per firm for standard working hours conditions and skilled occupations to +2.3 jobs for occupations with permanent contracts; while their change for the average hourly remuneration is about +2 euro. In contrast, for the non-treated applicants the changes for the same periods show a negative variation (-1.7 jobs) for the occupations with standard working hours and a zero or near-zero variation for the jobs with permanent contracts, skilled occupations, and average hourly remuneration.

Table 2 – Average change (per firm) of the outcome variables between the pre (t-1) and post (t+3) application times

| <i>Outcome variable Y</i> | $\Delta Y (t+3) - (t-1)$ | |
|--|---------------------------|-------------------------------|
| | <i>Treated applicants</i> | <i>Non treated applicants</i> |
| Jobs with standard working hours (1=1 job) | +2.9 | -1.7 |
| Skilled jobs (1=1 job) | +2.9 | 0.0 |
| Jobs with permanent contracts (1=1 job) | +2.3 | -0.3 |
| Average hourly remuneration (1= 1 €) | +2.0 | 0.3 |

4. EMPIRICAL STRATEGY

Our empirical strategy is designed to exploit the presence of the different multiple consecutive calls for applications under which the POE/PRIME program was administered. In these calls, applicants were ranked based on an application score, and the treatment assignment was determined by cut-off points in the ranking of the application scores.

During the lifetime of POE/PRIME, the operational rules for assigning the application score varied along the different calls for applications. In our period of observation, 2003-6, the score was based solely on two factors: a) an indicator that measures the present value (in percentage of the total investment value) of the sum of the projected taxable profits, the additional labor costs and the interests paid to financial institutions for the part of the investment not covered by the program funds; b) an indicator of financial autonomy, in terms of the ratio between the firm's equity and its total assets, intended to capture the projected risk of failure of the firm. The latter worked as an exclusion criterion: projects submitted by firms' whose ratio of financial autonomy was below a certain threshold (which varied between 20% and 25% across all calls) were not considered for further analysis.³²

These operational rules that were in place in the 2003-6 period were established in 2002 by the recently elected center-right government led by José Manuel Barroso³³ in an effort to simplify the treatment assignment procedure and reduce the administrative costs and application burdens. The Barroso Government's rules sharply deviated from the previous treatment-assignment procedures, that were based on multiple criteria and assigned a high weight to some qualitative features of the submitted projects, such as the contribution of the project to national and regional development goals also in terms of innovation intensity, export orientation, and quality of jobs to be generated by the projects. These multiple-criteria rules were gradually re-established after 2005³⁴, when a new government took office.

The 2003-6 period is ideal for the analysis also because during those years, likewise in other phases of the POE/PRIME program, all the information on the investment project submitted by the applicant firms were predominantly prepared by external consultants³⁵. The application packages prepared by the consultants included figures that were determined by exploiting all the degrees of discretion allowed by the program bookkeeping rules in order to increase the likelihood of success in receiving the support. Thus, the resulting application scores were significantly affected by such degrees of discretion used by the

32 The relevant pieces of legislation are Portaria n.º 865-A/2002 (published on the 22nd of July) and Portaria n.º 879-A/2002 (published on the 25th of July).

33 Two year later, by the end of 2004, Barroso became President of the European Commission.

34 The relevant pieces of legislation are Portaria n.º 88-D/2006 (published on the 24th of January) and Portaria n.º 130-A/2006 (published on the 14th of February).

35 In many cases, consulting firms were also the main sources of information for the firms on the availability of the support and on the advantages to apply for it (see Augusto Mateus & Associados, 2005).

external consultants, further weakening any endogeneity of the treatment assignment process with regard to the pre-application trends of the job quality outcomes of the application firms.

For these reasons, during the 2003-6 period, the cut-offs of the application scores that established the treatment status of the applicants did not sort out the firms based on characteristics that correlate with their pre-application trends of the job-quality outcomes. This is evident from the summary statistics of Table 3 that compares the pre-application trends of the job quality variables (pooled across the four different yearly waves of call for applications) between the treated and non-treated applicant firms. For all the job-quality outcomes, the difference between treated- and non-treated- applicants is quite small and not statistically significant. Moreover, for three of the four outcomes (jobs with standard working hours, permanent contract occupations and the average hourly remuneration) such slim difference is in favor of the non-treated applicants rather than the treated applicants, while the latter group is the one that display the most positive variation in the job-quality outcomes in the post-application periods, as highlighted in Table 2.

Table 3 – Pre-treatment trends of the job-quality outcome variables – average change (per firm) between the (t-3) and (t-1): Treated firms vs. non-treated applicants

| <i>Outcome variable Y</i> | $\Delta Y [(t-1) - (t-3)]$ | | |
|--|-------------------------------|-----------------------------------|----------------------------|
| | <i>Treated applicants (a)</i> | <i>Non treated applicants (b)</i> | <i>D (a)-(b) sd t-stat</i> |
| Jobs with standard working hours (1=1 job) | + 1.85 | + 2.41 | -0.56 0.43 -1.28 |
| Skilled jobs (1=1 job) | + 1.61 | + 1.29 | +0.32 0.35 0.91 |
| Jobs with permanent contracts (1=1 job) | + 1.27 | + 1.78 | -0.51 0.42 -1.21 |
| Average hourly remuneration (1= 1 €) | + 2.70 | + 3.36 | -0.64 0.39 -1.68 |

The exogenous treatment-assignment process of the POE/PRIME program in the 2003-6 period could be exploited in the analysis by means of a mere pre-post control group design, in terms of a dynamic *difference-in-differences* estimator separately applied to each consecutive yearly cohort of treated- and non-treated applicant firms. In the case of the POE/PRIME program, however, the number of applicants in each yearly wave of calls for applications is not large enough to guarantee that the pseudo-randomization conditions, deriving from the cut-offs in the exogenous application scores, achieve a full balancing of some crucial observable characteristics that may expose the firms to possible heterogeneous future influences that may differently

affect the post-treatment job-quality outcomes (such as: industrial sector, size class, regional location, age of the firm, pre-application employment growth-trend and pre-application job-quality level). Indeed, the accidental clustering of these crucial observable characteristics in one of the two groups (within each yearly cohort) could expose the firms to different external dynamics that if left uncontrolled for, may pose a threat to the validity of the parallel trend assumption.

Such unbalancing is similar to the issue that would arise from using small sample sizes in actual randomized experiments. For example, Kernan et al., (1999) show that even with pure random assignment the probability that the treatment and control group will differ significantly along some important dimensions becomes very relevant with sample sizes smaller than 200 units.³⁶ In such small-sample experiments, it is well established in the literature (see for example the review by Bruhn and McKenzie, 2008) that one viable solution for preventing unbalancing is to adopt a block-randomization design based on the major confounding factors that may affect the outcome variable of the analysis.

For these reasons, in order to best explore the presence of the multiple yearly waves of applicants in the context of these natural experiment conditions, we mimic a pre-test-post-test block-randomization design in terms of a dynamic matching estimator combined with a DD scheme. This is implemented by comparing the pre-post application job quality outcome change of the treated and non-treated applicants that are selected by means of an exact matching procedure implemented separately within each yearly round of calls for applications. Our estimation procedure yields a separate impact (in terms of Average treatment effect on the treated -ATT) for each cell of identical firms within each cohort of yearly calls for applications. The different cell-specific ATTs are then successively aggregated into local impact estimates at the level of the yearly calls for applications. The global ATTs of interest are then finally estimated through aggregating these local impact estimates across the different yearly cohorts of calls. Such an estimator is quite a novelty in evaluating firm-level support interventions, with a similar approach being implemented, to the best of our knowledge, only in Bondonio & Martini (2012)³⁷.

36 Kernan et al. (1999) show that when considering a binary covariate present in 30 percent of the sample, the probability that the treatment and control groups will differ by more than 10 percent is 38% with a sample size of 50 units, 27% with 100 units, 9% with 200 units, and 2% with 400 units.

37 Dynamic statistical matching for successive cohorts of treated units, however, is a well-established estimation model in the field of evaluating active labour market policies (Sianesi, 2004) and it has also been recently adopted in the context of evaluating disaster relief assistance programs (Bondonio & Greenbaum, 2018).

Compared to standard (static) matching approaches, the dynamic feature of our model avoids reliance on control variables that are fixed at an initial point in time for all subsequent annual cohorts of treatments. Rather, the matching process separately pinpoints with high accuracy the specific pre-treatment periods for each cohort of treated firms.

Exact matching is preferable to a one-dimensional matching based on a single propensity score (PS) variable because a good part of our covariates are categorical by nature (i.e. firms' industrial sector, regional location), while others represent risk-factors for selection bias mainly across discrete intervals (e.g. the pre-application size class of the firms and the distinction between relatively new firms versus incumbent firms at the time of application).

Moreover, each of these covariates can have an equally relevant potential impact on the future determinations of the outcome variable. In these circumstances, as also discussed elsewhere in the literature (e.g., Ho et al., 2007; King et al., 2011), a one-dimensional matching based on a PS variable can disguise some admissible degree of mismatching on single covariates.³⁸ Such mismatching can lead to severe bias on the impact estimates, and for this reason it is recommended (Ho et al., 2007) that absolute constraints are placed in the matching procedures to guarantee the perfect balancing of the covariates with the highest rank of importance for the future determination of the outcome variable.

Because of these considerations, the relatively small number of our equally-important covariates is conducive to adopting an exact matching stratification of the applicants. Such a procedure guarantees both a limited loss of external validity (with only a limited number of treated applicants outside the common support) and the avoidance of any hidden residual degree of unbalance on some single covariates that could be conducive to selection bias in the impact estimates.

In detail, the estimation procedure of our dynamic exact-matching DD model can be summarized as follows:

38 This is because the balancing property of the propensity score (PS) is satisfied when for the single covariates non-statistically significant differences are detectable between the treated and comparison units, conditional to similar (PS). However, hidden behind the non-statistically significance there could be unbalancing of single covariates that might induce large selection bias in the impact estimates if such covariates are of high rank of importance for the future determination of the outcome variable.

- (ii) All firms contained in the POE/PRIME-LEED database are sorted into applicants and non-applicant groups. The applicant group is further divided into successive cohorts (p) of applicants based on their year of application.
- (iii) To fully explore the natural experiment conditions involved in program procedures, the non-applicant firms are dropped from the analysis.
- (iv) Within each cohort (p) of applicants, we coarsen the relevant covariates (X) into a number of discrete categories. These categories represent the distinct firm-level characteristics that in the years before the period of observation of our analysis are shown in the Portuguese LEED database to correlate with different levels of the job-quality outcomes that we consider in the analysis:
 - (a) Industrial sector (coarsened into: low- and medium-low-tech manufacturing industries; high- and medium-high-tech manufacturing industries; construction; trade; accommodation and food service activities; knowledge-intensive market services);
 - (b) Pre-application firm size: 1-4 employees; 5-9; 10-19; 20-49; 50-99; 100-250;
 - (c) Age of the firm at the time of the application: 1-4 years of age; 5 years or more;
 - (d) Region: Norte; Algarve; Centro; Lisboa; Alentejo; Açores; Madeira; Inter-regional³⁹;
 - (e) Pre-application growth-trend of employment [in terms of employment change between (t-3) and (t-1), with t being the application year]: workforce increase; constant; decrease; firm not in existence at t-3;
 - (f) The firms' pre-application level of the job-quality outcome (Y) used in the analysis is coarsened into two categories: one below and one above the average of the firms within the same sector and size class.
- (v) Separately for each cohort (p) of applicants sharing the same application year, we match the treated with the non-treated applicants with the same exact values of the coarsened controls (X). The result of this matching procedure is a number of cells (c) with identical categories of the controls (X).

39 The listed regions correspond to the Portuguese NUTS 2 administrative units. The *Inter-regional* category is applied to all firms that do not sell the majority of their products and services on a regional market.

- (vi) Local ATTs for each cell (c) of identical treated and non-treated matched applicants sharing a same application year (p) are estimated through the following DD estimator:

$$t_{pc}^{DD} = E [Y_{t+3}^1 - Y_{t-1}^0 \mid T_c=1, X, p] - E [Y_{t+3}^0 - Y_{t-1}^0 \mid T_c=0, X, p].$$

- (vii) Average ATTs for each cohort (p) of treated applicants sharing a same application year are estimated as the following weighted averages of the local t_{pc}^{DD} :

$$t_p^{DD} = \sum_c [\tau_{DDpc} * W_{pc}], \text{ with weights } W_{pc} = \frac{N_{pc}}{\sum N_{pc}} \text{ computed as the number of treated firms in cell } (c) \text{ over the total number of treated firms in cohort } (p)^{40}.$$

- (viii) Global ATTs, across all yearly cohorts of applicants are estimated as the weighted average of t_{DDp} , with weights proportional to the number of treated firms in each cohort (p).

5. RESULTS

The impact estimates from our preferred estimation model described above are summarized in Table 4.

Table 4 – ATT estimates with Exact Matching on treated and non-treated applicants

| <i>Outcome variable (Y)</i> | <i>N. of treated firms</i> | <i>N. of non-treated firms</i> | <i>ATT</i> | <i>sd</i> |
|--|----------------------------|--------------------------------|------------|-----------|
| Jobs with standard working hours (1=1 job) | 386 | 308 | 4.9 ** | 1.65 |
| Skilled jobs (1=1 job) | 385 | 320 | 2.9 ** | 1.41 |
| Permanent jobs (1=1 job) | 389 | 311 | 2.0* | 1.57 |
| Avg. remuneration per hour (1= 1 €) | 387 | 290 | 2.2** | 1.05 |

** = Statistically significant at 5% level.

* = Statistically significant at 10% level.

The results from Table 4 indicate that the POE/PRIME program co-funded by EU structural funds generated a growth in the number of jobs with standard working hours that is higher than the counterfactual trend of no public support

40 Because rigorous internal validity calls for separately considering the different yearly cohort of treatments and due to the absence of reliable existing evidence available on the job-quality impact of EU funded programs, the main focus of the analysis is maintained on estimating ATTs parameters. Estimating heterogeneous impacts based on different characteristics of the treated firms, instead, would lead to an insufficiently large number of treated applicant firms within each cohort of calls for applications. Moreover, in the case of the POE/PRIME program, no clear theory of change can be highlighted to predict different impacts in terms of job-quality outcomes based on the limited heterogeneity of the admissible beneficiaries across the different program schemes.

by an average of 4.9 additional jobs per firm. Such a positive program impact on job quality indicators is confirmed also for the number of skilled and permanent jobs (with an estimated growth impact of +2.9 skilled jobs and +2.0 permanent jobs per firm, compared to the counterfactual trend of no public support), and for the change in the average pay per hour [with an estimated growth impact of +2.20€ per hour (+17.8%)⁴¹ compared to the counterfactual trend of no public support]. Given that the average value of the support assigned to each treated firm included in the analysis is about 78,720€, these impact estimates yield the following cost of the subsidies per each additional job generated by the program: 16,065€ for the jobs with standard working hours; 27,145€ for the skilled jobs; and 39,360€ for the permanent jobs.

6. SENSITIVITY ANALYSIS

Our preferred estimation model maximizes the controlling of selection bias in the impact estimates by excluding from the comparison group the firms that did not apply for the program incentives. This exclusion is vital to fully explore the desirable natural experiment conditions arising from the program treatment assignment procedures and to achieve the balancing of the important unobservable covariates captured by the decision to apply for the program (e.g. the firms' positions in their investment cycles, risk aversion, and management style).

Excluding non-applicant firms from the potential composition of the comparison groups, however, comes at a cost of modestly: i) decreasing the statistical efficiency and external validity of the impact estimates (due to the exclusion of some treated applicants that are outside the common support); ii) reducing the possibility of achieving a stricter balancing of the observable covariates. The latter point is due to the fact that the sample size of the comparison groups formed by the non-treated applicants is smaller than that of the comparison groups drawn from the general population of non-treated firms. For this reason, a matching procedure based on non-treated applicants has to use coarser categorizations of some observable covariates (e.g. size class and age of the firms) than a matching that uses the general population of non-treated firms.

41 The average per-hour remuneration in the treated firms was about 17.80€.

Because of such possible limitations i) and ii), we tested the robustness of our results by replicating the analysis with three additional estimation models that enlarge the composition of the comparison group to include the non-applicant firms as well. These models are the following: a) The same preferred estimation procedure replicated on an enlarged comparison group that also includes all the non-applicant firms; b) A Propensity Score Matching – PSM – (radius) procedure based on the same categorical covariates of the preferred estimation model, but applied to the enlarged comparison group; c) A PSM (radius) procedure applied to the enlarged comparison group, but with continuous functional forms for the following control variables: firm size (number of employees); firm age (in years); pre-intervention growth trend (employment change); pre-intervention level of the outcome variable.

Table 5 - Results from the sensitivity analysis - Comparison of ATT estimates from different estimation models

| <i>Output variable Y</i> | <i>Exact matching (All firms) a)</i> | <i>PSM-categorical controls (All firms) b)</i> | <i>PSM continuous controls (All firms) c)</i> |
|--|--------------------------------------|--|---|
| Jobs with standard working hours (1=1 job) | 5.9 *** | 5.8 *** | 5.0 *** |
| (sd) | (0.62) | (0.14) | (0.15) |
| Skilled jobs (1=1 job) | 4.1 *** | 3.9 *** | 4.4 *** |
| (sd) | (0.50) | (0.13) | (0.13) |
| Permanent jobs (1= 1 job) | 4.3 *** | 4.1 *** | 3.2 *** |
| (sd) | (0.49) | (0.13) | (0.13) |
| Avg. remuneration per hour (1= 1 €) | 1.8*** | 0.0 | 0.0 |
| (sd) | (0.44) | (0.03) | (0.03) |

*** = Statistically significant at 1% level.

** = Statistically significant at 5% level.

* = Statistically significant at 10% level.

The results from these three additional estimation models, summarized in Table 5, show a low volatility of the impact estimates across the different methodological options (with the only exception of the two PSM impact estimates on average hourly remuneration), confirming an overall adequate robustness for our findings.

7. CONCLUDING REMARKS AND POLICY IMPLICATIONS

In this paper we exploited the availability of linked employer-employee data and the presence of desirable natural experiment conditions offered by a large program co-funded by the European Regional Development Fund in Portugal (the POE/PRIME program) to estimate the impact of firm investment support policies on a number of job quality outcomes. To the best of our knowledge, this is the first paper providing robust empirical evidence regarding the impact of EU-funded enterprise support policies on job-quality outcomes at the firm level.

The linked employer-employee data (LEED) that we use in the analysis, merged with the activity data from the POE/PRIME program, offer ideal impact identification conditions for the following reasons. First, our LEED data come from administrative records that are free of the many attrition bias and measurement error limitations that plague the firm-level balance sheet data commonly used in counterfactual evaluation of firm level support programs. Second, the POE/PRIME program activity data ensure an adequate external validity for the results and limit the treatment contamination issues deriving from the concurring presence of other policy interventions. This is because the POE/PRIME program is a very prominent EU cohesion policy intervention, with a total budget of nearly €8.3 billion (and direct support to firms that represented nearly $\frac{3}{4}$ of the program's total public funds) and it was by far the most prominent source of support for Portuguese firms during our period of observation.

The natural experiment conditions that we exploit in the analysis arise from the exogeneity of the application scores that determined the treatment status of the firms in a particular period of the program implementation (i.e. the years 2003-6). These conditions ensure the possibility of avoiding the selection bias issues that are commonly encountered in impact evaluation studies of enterprise support programs when no exogenous treatment exclusions are in place.

We exploit these favorable impact identification conditions by means of a preferred estimation model that mimics a block-randomized pre-test-post-test design in the form of an exact matching procedure, combined with a *difference-in-differences* scheme that is adopted separately for each of the successive cohorts of applicant firms. The robustness of the estimates from this preferred

model is thoroughly tested by means of a wide sensitivity analysis in terms of using a different comparison group (the universe of all non-treated firms, instead of non-treated applicant firms) and using an alternative matching procedure (propensity score versus exact matching).

Our results indicate that the POE/PRIME program has generated positive impacts on job quality outcomes, compared to a counterfactual scenario of no public support, with each treated firm creating on average 4.9 additional standard-working-time jobs, 2.9 skilled jobs, and 2.0 permanent contract jobs. We also estimate that these job-quality gains were coupled by positive impacts on the average per-hour remuneration of the treated firms (+2.20€, equal to +17.8%). In terms of cost per unit of impact, these results highlight that about 16,100€ of public subsidies are needed to generate each additional job with standard working hours, while this cost is about 27,100€ for the skilled jobs; and about 39,400€ for the permanent jobs. In the existing literature on business incentive programs in Europe, the cost-per-job figures related to general employment outcomes tend to be within a range of 6,300€ -77,500€ (Bondonio & Greenbaum, 2006, 2014; Bondonio & Martini, 2012, 2019; Cerqua & Pellegrini, 2014; Criscuolo et al., 2012). In this regard, our results indicate that for the case of the POE/PRIME program, EU funded enterprise support interventions generated favourable job-quality outcomes at a cost (per unit of impact) that remains in lines with that attributable to general employment gains.

Although the scope of the analysis recommends some caution in extracting strong policy conclusions at the European level, our findings bear adequate external validity and have some noticeable implications that should be considered. First, they highlight the fact that increasing job quality is a policy goal that can be pursued, at a reasonable cost, by means of EU industrial policies in the form of enterprise support interventions (even if job-quality outcomes were not an explicit goal of these policies). Thus, our results indicate that industrial policies can have noticeable impacts on job-quality outcomes in the same vein as labor market regulations, social dialogue, and active labor market policies that were traditionally be regarded as the main instruments to achieve desirable job-quality outcomes (International Labour Organization, 2014; World Bank, 2013).

Second, our results implicates that the regional policies that attempt to increase the endowment of human capital through individual-based policies, as opposed

to place-based ones (Barca et al., 2012) – as is often the case with education and training policies – can be successfully complemented with policy instruments that aim to expand the demand for high quality labour. Such demand expansion can derive from policy interventions aimed at supporting firm-level investment projects that may increase the stock of more innovative and sophisticated production procedures and organizational schemes.

These considerations are of particular relevance in the present times in which an increasing pressure is placed on the EU Cohesion Policy to effectively foster the creation of decent job opportunities in the less developed regions that have to deal with enduring high unemployment rates and slow economic growth.

CHAPTER 6

Conclusion



1. OVERVIEW AND RESEARCH CONTRIBUTIONS

We care about jobs. Their quantity and quality – productive, safe, fairly remunerated, life-work balanced – mirrors an economy's capacity to use available human capital in the creation of value for all. But jobs move, and jobs change, in a geographically uneven way. Long ago, employment expanded in the manufacturing sector, while draining from agriculture. More recently, Artificial Intelligence and Modern Robotics affect labour in many economic activities, including services. In the evolution of work, some places lose more employment than other places due to automation, recessionary shocks, or other adverse events. Some places end up having better jobs than others. Why? Despite a growing interest of scholars and policy makers, this remains poorly understood.

We know that the particular portfolio of jobs in each region conditions the geography of jobs (Moretti, 2012). But not alone. The way jobs relate to each other – relatedness – also shapes their evolution. For instance, regions tend to diversify into new job specializations that are highly related to previously existing ones (Alabdulkareem et al., 2018; Muneeppeerakul et al., 2013; Hidalgo et al., 2018). Relatedness also seems to prevent the exit of job specializations in a region, conditioning its capacity to cope with adverse events (Boschma, 2015). Finally, policy makers often make use of local capabilities to reach desired outcomes (Balland et al., 2019; Fagerberg & Verspagen, 2009).

This thesis aims to further advance our knowledge on how relatedness determines the labour dynamics that operate in each region, continuously shaping the geography of work. Several related bodies of literature – Evolutionary Economic Geography (EEG), Labour Economics, Urban Scaling, Innovation Studies, Regional Policy – offered valuable and robust foundations I could build on. The contributions presented in this thesis are bounded by the following research question:

RQ: How does relatedness between jobs shape the geography of work?

This is the main research question that bounds the research work in this thesis. It unfolds in four specific ones. They are listed below:

RQ1: Does relatedness scale with city size?

RQ2: Which dimension(s) of relatedness drive the entry and exit of jobs in cities?

RQ3: What is the impact of automation through relatedness on job growth in cities?

RQ4: What is the impact of policy incentives for business through relatedness on job quality?

The first two (*RQ1* and *RQ2*) focus on how relatedness “pulls” capabilities together in a region – “magnet effects. *RQ3* addresses how impacts from automation, may spread in the labour system through the structure of relatedness between jobs – “diffusion effects”. *RQ4* refers to how relatedness can be incorporated in the implementation of regional policy to boost job quality in local firms’ workforce. These four research questions are addressed one by one in each empirical chapter of this thesis.

1.1. On the “magnet effects” of relatedness

Multiple studies have recently shown how relatedness attracts new capabilities while preventing the exit of existing ones (Hidalgo et al., 2018) – “magnet effects”. If new capabilities in a city tend to be related with previously existing ones, how much does relatedness grow with diversity? In Chapter 2 (*RQ1*), we use employment data at the city-job level from six industrialized countries (United States, Canada, Sweden, Portugal, Australia, and Germany). First, we create a network representation for each urban labour system (as in Muneeppeerakul et al., 2013). In this network, the “Job Space” (also referred as occupational space), nodes are job specializations of a city, and links depict relatedness between them, measured as geographical relatedness (as in Hidalgo et al., 2007). We modelled the growth of relatedness between jobs in a city as a superlinear function of city size (number of job specializations), as follows:

$$D_i = \alpha N_i^\beta, \beta > 1 \quad (1)$$

where D is network relatedness density, i denotes time in a particular city, α captures the effects of technology and institutional arrangements, N is the number of job specializations, and the exponent β is an elasticity parameter for how much the network relatedness varies with network size. We test this exponential model, where network density scales with network size, against the null hypothesis of proportional grow.

Relatedness scales with city size. Our results indicate that relatedness grows superlinearly with city diversity in all six countries. The estimated scaling exponents range from 1.17 in Swedish cities to 2.35 in USA cities. This means that bigger cities have a higher share in job specializations that have higher levels of relatedness to neighboring jobs in the local economy.

Relatedness self-reinforces as a city grows. By confirming model (1), showing a β consistently higher than one, our empirical results indicate that, rather than a proportional increase in infrastructural and input complementarities (where β would equal one), additional links are formed. This implies that relatedness tends to self-reinforce itself as the city grows, creating more than proportional economic diversification opportunities for bigger cities.

In Chapter 3 (RQ2), we use employment data (city-job, industry-job, and task-job levels of analysis) from US BLS and O*NET (2005 to 2016). As before, nodes are job specializations and links measure the level of relatedness between them. But now, the network has three types of links, one for each dimension of relatedness. This network is displayed as a multi-layer network, with one layer for each relatedness dimension. Complementarity measures how often two job specializations co-occur in the same input-output chain. Similarity measures their overlap in the same tasks. And synergy is measured by regressing geographical relatedness on local complementarity and similarity, and saving the residuals as local synergies. We estimated three-way fixed effects regression models to capture the “magnet effects” of each relatedness dimension.

Relatedness drives diversification in three distinct ways. Our results confirm that relatedness significantly favours entry of new job specializations in a city, while preventing exit of existing ones. But it does so in different magnitudes. Synergy shows far stronger “magnet effects”, while the effect of similarity is just slightly higher than the effect of complementarity.

Relatedness in three dimensions. Three forces of agglomeration seem to bring firms together in a city (Marshall, 1920): access to specialized suppliers (complementarities along input-output chains), labour pools (similarity of knowledge and skills), and knowledge spillovers that form economically vibrant ecosystems (Jacobs, 1969). In order to compare the “magnet effects” for each dimension of relatedness, we unpack it as follows:

$$GeoRelatedness_{i,j} = \beta_1 Similarity_{i,j} + \beta_2 Complementarity_{i,j} + Synergy_{i,j} \quad (2)$$

where *GeoRelatedness* is the widely used geographical measure of (total) relatedness between job specializations *i* and *j*. β_1 and β_2 correspond, respectively, to the shares of similarity and complementarity links that imply collocation (i.e., local similarity and complementary). This equation (2) is an identity model, assuming that synergy links are due to pure geographical collocation that does not result from similarity or complementarity. The resulting statistical distribution for each relatedness density shows that local synergy grows with complementarity and similarity, but not linearly. The very few job specializations that have high local synergies tend to have medium levels of local complementarities, and medium-high levels of local similarities.

Measuring local synergies between local capabilities. Geographical collocation of economic capabilities has become widely used to measure relatedness. However, it is a “black box” kind of measure. It captures how often two capabilities collocate in a region, thus, the geographical result of all “magnet effects” of relatedness combined. Some authors use alternative measures, often without identifying which dimension(s) of relatedness is being captured, and/or not covering the remaining ones. Moreover, the idiosyncratic nature of relatedness within vibrant economic niches, local synergies, makes them particularly difficult to measure directly. On the other hand, synergies are intrinsically local, as well as the links in geographical relatedness. Thus, the third dimension of relatedness can be measured indirectly. We did so, in Chapter 3, as follows:

$$Synergy_{i,j} = GeoRelatedness_{i,j} - (\beta_1 Similarity_{i,j} + \beta_2 Complementarity_{i,j}) \quad (3)$$

For network visualization purposes, we solve (3) by isolating high *GeoRelatedness* links that have low *Similarity* and low *Complementarity*. Whereas for the main analysis, we measure *Synergy* as a residual from regressing geographical relatedness density at the job-city level on complementarities and similarities (with resulting high R-square of 0.88).

1.2. On the “diffusion effects” of relatedness

Chapter 4 (RQ3) extended the Job Space once more, this time to account for different types of jobs specializations (nodes) with regard to their technical risk of automation (Atkinson, 2017). Although this *Job Space Under Automation* is not a directed network, from the perspective of each node (source node), both its

links and the destination of its links (target node) matter, having high technical risk of automation or not. The paper uses data from Atkinson (2017), BLS and O*NET, 2005 to 2016, to regress job employment growth in cities between 2007 and 2016 on complementarities to high-risk jobs and similarities to high-risk jobs.

Impacts from automation diffuse locally. Results confirm that relatedness promotes the diffusion of impacts from automation between professions. High complementary to high-risk jobs, and low similarity to high-risk jobs, tend to increase employment in jobs in US cities. It confirms recent literature suggesting that the demand for labour increases in jobs complemented, and not substituted, by the new technologies (Autor, 2015; Bessen et al., 2019). The process from technical potential to actual adoption of technology seems to depend on a myriad of local factors. They affect the displacement costs and productivity gains that weight against each other in each firm's choices between capital and labour (Acemoglu & Restrepo, 2019; Lordan & Neumark, 2018). But at the end, displacement costs seem to be conditioned by proximity to similar jobs that are already getting automated. The more similar nearby jobs, the easier, and cheaper, to adapt the technology to automate those jobs too (Nedelkoska & Quintini, 2018). While the transference of productivity gains must be higher where more jobs collaborate with potentially automatable jobs (Brynjolfsson & Mitchell, 2017). Thus, solely considering the composition of tasks within jobs is not enough to determine jobs' risk of automation. The structure of similarities and complementarities between automatable and non-automatable jobs in each city must also be accounted for.

"Diffusion effects" of relatedness are context dependent. In addition, the results have, at least, two theoretical-conceptual implications. First, it confirms the existence of "diffusion effects" of relatedness found in previous studies for other levels of analysis (Foster-McGregor & Pöschl, 2016; Jun et al., 2019; Morrison, 2008; Morrison et al., 2013). Second, diffusion effects depend on characteristics of the source and target nodes (each link connects two nodes, target and source).

Measuring jobs network risk of automation. The diffusion of impacts from automation does not seem to spread homogeneously to all neighbouring local capabilities. Rather, it selects diffusion targets among neighbouring jobs. Jobs that become automated seem to spread automation risk to neighbouring similar

jobs and provide productivity gains to complementary jobs. This means that, in order to capture “diffusion effects” of relatedness, we need a new measure that accounts for the links of each job (source node), but also for the risk level of neighbouring jobs (target nodes) and the kind of relatedness they have. A simple transformation of a widely used measure can serve this purpose. More concretely, this paper introduces a new filter (whether jobs have high technical risk of automation or not) to measure relatedness density at the job-city level, as follows:

$$RelatednessToHighRisk_{i,c} = \frac{\sum_{i \neq j, j \in c, Risk_j = high} Relatedness_{i,j}}{\sum_{i \neq j} Relatedness_{i,j}} * 100 \quad (4)$$

where *RelatednessToHighRisk* measures, for each job *i* in city *c*, the share of links to neighboring jobs ($j \neq i$ that predominate in city *c*) that are at high risk, over the total relatedness of job *i*. Equation (4) can be adapted to any type of relatedness link, or other adverse events. For instance, under automation, in line with the empirical results in the paper, similarities and complementarities seem to be particularly relevant, yet with opposite impacts. Moreover, based on (4), the paper adapts Atkinson (2017)’s technical risk of automation to build a novel (preliminary) index for US cities, that accounts for network risk (available in tfarinha.wixsite.com/tfarinha).

1.3. On how innovation policy can boost job quality

In Chapter 5 (*RQ4*), we use Portuguese linked-employer-employee data (2000 to 2009) and data on EU supported projects. The paper combines distinct Counterfactual Impact Evaluation methods to adjust to the specificities of the evaluation case and includes a valid control group (dynamic matching, difference-in-differences, and block-randomized pretest-posttest design). We quantify the impact of business policy incentives (co-funded by European Regional Development Fund) on Portuguese firms’ job quality in four dimensions: security, work-life balance, compensation, and specialization level.

Innovation policy incentives can boost job quality. Our results show positive impacts for all four job quality variables. More concretely, after treatment, in comparison with the control group of no treatment, each treated firm seems to have generated an average of 5 additional standard-working-time jobs, 3 skilled jobs, 2 permanent-contract jobs, and 17.8% higher remuneration (+2.20€/hour). This confirms the importance of geography in such policy intervention for job quality. Policy business incentives to innovation do not directly target job quality

outcomes, and yet it affects them, as a by-product of the success of innovation projects in treated firms (beyond the project's implementation period). This is, in great part, because highly specialized jobs tend to have a particularly dense web of relatedness links and multiplier effects within and across local firms (Moretti, 2012). This can be stimulated by such policy instruments.

Policy as an actor of change. The results in this paper confirm that the high density of relatedness within innovative sectors can be stimulated to boost employment and job quality in local firms, and ultimately in the regional labour force. This means that policy intervention can interact with (and/or creates) labour dynamics, driving the evolution and geography of work.

2. LIMITATIONS AND FUTURE RESEARCH

The recent and rich topic of relatedness still has much to be investigated (Boschma 2017). The contributions presented in this thesis advance our knowledge regarding relatedness between jobs and how it shapes the evolution of work and its geography. As any research, this study is not without limitations. I will highlight and discuss a number of relevant aspects that I did not cover, followed by suggestions for further research.

We do not directly observe the formation of relatedness links in cities. In Chapter 2, we interpret our results (relatedness grows superlinearly with diversity of jobs in a city) based on theoretical assumptions that require stronger empirical support. Relying on existing literature, we assume that a proportional increase would simply correspond to a higher division of labour within input-output chains. Whereas a more than proportional increase (superlinear) indicates that additional relatedness links are being formed, outside input-output chains, by other types of relatedness than complementarities. But Chapter 2 does not directly observe the formation of new links, nor does it show which type of relatedness they represent. The collection of primary data and empirics at the firm level could bring clarity to this, distinguishing and comparing the different types relatedness links that are formed in the firm/cluster/city by new hires.

Local synergy needs further investigation. Chapter 3 measures local synergy, similarity and complementarity in one single paper. Synergy accounts for interdependencies and flows of knowledge (Jacobs, 1969; Marshall, 1920) between local capabilities in vibrant economic niches. Data wise, this is very

difficult to capture directly, and our indirect measure can only retain stronger links (e.g., paralegals often use taxi services to gather relevant information and urgent errands, or tailors and athletes, etc.). As a residual measure, it fails to capture local synergy between two jobs that is weaker than their complementarity and similarity. Also, it does not distinguish specific institutional arrangements or natural resources that make certain jobs often co-occur in cities. We also do not know how much the results are specific to North America. And, as few countries have the required data to replicate the method in this paper, alternative measures of local synergy are needed. New case studies and qualitative studies can show what comprises local synergy and which instrumental variables could represent it at a meso or macro level.

We do not know if local synergy holds as the main driver of city diversification for other levels of analysis. Existing literature suggests that local synergies would have even stronger “magnet effects” within innovative industries, given a higher concentration of specialized jobs (Moretti, 2012). Can the relative magnitudes of the three dimensions of relatedness change according to the type of industry? Would local synergies be weaker as a driver of diversification in less developed regions? Both a static and a dynamic approach could shed light into this. For instance, studying how local synergies overlap with specialized economic activities. Or repeating the analysis in Chapter 3 yet filtering by type of city, from low to high levels of economic complexity. This also opens another question. Would the results still hold, for instance, at the product level? In other words, are new product specializations mostly synergic with existing ones? Or mostly similar, or complementary? The method in this paper can be applied to products for sure.

I do not directly observe the diffusion of impacts from automation between workers, in Chapter 4. Empirics at more granular levels of analysis (case studies), or for different countries (such as developing countries) can reveal specific industrial dynamics not captured here. For instance, the impact of automation on reshoring (Foster-McGregor, Nomaler, & Verspagen, 2019), e.g., factories brought “back home”, substituting offshored jobs by robots in homeland (Belfiore, 2016; De Backer, Menon, Desnoyers-James, & Moussiégt, 2016; Ernst, Merola, & Samaan, 2018; Hallward-Driemeier & Nayyar, 2019).

Crossed diffusion effects from distinct events. More investigation is needed to understand how effects from distinct economic events might interact with each

other, when they occur simultaneously. In robustness analysis, the interaction of the Great Recession with automation was analysed. But technological transitions unravel in long time spans, and meanwhile, interact with other adverse events. For instance, soon we might have enough data to investigate whether the current Covid19 health pandemic accelerated automation. Or if it rather added new diffusion effects, contrary to the automation ones.

Theoretical formulation of diffusion effects. The findings in Chapter 4 ask for a theoretical formulation on how relatedness shapes labour dynamics under automation. When a fully new technology enters the city to automate a certain type of job (*job0*), the transmission of impacts to other jobs (*job1*) is not a homogeneous dissipation of impacts (rather selecting which neighbouring jobs will be *job1*, and subsequently *job2, job3*, etc.). At the end, each job has an overall risk of automation that is determined by both individual and systemic risks. This could be investigated by adapting Network Epidemics Modelling (Barabási, 2016) to the diffusion of impacts from automation between job specializations in each local labour system (city, region, or country). Also, Agent Based Modelling could be applied to forecast the net result of workers individual choices under automation (training programs, job mobility, etc.), given the specificities of each place (capabilities, magnet and diffusion effects of relatedness, institutional arrangements, etc.).

Connectivity between cities. While we focus on relatedness within cities, we do not cover relatedness between them. And certainly, relatedness also conditions global labour flows and other labour dynamics between different places. This is also an expanding research avenue in the EEG literature (Boschma et al., 2009; Henning, 2020; Timmermans and Boschma, 2014), that can inspire further research regarding labour flows and the evolving geography of work. For instance, the “magnet effects” could be further explored in light of migration trends. Whereas the “diffusion effects” could be investigated regarding the impacts of automation in global value chains.

The role of institutions. Laws, culture, taxes, labour unions, and also policy intervention shape labour dynamics in each particular place. Chapters 2 to 4 do not cover this directly, rather assuming that institutions get represented in the local structure of jobs, and are reflected in net employment growth within job classes in a city. For instance, digital technologies might increase institutions’ demand for law and big data specialists, and will create new laws that protect

individual data. But some cities may have very different institutional settings, that affect labour dynamics in particular ways. For instance, social capital has been shown to promote industry diversification in EU regions (Cortinovis, Xiao, Boschma, & van Oort, 2017). Also, relatedness between product specializations seem to affect regional diversification in different magnitudes in different varieties of capitalism (Boschma & Capone, 2015). Such forms of institutional settings certainly affect labour dynamics in cities. The replication of such studies, regarding relatedness between jobs, or case studies and qualitative studies, might advance our understanding of the role of institutions in local labour dynamics. Chapter 5 addresses policy intervention. It assumes the interaction of relatedness with the policy instrument under analysis. However, it does not observe it. Surveys and firm case studies can shed light on which particular relatedness links are stimulated (formed or strengthened) by business policy incentives to innovation.

New dimensions of job quality. As the new digital technologies alter the nature of work, the concept of job quality is also changing. For instance, specialized skills that are easily programable with new digital technologies will no longer increase productivity and labour demand when firms decide to automate those skills. This means that the job quality dimension of productive jobs is no longer well proxied by workers' academic degree or years of experience. Future policy evaluation, regarding job quality outcomes, should adapt accordingly, for instance by using firm's share in specialized jobs that have low risk of automation, or firms' investment in training in non-automatable skills. And we need to test the impacts of policy business incentives to innovation on these new variables for job quality.

3. POLICY IMPLICATIONS

The future workforce will be mostly based on skills that are the bottlenecks of the new digital technologies, such as creativity, social and emotional intelligence, and cognitive flexibility. The local labour systems that can renew themselves towards a workforce that rely on those skills, will succeed, otherwise face high unemployment levels. Which means that the workforce must adapt at a quicker pace than the automation of tasks, in order to keep employment levels. This can also generate labour demand for skills that can assist a sustainable transition, such as research and management skills. Certainly, the final outcomes of

automation depend on how the transition process unravels and what we can do about it (Frank et al., 2019).

While technical feasibility determines the potential for automation, local factors and labour dynamics will determine how, when, and where new technologies get adopted (Brynjolfsson et al., 2017). Scholars are asked to deepen our knowledge on labour dynamics and harness policy with new analytical tools (OECD-NAEC, 2019) to foster employment resilience in economies more negatively affected. Our findings in this thesis might help identify which jobs in which cities might be at greater risk of automation, and where relatedness might promote greater labour productivity, and other dimensions of job quality, and thus, higher labour demand.

For instance, for more effective social protection to displaced workers and supporting training programs, policy makers need to be able to estimate local impacts from automation and identify most affected jobs and cities. Chapter 4 provides insights in how to map the diffusion of automation impacts between job specializations and use it to neutralise negative effects. Automation impacts seem to diffuse through the existing structure of relatedness between jobs in a region – “diffusion effects” of relatedness under automation – as follows. More automatable jobs tend to locally diffuse automation risks through similarities, and automation benefits through complementarities. Conversely, local similarity to less automatable jobs would provide a job with easier adaptation path towards non automatable skills. This means that policy intervention will be required especially where similarities to high-risk-jobs surpass complementarities to high-risk-jobs. This also points to the need of a two-folded strategy: help workers acquire non-automatable skills (minimize automation risks), while promoting work collaboration with AI solutions, to increase their productivity and labour demand (maximize automation benefits).

But policy can also stimulate the local effects of relatedness that affect labour dynamics, in a way that leads to a quicker renewal of skills in the local labour force and promote high-quality job employment. For instance, according to our results in Chapter 5, EU policy incentives to firms’ innovation projects have an important by-product – job quality, as measured by the increase of specialized jobs and improved working conditions in local firms. This confirms that policy can play an active role in the renewal of the workforce.

However, there is still no theory of change identifying the chain of effects from the implementation of policy business incentives to outcomes in job quality of local firms. This is crucial to support the design of policy instruments (Weiss, 1995) where job quality is a central policy goal, and the effects of relatedness on labour dynamics are efficiently explored. This is of crucial relevance under the current wave of automation, as it may severely affect employment and job quality in many regions.

The empirical findings in Chapters 2 and 3 can contribute to this. Besides diffusion effects, relatedness also affects labour dynamics through “magnet effects”. Relatedness has extensively been shown to favour the entry of new job specializations that are highly related with existing ones in the city, while also preventing exit of existing job specializations. Chapter 3 further shows that relatedness pulls jobs together in a city in three distinct ways: jobs might be attracted because they are complementary to existing ones, and/or similar, and/or synergic to them. The latter dimension of relatedness, local synergy, showed particularly strong “magnet effects”. Moreover, results in Chapter 2 show that bigger cities have more than proportionally higher levels of relatedness, in comparison with smaller cities. This seems to indicate that relatedness, and its effects in labour dynamics, reinforces itself as cities grow in diversity.

This might, in part, explain the increasing economic divergence between regions in the USA, for instance (Moretti, 2012). Regions with less relatedness density, especially of the local synergy kind, have less capacity for future job diversification and less job quality. Chapter 5 shows that policy can help lagging regions by stimulating the structure of relatedness around specialized jobs in innovative sectors. Which would also help lagging regions to reach higher resilience capacity to undergo a necessary transition to non-automatable skills and keep the labour force employed.

The research contributions presented in this thesis support the claim that relatedness between jobs shapes the constantly evolving geography of work. It does so in several ways, resulting in magnet and diffusion effects. The percolation of impacts from automation is also conditioned by the local structure of relatedness. The effectiveness of policy intervention is also bounded by relatedness linkages. We join scholars and policy makers defending that policy can, to a certain extent, guide the transition of skills in regional labour systems. It will depend on how well we understand how policy instruments interact

with existing labour dynamics. How the structure of relatedness in each place bounds the diffusion of knowledge and impacts from automation, attracting new jobs, and accelerating automation in other jobs. I hope this thesis inspires future policy-science collaboration to advance our knowledge and nurture a digital future with good jobs for all.

BIBLIOGRAPHY

- Acemoglu, D., & Autor, D. (2011). *Skills, tasks and technologies: Implications for employment and earnings*. *Handbook of Labor Economics* (Vol. 4). [https://doi.org/10.1016/S0169-7218\(11\)02410-5](https://doi.org/10.1016/S0169-7218(11)02410-5)
- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The Network Origins of Aggregate Fluctuations. *Econometrica*, 80(5), 1977–2016. <https://doi.org/10.3982/ecta9623>
- Acemoglu, D., LeLarge, C., & Restrepo, P. (2020). *Competing with Robots: Firm-Level Evidence from France*. Cambridge, MA. <https://doi.org/10.3386/w26738>
- Acemoglu, D., & Restrepo, P. (2017). *Robots and Jobs: Evidence from US Labor Markets*. Cambridge, MA. <https://doi.org/10.3386/w23285>
- Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*. <https://doi.org/10.1257/jep.33.2.3>
- Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines : the simple economics of artificial intelligence*.
- Alabdulkareem, A., Frank, M. R., Sun, L., AlShebli, B., Hidalgo, C., & Rahwan, I. (2018). Unpacking the polarization of workplace skills. *Science Advances*, 4(7). <https://doi.org/10.1126/sciadv.aao6030>
- Arntz, M., Gregory, T., & Zierahn, U. (2016). The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. <https://doi.org/10.1787/5jlz9h56dvq7-en>
- Arntz, M., Gregory, T., & Zierahn, U. (2017). Revisiting the risk of automation. *Economics Letters*, 159, 157–160. <https://doi.org/10.1016/J.ECONLET.2017.07.001>
- Atalay, E., Hortaçsu, A., Roberts, J., & Syverson, C. (2011). Network structure of production. *Proceedings of the National Academy of Sciences of the United States of America*, 108(13), 5199–5202. <https://doi.org/10.1073/pnas.1015564108>
- Atkinson, R. D. (2017). Unfortunately, Technology Will Not Eliminate Many Jobs. Information Technology and Innovation Foundation. Retrieved from <https://itif.org/publications/2017/08/07/unfortunately-technology-will-not-eliminate-many-jobs>
- Augusto Mateus & Associados. (2005). *Atualização da Avaliação Intercalar do Programa de Incentivos Modernização da Economia*.
- Autor, D. (2010). *The Polarization of Job Opportunities in the U.S. Labor Market Implications for Employment and Earnings*.
- Autor, D. (2014, May 23). Skills, education, and the rise of earnings inequality among the “other 99 percent.” *Science*. American Association for the Advancement of Science. <https://doi.org/10.1126/science.1251868>
- Autor, D. (2015). Why are there still so many jobs? the history and future of workplace automation. In *Journal of Economic Perspectives* (Vol. 29, pp. 3–30). <https://doi.org/10.1257/jep.29.3.3>
- Autor, D., & Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5), 1553–1597. <https://doi.org/10.1257/aer.103.5.1553>
- Autor, D., Levy, F., & Murnane, R. (2003). *The skill content of recent technological change: an empirical exploration*. Retrieved from <https://economics.mit.edu/files/11574>

- Balland, P.-A., & Rigby, D. (2017). The Geography of Complex Knowledge. *Economic Geography*, 93(1), 1–23. <https://doi.org/10.1080/00130095.2016.1205947>
- Balland, P. A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. *Regional Studies*, 53(9), 1252–1268. <https://doi.org/10.1080/00343404.2018.1437900>
- Barabási, A.-L. (2016). Network science introduction. *Network Science*, 1–27. <https://doi.org/10.1007/978-1-4614-6744-1>
- Barbour, E., & Markusen, A. (2007). Regional Occupational and Industrial Structure: Does One Imply the Other? *International Regional Science Review*, 30(1), 72–90. <https://doi.org/10.1177/0160017606296727>
- Barca, F., Mccann, P., & Rodríguez-Pose, A. (2012). The case for regional development intervention: Place-based versus place-neutral approaches. *Journal of Regional Science*, 52(1), 134–152. <https://doi.org/10.1111/j.1467-9787.2011.00756.x>
- Bartik, T. J. (2012). The Future of State and Local Economic Development Policy: What Research Is Needed. *Growth and Change*, 43(4), 545–562. <https://doi.org/10.1111/j.1468-2257.2012.00597.x>
- Beaudry, P., Green, D. A., & Sand, B. (2012). DOES INDUSTRIAL COMPOSITION MATTER FOR WAGES? A TEST OF SEARCH AND BARGAINING THEORY. *Econometrica*, 80(3), 1063–1104. <https://doi.org/10.3982/ECTA8659>
- Bechichi, N., Grundke, R., Jamet, S., & Squicciarini, M. (2018). *Moving between jobs: an analysis of occupation distances and skill needs*. Retrieved from www.oecd.org/going-digital
- Beckstrom, R. (2008). *The Economics of Networks and Cybersecurity*.
- Belfiore, M. (2016, April 13). Automation Brings Manufacturing Back Home | Automation World. Retrieved July 11, 2020, from <https://www.automationworld.com/factory/iiot/article/13315185/automation-brings-manufacturing-back-home>
- Benzell, S. G., Brynjolfsson, E., Maccrory, F., & Westerman, G. (2019). Identifying the Multiple Skills in Skill-Biased Technical Change. *The MIT Press*.
- Bernini, C., & Pellegrini, G. (2011). How are growth and productivity in private firms affected by public subsidy? Evidence from a regional policy. *Regional Science and Urban Economics*, 41(3), 253–265. <https://doi.org/10.1016/j.regsciurbeco.2011.01.005>
- Berriman, R., Hawksorth, J., Kelly, C., & Foyster, R. (2017). Will robots steal our jobs? The potential impact of automation. *PwC Report: UK Economic Outlook March 2017*, (March), 30–48. Retrieved from <https://www.pwc.co.uk/economic-services/ukeyo/pwcukeo-section-4-automation-march-2017-v2.pdf>
- Bessen, J., Goos, M., Salomons, A., & van den Berge, W. (2020). Firm-Level Automation: Evidence from the Netherlands. *AEA Papers and Proceedings*, 110, 389–393. <https://doi.org/10.1257/pandp.20201004>
- Bessen, J., Goos, M., Salomons, A., & Van Den Berge, W. (2019). *What happens to workers at firms that automate CPB Discussion Paper Change in annual income year relative to automation Automatic Reaction-What Happens to Workers at Firms that Automate? **. Retrieved from <https://www.cpb.nl/sites/default/files/omnidownload/CPB-Discussion-Paper-390-Automatic-Reaction-What-Happens-to-Workers-at-Firms-that-Automate.pdf>

- Bettencourt, L M A. (2013). The Origins of Scaling in Cities. *Science*, 340(6139), 1438–1441. <https://doi.org/10.1126/science.1235823>
- Bettencourt, Luis M.A. (2014). Impact of changing technology on the evolution of complex informational networks. *Proceedings of the IEEE*, 102(12), 1878–1891. <https://doi.org/10.1109/JPROC.2014.2367132>
- Bettencourt, Luís M.A. (2013). The origins of scaling in cities. *Science*, 340(6139), 1438–1441. <https://doi.org/10.1126/science.1235823>
- Bettencourt, Luís M.A., Samaniego, H., & Youn, H. (2014). Professional diversity and the productivity of cities. *Scientific Reports*, 4(1), 1–6. <https://doi.org/10.1038/srep05393>
- Beyers, W. B. (2013). The Great Recession and State Unemployment Trends. *Economic Development Quarterly*, 27(2), 114–123. <https://doi.org/10.1177/0891242413479653>
- Bia, M., & Mattei, A. (2012). Assessing the effect of the amount of financial aids to Piedmont firms using the generalized propensity score. *Statistical Methods and Applications*, 21(4), 485–516. <https://doi.org/10.1007/s10260-012-0193-4>
- Bondonio, D., & Greenbaum, R. T. (2006). Do Business Investment Incentives Promote Employment in Declining Areas? Evidence from EU Objective-2 Regions. *European Urban and Regional Studies*, 13(3), 225–244. <https://doi.org/10.1177/0969776406065432>
- Bondonio, D., & Greenbaum, R. T. (2014). Revitalizing regional economies through enterprise support policies: an impact evaluation of multiple instruments. *European Urban and Regional Studies*, 21(1), 79–103. <https://doi.org/10.1177/0969776411432986>
- Bondonio, D., & Greenbaum, R. T. (2018). Natural disasters and relief assistance: Empirical evidence on the resilience of U.S. counties using dynamic propensity score matching. *Journal of Regional Science*, 58(3), 659–680. <https://doi.org/10.1111/jors.12379>
- Bondonio, D., & Martini, A. (2012). Counterfactual impact evaluation of cohesion policy: impact and cost-effectiveness of investment subsidies in Italy. *Final Report to DG Regional Policy*.
- Bondonio, D., & Martini, A. (2019). Are they worth it?: A counterfactual impact evaluation of a decade of investment subsidies to Italian firms. *Scienze Regionali*, 18(2), 261–290. <https://doi.org/10.14650/93650>
- Boschma, R. (2015). Regional resilience from an evolutionary perspective. *Special Issue Paper Regional Studies*, 1–23.
- Boschma, R. (2015). Towards an Evolutionary Perspective on Regional Resilience. *Regional Studies*, 49(5). <https://doi.org/10.1080/00343404.2014.959481>
- Boschma, R. (2017). Relatedness as driver of regional diversification: a research agenda. *Regional Studies*, 51(3), 351–364. <https://doi.org/10.1080/00343404.2016.1254767>
- Boschma, R. A., & Frenken, K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography*, 6(3), 273–302. <https://doi.org/10.1093/jeg/lbi022>

- Boschma, R., Balland, P. A., & Kogler, D. F. (2015). Relatedness and technological change in cities: The rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. *Industrial and Corporate Change*. <https://doi.org/10.1093/icc/dtu012>
- Boschma, R., & Capone, G. (2015). Institutions and diversification: Related versus unrelated diversification in a varieties of capitalism framework. *Research Policy*, 44(10), 1902–1914. <https://doi.org/10.1016/j.respol.2015.06.013>
- Boschma, R., Eriksson, R., & Lindgren, U. (2009). How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity. *Journal of Economic Geography*, 9(2), 169–190. <https://doi.org/10.1093/JEG/LBN041>
- Boschma, R., Minondo, A., & Navarro, M. (2013). The Emergence of New Industries at the Regional Level in Spain: A Proximity Approach Based on Product Relatedness. *Economic Geography*, 89(1), 29–51. <https://doi.org/10.1111/j.1944-8287.2012.01170.x>
- Boxall, P., & MacKy, K. (2009). Research and theory on high-performance work systems: Progressing the high-involvement stream. *Human Resource Management Journal*, 19(1), 3–23. <https://doi.org/10.1111/j.1748-8583.2008.00082.x>
- Brachert, M. (2016). *The Rise and Fall of Occupational Specializations in German Regions from 1992 to 2010. "Relatedness as driving force of human capital dynamics.*
- Bradberry, T. (2017). Emotional intelligence: What it is and why you need it. Retrieved August 1, 2019, from <https://www.weforum.org/agenda/2017/02/why-you-need-emotional-intelligence/>
- Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. *Research Policy*, 32(1), 69–87. [https://doi.org/10.1016/S0048-7333\(02\)00004-5](https://doi.org/10.1016/S0048-7333(02)00004-5)
- Bresnahan, T., & Greenstein, S. (1996). Technical Progress and Co-invention in Computing and in the Uses of Computers. *Brookings Papers on Economic Activity*, 27(1996 Microeconomics), 1–83.
- Brian J. L. Berry. (1968). *Theories of urban location*. Washington, D.C.: Association of American Geographers. Retrieved from <https://catalogue.nla.gov.au/Record/2088018>
- Bronzini, R., & de Blasio, G. (2006). Evaluating the impact of investment incentives: The case of Italy's Law 488/1992. *Journal of Urban Economics*, 60(2), 327–349. <https://doi.org/10.1016/j.jue.2006.03.005>
- Bruhn, M., & McKenzie, D. (2008). *In Pursuit Of Balance: Randomization In Practice In Development Field Experiments* (Policy Research Working Papers). The World Bank. <https://doi.org/10.1596/1813-9450-4752>
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age : work, progress, and prosperity in a time of brilliant technologies*. Retrieved from <https://www.norton.com/books/The-Second-Machine-Age/>
- Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *American Association for the Advancement of Science*. Retrieved from <http://science.sciencemag.org/>
- Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What Can Machines Learn, and What Does It Mean for Occupations and the Economy? *AEA Papers and Proceedings*, 108, 43–47. <https://doi.org/10.1257/pandp.20181019>

- Brynjolfsson, E., Rock, D., & Syverson, C. (2017). Artificial intelligence and the modern productivity paradox: a clash of expectations and statistics. *NBER WORKING PAPER SERIES*. Retrieved from <http://www.nber.org/papers/w24001>
- Cappelli, R., Montobbio, F., & Morrison, A. (2018). Unemployment resistance across EU regions: the role of technological and human capital. *Papers in Evolutionary Economic Geography (PEEG)*. Retrieved from <https://ideas.repec.org/p/egu/wpaper/1831.html>
- Cerqua, A., & Pellegrini, G. (2014). Do subsidies to private capital boost firms' growth? A multiple regression discontinuity design approach. *Journal of Public Economics*, 109, 114–126. <https://doi.org/10.1016/j.jpubeco.2013.11.005>
- Charles, K. K., Hurst, E., & Notowidigdo, M. J. (2016). The masking of the decline in manufacturing employment by the housing bubble. *Journal of Economic Perspectives*, 30(2), 179–200. <https://doi.org/10.1257/jep.30.2.179>
- Chen, H., Li, X., Frank, M., Qin, X., Xu, W., Cebrian, M., & Rahwan, I. (2019). Automation Impacts on China's Polarized Job Market. *Arxiv*. Retrieved from <http://arxiv.org/abs/1908.05518>
- Christaller, W., & Baskin, C. (1966). *Central Places in Southern Germany*. Pp. 230. N.J.: Prentice-Hall, 1966. \$9.95. (E. Cliffs, Ed.). NJ, USA: Prentice-Hall.
- Clark, A. (2005). What Makes a Good Job? Evidence from OECD Countries. In *Job Quality and Employer Behaviour* (pp. 11–30). Palgrave Macmillan UK. https://doi.org/10.1057/9780230378643_2
- Coffey, W. (1979). Allometric growth in urban and regional social-economic systems. *Canadian Journal of Regional Science*, 11(1), 49–65. Retrieved from <http://www.cjrs-rcsr.org/archives/2-1/Coffey.pdf>
- Cortinovis, N., Xiao, J., Boschma, R., & van Oort, F. G. (2017). Quality of government and social capital as drivers of regional diversification in Europe. *Journal of Economic Geography*, 17(6), 1179–1208. <https://doi.org/10.1093/JEG/LBX001>
- Craglia, M., Annoni, A., Benczur, P., Bertoldi, P., Delipetrev, P., De Prato, G., ... Vesnic, A. L. (2018). *Artificial intelligence - A European perspective*. <https://doi.org/10.2760/936974>
- Criscuolo, C., Martin, R., Overmann, H., & Van Reenen, J. (2012). The Causal Effects of an Industrial Policy. *NBER Working Paper Series, Working Pa.*
- D P Reed. (2001). The Law of the Pack. *Harv Bus Rev*, 23–24. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/11213694/>
- Dahlin, E. (2019). Are Robots Stealing Our Jobs? *Socius: Sociological Research for a Dynamic World*, 5, 237802311984624. <https://doi.org/10.1177/2378023119846249>
- Davies, B., & Maré, D. C. (2019). *Relatedness, Complexity and Local Growth*. Retrieved from www.iza.org
- Davies, J. (2019). Creative industries research requires a new approach to data analysis. In *A Research Agenda for Creative Industries* (pp. 76–92). Edward Elgar Publishing. <https://doi.org/10.4337/9781788118583.00012>
- De Backer, K., Menon, C., Desnoyers-James, I., & Moussiégt, L. (2016). *Reshoring: Myth or Reality?* (OECD Science, Technology and Industry Policy Papers No. 27). *OECD Science, Technology and Industry Policy Papers* (Vol. 26). Paris. <https://doi.org/10.1787/5jm56frbm38s-en>

- Decker, M., Fischer, M., & Ott, I. (2017). Service Robotics and Human Labor: A first technology assessment of substitution and cooperation. *Robotics and Autonomous Systems*, 87, 348–354. <https://doi.org/10.1016/j.robot.2016.09.017>
- del Rio-Chanona, R. M., Mealy, P., Beguerisse-Díaz, M., Lafond, F., & Farmer, J. D. (2019). Automation and occupational mobility: A data-driven network model. Retrieved from <http://arxiv.org/abs/1906.04086>
- Delgado, M., Porter, M. E., & Stern, S. (2014). Clusters, convergence, and economic performance. *Research Policy*, 43(10), 1785–1799. <https://doi.org/10.1016/j.respol.2014.05.007>
- Dicken, P., & Lloyd, P. (1990). *Location in Space: Theoretical Perspectives in Economic Geography*. New York, New York: Harper & Row. Retrieved from <https://www.amazon.com/Location-Space-Theoretical-Perspectives-Geography/dp/0060416777>
- Dolfmanm, M. L., Inasco, M., & Holden, R. J. (2018). Healthcare jobs and the Great Recession. *BLS Monthly Labor Review*, 2018(6). <https://doi.org/10.21916/mlr.2018.17>
- Duranton, G., & Puga, D. (2003). Micro-foundations of urban agglomeration economies. *NATIONAL BUREAU OF ECONOMIC RESEARCH*. Retrieved from <http://www.nber.org/papers/w9931.pdf>
- Duranton, Gilles, & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. Vernon Henderson & Jacques-François Thisse (Eds.), *Handbook of Regional and Urban Economics* (1.4, pp. 2063–2117). Amsterdam: Elsevier. Retrieved from <http://cep.lse.ac.uk/~duranton>
- Eck, N. J. van, & Waltman, L. (2009). How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology*, 60(8), 1635–1651. <https://doi.org/10.1002/asi.21075>
- Ernst, E., Merola, R., & Samaan, D. (2018). *The economics of artificial intelligence: Implications for the future of work*. Retrieved from www.ilo.org/publns.
- Erwin, D. H. (2015). Novelty and innovation in the history of life. *Current Biology*, 25(19), R930–R940. <https://doi.org/10.1016/j.cub.2015.08.019>
- Essletzbichler, J. (2015). Relatedness, Industrial Branching and Technological Cohesion in US Metropolitan Areas. *Regional Studies*, 49(5), 752–766. <https://doi.org/10.1080/00343404.2013.806793>
- European Commission. (2008). *Employment in Europe 2008*. Brussels: European Commission, Directorate-General for Employment, Social Affairs and Equal Opportunities.
- European Commission. (2014). *Sixth report on economic, social and territorial cohesion: investment for jobs and growth - Promoting development and good governance in EU regions and cities*. Brussels: European Commission, DG for Regional and Urban Policy and DG for Employment, Social Affairs and Inclusion.
- Everett, M. G., & Borgatti, S. P. (2014). Networks containing negative ties. *Social Networks*, 38(1), 111–120. <https://doi.org/10.1016/j.socnet.2014.03.005>
- Fagerberg, J., & Verspagen, B. (2009). Innovation studies-The emerging structure of a new scientific field. In *Innovation, Economic Development and Policy: Selected Essays* (Vol. 38, pp. 218–233). Edward Elgar Publishing Ltd. <https://doi.org/10.1016/j.respol.2008.12.006>

- Fagiolo, G., Reyes, J., & Schiavo, S. (2010). The evolution of the world trade web: A weighted-network analysis. *Journal of Evolutionary Economics*, 20(4), 479–514. <https://doi.org/10.1007/s00191-009-0160-x>
- Farinha, T., Balland, P.-A., Morrison, A., & Boschma, R. (2019). What drives the geography of jobs in the US? Unpacking relatedness. *Industry and Innovation*. <https://doi.org/10.1080/13662716.2019.1591940>
- Feng, A., & Graetz, G. (2015). Rise of the Machines: The Effects of Labor-Saving Innovations on Jobs and Wages. Retrieved from <http://ftp.iza.org/dp8836.pdf>
- Feser, E. J. (2003). What Regions Do Rather than Make: A Proposed Set of Knowledge-based Occupation Clusters. *Urban Studies*, 40(10), 1937–1958. <https://doi.org/10.1080/0042098032000116059>
- Florida, R. (2012). *The Rise of the Creative Class--Revisited: 10th Anniversary Edition--Revised and Expanded*. Retrieved from <http://www.amazon.com/The-Rise-Creative-Class-Revisited-Edition-Revised/dp/0465029930>
- Florida, R., Mellander, C., & Stolarick, K. (2008). Inside the black box of regional development—human capital, the creative class and tolerance. *Journal of Economic Geography*, 8(5), 615–649. <https://doi.org/10.1093/JEG/LBN023>
- Foster-McGregor, N., & Pöschl, J. (2016). Productivity effects of knowledge transfers through labour mobility. *Journal of Productivity Analysis*, 46(2–3), 169–184. <https://doi.org/10.1007/s11123-016-0478-y>
- Foster-McGregor, N., Nomaler, Ö., & Verspagen, B. (2019). *Measuring the Creation and Adoption of New Technologies using Trade and Patent Data* (2019-053 No. ISSN 1871-9872). *UNU-MERIT Working Papers*. <https://doi.org/10.20955/r.85.67>
- Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., ... Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. *PNAS*. <https://doi.org/10.1073/pnas.1900949116>
- Frank, M. R., Sun, L., Cebrian, M., Youn, H., & Rahwan, I. (2018). Small cities face greater impact from automation. *Journal of The Royal Society Interface*, 15(139), 20170946. <https://doi.org/10.1098/rsif.2017.0946>
- Frenken, K., Van Oort, F., & Verburg, T. (2007). Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies*, 41(5), 685–697. <https://doi.org/10.1080/00343400601120296>
- Frey, C. B., & Osborne, M. A. (2013). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280. <https://doi.org/10.1016/j.techfore.2016.08.019>
- Fujiwara, Y., & Aoyama, H. (2010). Large-scale structure of a nation-wide production network. *The European Physical Journal B* 2010 77:4, 77(4), 565–580. <https://doi.org/10.1140/EPJB/E2010-00275-2>
- Gereffi, G., & M. Korzeniewicz. (1994). *Commodity Chains and Global Capitalism*. Westport, CT: Praeger. Retrieved from https://books.google.nl/books?hl=en&lr=&id=A86j9pWfTcAC&oi=fnd&pg=PR5&ots=glj-Y2MzU9&sig=6jQ3foYUDpYDaFVtHw-ay-b9VBM&redir_esc=y#v=onepage&q&f=false
- Glaeser, E. (2011). *Triumph of the City: How our greatest invention makes us richer, smarter, greener, healthier and happier*. London: Penguin Press. Retrieved from <https://www.penguinrandomhouse.com/books/303439/triumph-of-the-city-by-edward-glaeser/>

- Glaeser, E. L., Scheinkman, J. A., & Shleifer, A. (1995). Economic Growth in a Cross-Section of Cities. *Journal of Monetary Economics*, Volume 36, August 1995, Pp. 117-144. Retrieved from <http://www.nber.org/papers/w5013>
- Goodman, C. J., & Mance, S. M. (2011). Employment loss and the 2007–09 recession: an overview. *The 2007–09 Recession, BLS Monthly Labor Review*. Retrieved from <https://www.bls.gov/opub/mlr/2011/04/art1full.pdf>
- Goos, M., Manning, A., & Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8), 2509–2526. <https://doi.org/10.1257/aer.104.8.2509>
- Graetz, G., & Michaels, G. (2018). ROBOTS AT WORK. *The Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_00754
- Green, F. (2006). *Demanding Work: The Paradox of Job Quality in the Affluent Economy on JSTOR*. Princeton University Press. Retrieved from <https://www.jstor.org/stable/j.ctt4cgbjg>
- Green, F. (2008). Work Effort and Worker Well-Being in the Age of Affluence. *Political Science*. Retrieved from <https://www.semanticscholar.org/paper/Work-Effort-and-Worker-Well-Being-in-the-Age-of-Green/f3a9c1538fc53a9a75364fedc02309745252e44f#paper-header>
- Griliches, Z. (1969). Capital-Skill Complementarity. *The Review of Economics and Statistics*, 51(4), 465. <https://doi.org/10.2307/1926439>
- Grossman GM, & Helpman E. (1993). *Innovation and Growth in the Global Economy*. Cambridge, MA: The MIT Press. Retrieved from <https://mitpress.mit.edu/books/innovation-and-growth-global-economy>
- Guimei Liu, Limsoon Wong, & Hon Nian Chua. (2009). Complex discovery from weighted PPI networks. *Bioinformatics*, 25(15), 1891–1897. <https://doi.org/10.1093/BIOINFORMATICS>
- Hallward-Driemeier, M., & Nayyar, G. (2019). *Have Robots Grounded The Flying Geese?: Evidence From Greenfield FDI In Manufacturing* (Policy Research Working Papers). The World Bank. <https://doi.org/10.1596/1813-9450-9097>
- Harris, S. D., & Krueger, A. B. (2015). *Proposal for Modernizing Labor Laws for Twenty-First-Century Work: The “Independent Worker.”*
- Harter, J. K., Schmidt, F. L., & Keyes, C. L. M. (2004). Well-being in the workplace and its relationship to business outcomes: A review of the Gallup studies. In *Flourishing: Positive psychology and the life well-lived*. (pp. 205–224). American Psychological Association. <https://doi.org/10.1037/10594-009>
- Hasan, S., Ferguson, J. P., & Koning, R. (2015). The lives and deaths of jobs: Technical interdependence and survival in a job structure. *Organization Science*, 26(6), 1665–1681. <https://doi.org/10.1287/orsc.2015.1014>
- Hausmann, R., & Hidalgo, C. (2011). The Network Structure of Economic Output. *Springer*. Retrieved from <http://arxiv.org/ftp/arxiv/papers/1101/1101.1707.pdf>
- Hayek, F. A. (1945). The Use of Knowledge in Society. *The American Economic Review*, 35, 519–530. <https://doi.org/10.2307/1809376>
- He, C., Guo, Q., & Rigby, D. (2015). Industry Relatedness, Agglomeration Externalities and Firm Survival in China. Retrieved from <https://econpapers.repec.org/RePEc:egu:wpaper:1528>
- Hendler, J., & Golbeck, J. (2008). Metcalfe’s law, Web 2.0, and the Semantic Web. *Web Semantics*, 6(1), 14–20. <https://doi.org/10.1016/j.websem.2007.11.008>

- Henning, M. (2020). Regional labour flows between manufacturing and business services: Reciprocal integration and uneven geography. *European Urban and Regional Studies*, 27(3), 290–302. <https://doi.org/10.1177/0969776419834065>
- Henrich, J. (2015). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton: Princeton University Press. Retrieved from <https://press.princeton.edu/books/paperback/9780691178431/the-secret-of-our-success>
- Hidalgo, C., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science (New York, N.Y.)*, 317(5837), 482–487. <https://doi.org/10.1126/science.1144581>
- Hidalgo, C. (2015). *Why Information Grows: The Evolution of Order, from Atoms to Economies*. New York: Basic Books. Retrieved from <http://ecomplexity.ir/wp-content/uploads/2018/08/1.pdf>
- Hidalgo, C. A., Balland, P. A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., ... Zhu, S. (2018). The Principle of Relatedness. In *Springer Proceedings in Complexity* (pp. 451–457). Springer. https://doi.org/10.1007/978-3-319-96661-8_46
- Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26), 10570–10575. <https://doi.org/10.1073/pnas.0900943106>
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*, 15, 199–236. <https://doi.org/10.1093/pan/mpi013>
- Ho, D. E., Imai, K., King, G., Stuart, E. A., Abadie, A., Beck, N., ... Rubin, D. (2007). Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*, 15, 199–236. <https://doi.org/10.1093/pan/mpi013>
- International Labour Organization. (2014). *World of Work 2014: Developing with jobs*. Retrieved from http://www.ilo.org/global/research/global-reports/world-of-work/2014/WCMS_243961/lang--en/index.htm
- Isard, W. (1956). *Location and Space-Economy: A General Theory Relating to Industrial Location, Market Areas, Land Use, and Urban Structure*. Cambridge, Massachusetts: MIT Press.
- Jacobs, J. (1969). *The Economy of Cities*. Penguin Random House. Retrieved from <https://www.penguinrandomhouse.com/books/86059/the-economy-of-cities-by-jane-jacobs/>
- Jaimovich, N., & Siu, H. E. (2012). Job Polarization and Jobless Recoveries. *NBER Working Paper Series*. Retrieved from <http://www.nber.org/papers/w18334>
- James Manyika, Susan Lund, Michael Chui, Jacques Bughin, Jonathan Woetzel, Parul Batra, ... Saurabh Sanghvi. (2017). *Jobs lost, jobs gained: workforce transitions in a time of automation*. Retrieved from www.mckinsey.com/mgi
- Jones, C. I., & Romer, P. M. (2010). The new kaldor facts: Ideas, institutions, population, and human capital. *American Economic Journal: Macroeconomics*, 2(1), 224–245. <https://doi.org/10.1257/mac.2.1.224>
- Jun, B., Alshamsi, A., Gao, J., & Hidalgo, C. A. (2019). Bilateral relatedness: knowledge diffusion and the evolution of bilateral trade. *Journal of Evolutionary Economics*, 30(2), 247–277. <https://doi.org/10.1007/s00191-019-00638-7>

- Kaufman, B. E. (2010). The Theoretical Foundation of Industrial Relations and its Implications for Labor Economics and Human Resource Management. *ILR Review*, 64(1), 74–108. <https://doi.org/10.1177/001979391006400104>
- Kenneth J. Arrow. (1974). *The Limits of Organization*. New York: W. W. Norton & Company Ltd. Retrieved from <https://www.wwnorton.co.uk/books/9780393093230-the-limits-of-organization>
- Kernan, W. N., Viscoli, C. M., Makuch, R. W., Brass, L. M., & Horwitz, R. I. (1999). Stratified randomization for clinical trials. *Journal of Clinical Epidemiology*, 52(1), 19–26. [https://doi.org/10.1016/S0895-4356\(98\)00138-3](https://doi.org/10.1016/S0895-4356(98)00138-3)
- King, G., Nielsen, R., Coberley, C., & Pope, J. E. (2011). Comparative Effectiveness of Matching Methods for Causal Inference. *Working Paper*.
- King, G., Nielsen, R., Coberley, C., Pope, J. E., & Aaron Wells. (2011). *Comparative Effectiveness of Matching Methods for Causal Inference*. Healthways, Inc., 701 Cool Springs Blvd. Retrieved from <http://people.fas.harvard.edu/>
- Kline, P., & Moretti, E. (2013). Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics*, 129(1), 275–331. <https://doi.org/10.3386/w19293>
- Klinger, J., Mateos-Garcia, J., & Stathoulopoulos, K. (2018). Deep learning, deep change? Mapping the development of the Artificial Intelligence General Purpose Technology. *SSRN Electronic Journal*. Retrieved from <http://arxiv.org/abs/1808.06355>
- Kochan, T. (2016). What Was the Postwar Social Contract, Where Did It Come From, and What Made It Work for Three Decades? In T. Kochan (Ed.), *Shaping the future of work (15.662x)*. Cambridge: The MIT Press. Retrieved from <http://www.digitalhistory.uh.edu>
- Kogler, D. F., Rigby, D. L., & Tucker, I. (2013). Mapping Knowledge Space and Technological Relatedness in US Cities. *European Planning Studies*, 21(9), 1374–1391. <https://doi.org/10.1080/09654313.2012.755832>
- Kosfeld, R., & Werner, A. (2012). German Labour Markets—New Delineation after the Reforms of German District Boundaries 2007–2011. *Raumforschung Und Raumordnung*, 70(1), 49–64. <https://doi.org/10.1007/s13147-011-0137-8>
- Koski, H., & Pajarinen, M. (2013). The role of business subsidies in job creation of start-ups, gazelles and incumbents. *Small Business Economics*, 41(1), 195–214. <https://doi.org/10.1007/s11187-012-9420-5>
- Kremer, M. (1993). The O-Ring Theory of Economic Development. *The Quarterly Journal of Economics, JSTOR*, 108(3), 551–575. Retrieved from <https://notendur.hi.is/ajonsson/kennsla/O-ring.pdf>
- Krugman, P. (1996). *The Self Organizing Economy*. Cambridge, Massachusetts: Wiley-Blackwell. Retrieved from <https://www.wiley.com/en-us/The+Self+Organizing+Economy-p-9781557866981>
- Labianca, G. (2014). Negative ties in organizational networks. *Research in the Sociology of Organizations*, 40, 239–259. [https://doi.org/10.1108/S0733-558X\(2014\)0000040012](https://doi.org/10.1108/S0733-558X(2014)0000040012)
- Licklider. (1960). Man-Computer Symbiosis. Retrieved April 30, 2020, from <https://groups.csail.mit.edu/medg/people/psz/Licklider.html>

- Lin, J. (2011). Technological adaptation, cities, and New Work. *Review of Economics and Statistics*, 93(2), 554–574. https://doi.org/10.1162/REST_a_00079
- Lobo, J., Mellander, C., Stolarick, K., & Strumsky, D. (2014). The Inventive, the Educated and the Creative: How Do They Affect Metropolitan Productivity? *Industry and Innovation*, 21(2), 155–177. <https://doi.org/10.1080/13662716.2014.896602>
- Lordan, G., & Neumark, D. (2018). People versus machines: The impact of minimum wages on automatable jobs. *Labour Economics*, 52. <https://doi.org/10.1016/j.labeco.2018.03.006>
- Lösch, A. (1954). *The Economics of Location*. New Haven, CT, USA: Yale University Press.
- Lu, M. (2020). These are the jobs most at risk from COVID-19 transmission | World Economic Forum. Retrieved June 5, 2020, from <https://www.weforum.org/agenda/2020/04/occupations-highest-covid19-risk/>
- Lucas, R. E. (1993). Making a Miracle. *Econometrica*, 61(2), 251. <https://doi.org/10.2307/2951551>
- Lund, S., Manyika, J., Hilton Segel, L., Dua, A., Hancock, B., Rutherford, S., & Macon, B. (2019). The future of work in America People and places, today and tomorrow, (July), 124. Retrieved from www.mckinsey.com/mgi/publications/multimedia/
- Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., & Dewhurst, M. (2017). *A future that works: automation, employment, and productivity*. Retrieved from www.mckinsey.com/mgi.
- Markusen, A. (2001). *High-tech and I-tech : how metros rank and specialize*. Minneapolis MN: Humphrey Institute of Public Affairs University of Minnesota.
- Markusen, A. (2004). Targeting occupations in regional and community economic development. *Journal of the American Planning Association*, 70(3), 253–268. <https://doi.org/10.1080/01944360408976377>
- Markusen, A., & Schrock, G. (2006). The Distinctive City: Divergent Patterns in Growth, Hierarchy and Specialisation. *Urban Studies*, 43(8), 1301–1323. <https://doi.org/10.1080/00420980600776392>
- Markusen, A., Wassall, G. H., DeNatale, D., & Cohen, R. (2008). Defining the Creative Economy: Industry and Occupational Approaches. *Economic Development Quarterly*, 22(1), 24–45. <https://doi.org/10.1177/0891242407311862>
- Marshall, A. (1920). *Principles of Economics* (8 edition). London: Macmillan. Retrieved from <https://oll.libertyfund.org/titles/marshall-principles-of-economics-8th-ed>
- Martin, R., & Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of Economic Geography*, 6(4), 395–437. <https://doi.org/10.1093/JEG/LBL012>
- McAfee, A., & Brynjolfsson, E. (2016). Human Work in the Robotic Future | Foreign Affairs. Retrieved April 30, 2020, from <https://www.foreignaffairs.com/articles/2016-06-13/human-work-robotic-future>
- Mehta, A. (2014). 5 Ways to Lessen Inequality as Demand for Labor Decreases Worldwide | HuffPost. Retrieved April 30, 2020, from https://www.huffpost.com/entry/5-ways-lesson-inequality_b_6215708?guccounter=1
- Merritt, H. (2018, November 8). New technology and employment in Mexico. Retrieved from <https://ipn.elsevierpure.com/es/publications/new-technology-and-employment-in-mexico>
- Moretti, E. (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.

- Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, 5(1), 65–103. <https://doi.org/10.1257/app.5.1.65>
- Moretti, E., & Thulin, P. (2013). Local multipliers and human capital in the united states and sweden. *Industrial and Corporate Change*. <https://doi.org/10.1093/icc/dts051>
- Morgan, M. G., & Rosenberg, N. (1983). Inside the Black Box: Technology and Economics. *Journal of Policy Analysis and Management*, 3(1), 156. <https://doi.org/10.2307/3324062>
- Mori, T., Nishikimi, K., & Smith, T. E. (2008). THE NUMBER-AVERAGE SIZE RULE: A NEW EMPIRICAL RELATIONSHIP BETWEEN INDUSTRIAL LOCATION AND CITY SIZE*. *Journal of Regional Science*, 48(1), 165–211. <https://doi.org/10.1111/j.1467-9787.2008.00550.x>
- Morrison, A. (2008). Gatekeepers of Knowledge within Industrial Districts: Who They Are, How They Interact. *Regional Studies*, 42(6), 817–835. <https://doi.org/10.1080/00343400701654178>
- Morrison, A., Rabelotti, R., & Zirulia, L. (2013). When Do Global Pipelines Enhance the Diffusion of Knowledge in Clusters? *Economic Geography*, 89(1), 77–96. <https://doi.org/10.1111/j.1944-8287.2012.01167.x>
- Muneepeerakul, R., Lobo, J., Shuttters, S. T., Gómez-Liévano, A., & Qubbaj, M. R. (2013). Urban Economies and Occupation Space: Can They Get “There” from “Here”? *PLoS ONE*, 8(9). <https://doi.org/10.1371/journal.pone.0073676>
- NBER. (2010). *US Business Cycle Expansions and Contractions*. Cambridge. Retrieved from <http://www.nber.org/cycles/cyclesmain.html>
- Nedelkoska, L., Diodato, D., & Neffke, F. (2018). *Is Our Human Capital General Enough to Withstand the Current Wave of Technological Change?* Retrieved from https://growthlab.cid.harvard.edu/files/growthlab/files/humancapital_automation_cidrfwp93.pdf
- Nedelkoska, L., & Quintini, G. (2018). Automation, skills use and training. <https://doi.org/10.1787/2e2f4eea-en>
- Neffke, F. M. H., Otto, A., & Hidalgo, C. (2018). The mobility of displaced workers: How the local industry mix affects job search. *Journal of Urban Economics*, 108, 124–140. <https://doi.org/10.1016/j.jue.2018.09.006>
- Neffke, F., & Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, 34(3), 297–316. <https://doi.org/10.1002/smj.2014>
- Neffke, F., Henning, M., & Boschma, R. (2011). *How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. Papers in Evolutionary Economic Geography*.
- Neffke, Frank. (2017). *Coworker Complementarity*. *SSRN Electronic Journal*. Elsevier BV. <https://doi.org/10.2139/ssrn.2929339>
- OECD-NAEC. (2019). New Economic Policymaking - New Approaches for Economic Challenges. In *New Analytical Tools and Techniques for Economic Policymaking*. Paris. Retrieved from <http://www.oecd.org/naec/new-economic-policymaking/>
- OECD. (2018). *Good Jobs for All in a Changing World of Work. Good Jobs for All in a Changing World of Work*. OECD. <https://doi.org/10.1787/9789264308817-en>
- Olsen, K. M., Kalleberg, A. L., & Nesheim, T. (2010). Perceived job quality in the United States, Great Britain, Norway and West Germany, 1989-2005. *European Journal of Industrial Relations*, 16(3), 221–240. <https://doi.org/10.1177/0959680110375133>

- Perrault, R., Shoham, Y., Brynjolfsson, E., Clark, J., Etchemendy, J., Grosz Harvard, B., ... Mishra, S. (2019). *2019 annual report ar intelligence index Project Manager and Report Editor-in-Chief Steering Committee Steering Committee Introduction Report Highlights Acknowledgements*. Stanford. Retrieved from https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf
- Petralia, S., Baland, P. A., & Morrison, A. (2017). Climbing the ladder of technological development. *Research Policy*, 46(5), 956–969. <https://doi.org/10.1016/j.respol.2017.03.012>
- Porter, M. E. (1990). The Competitive Advantage of Nations. *Harvard Business Review*, 68(2), 73–93.
- Pumain, D., Paulus, F., Vacchiani-Marcuzzo, C., & Lobo, J. (2011). An evolutionary theory for interpreting urban scaling laws. *European Journal of Geography*. <https://doi.org/10.4000/CYBERGEO.2519>
- Rao, A., & Verweij, G. (2018). *Sizing the prize What's the real value of AI for your business and how can you capitalise?* Retrieved from <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf>
- Renski, H., Koo, J., & Feser, E. (2007). Differences in labor versus value chain industry clusters: An empirical investigation. *Growth and Change*, 38(3), 364–395. <https://doi.org/10.1111/j.1468-2257.2007.00375.x>
- Rigby, D. L. (2015). Technological Relatedness and Knowledge Space: Entry and Exit of US Cities from Patent Classes. *Regional Studies*, 49(11), 1922–1937. <https://doi.org/10.1080/00343404.2013.854878>
- Rodriguez, F., & A. Jayadev. (2010). The Declining Labor Share of Income. *UNDP Human Development Reports Research Paper, 2010/36*. <https://doi.org/10.1515/jgd-2012-0028>
- Roto J, S. J., & Claessen H. (2012). *Demographic trends in the Nordic local labour markets*. Nordic Centre for Spatial Development. Stockholm: Nordic Centre for Spatial Development. Retrieved from <https://nordregio.org/publications/demographic-trends-in-the-nordic-local-labour-markets/>
- Roux, D. (2018). Automation and employment: The case of South Africa. *African Journal of Science, Technology, Innovation and Development*, 10(4), 507–517. <https://doi.org/10.1080/20421338.2018.1478482>
- Sankar, S. (2012). The rise of human-computer cooperation | TED Talk. Retrieved April 30, 2020, from https://www.ted.com/talks/shyam_sankar_the_rise_of_human_computer_cooperation?language=en
- Shapiro, C., & Varian, H. R. (1998). *Information Rules: A Strategic Guide to the Network Economy*. Boston: Harvard Business School Press. Retrieved from <https://store.hbr.org/product/information-rules-a-strategic-guide-to-the-network-economy/863X>
- Shutters, S.T., Lobo, J., Muneeppeerakul, R., Strumsky, D., Mellander, C., Brachert, M., ... Bettencourt, L. M. A. (2018). Urban occupational structures as information networks: The effect on network density of increasing number of occupations. *PLoS ONE*, 13(5). <https://doi.org/10.1371/journal.pone.0196915>
- Shutters, Shade T., Muneeppeerakul, R., & Lobo, J. (2015). Quantifying urban economic resilience through labour force interdependence. *Palgrave Communications*, 1(1), 1–7. <https://doi.org/10.1057/palcomms.2015.10>

- Shutters, Shade T, Muneeppeerakul, R., & Lobo, J. (2015). How hard is it for urban economies to become 'green'? *Environment and Planning B: Planning and Design*, 43(1), 198–209. <https://doi.org/10.1177/0265813515600108>
- Shutters, Shade T, Muneeppeerakul, R., & Lobo, J. (2016). Constrained pathways to a creative urban economy. *Urban Studies*, 53, 3439–3454. <https://doi.org/10.2307/26151296>
- Sianesi, B. (2004). An evaluation of the Swedish system of active labor market programs in the 1990s. In *Review of Economics and Statistics* (Vol. 86, pp. 133–155). <https://doi.org/10.1162/003465304323023723>
- Statistics Canada. (2015). CMA and CA: Detailed definition. Retrieved July 9, 2020, from <http://www.statcan.gc.ca/pub/92-195-x/2011001/geo/cma-rmr/def-eng.htm>
- Steijn, M. (2018). *Improvement on the Association Strength: Implementing a Probabilistic Measure Based on Combinations without Repetition* (15 No. 1). Utrecht.
- Stokey, N. L. (1988). Learning by Doing and the Introduction of New Goods. *Journal of Political Economy*, 96(4), 701–717. <https://doi.org/10.1086/261559>
- Stokey, N. L. (1991). Human Capital, Product Quality, and Growth. *The Quarterly Journal of Economics*, 106(2), 587. <https://doi.org/10.2307/2937948>
- Szell, M., & Thurner, S. (2010). Measuring social dynamics in a massive multiplayer online game. *Social Networks*, 32(4), 313–329. <https://doi.org/10.1016/j.socnet.2010.06.001>
- Tanner, A. N. (2014). Regional Branching Reconsidered: Emergence of the Fuel Cell Industry in European Regions. *Economic Geography*, 90(4), 403–427. <https://doi.org/10.1111/ecge.12055>
- Thompson, W. (1985). From industries to occupations : rethinking local economic development. *Economic Development Commentary*.
- Thompson, W. R., & P. R. Thompson. (1987). Alternative Paths to the Revival of Industrial Cities. In G. Gappert (Ed.), *The Future of Winter Cities* (pp. 233–250). Newbury Park, CA: Sage. Retrieved from <https://books.google.nl/books?id=fWgFAQAIAAJ&q=%22Alternative+paths+to+the+revival+of+industrial+cities%22&dq=%22Alternative+paths+to+the+revival+of+industrial+cities%22&hl=en&sa=X&ved=2ahUKEwj2p8X6qsDqAhVQ-qQKHZokAwQQ6AEwA3oECAAQAg>
- Timmermans, B., & Boschma, R. (2014). The effect of intra-and inter-regional labour mobility on plant performance in Denmark: the significance of related labour inflows. *Journal of Economic Geography*, 14(2), 289–311. <https://doi.org/10.1093/jeg/lbs059>
- Tokuyama, T. (Ed.). (2008). Algorithms and Computation: 18th International Symposium, ISAAC 2007, Sendai, Japan, December 17–19, 2007. In *Algorithms and Computation* (pp. 1–1). Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-540-77120-3_1
- US Bureau of Labor Statistics. (2014, July 24). Occupational Employment Survey. Retrieved July 9, 2020, from <http://www.bls.gov/oes>
- US Census Bureau. (2015, December). Metropolitan and Micropolitan Statistical Areas. Retrieved July 9, 2020, from www.census.gov/population/metro/

- US Census Bureau. (2020). Metropolitan and Micropolitan Statistical Areas. Retrieved May 7, 2020, from <https://www.census.gov/programs-surveys/metro-micro/about.html>
- Warr, P. (1999). Well-being and the workplace. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 392–412). Russell Sage Foundation. Retrieved from <https://psycnet.apa.org/record/1999-02842-020>
- Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. (M. Granovetter, Ed.), *Canadian Journal of Criminology and Criminal Justice*. New York: Cambridge University Press. Retrieved from www.cambridge.org
- WEF. (2016). *The Future of Jobs: Employment, Skills and WAAorkforce Strategy for the Fourth Industrial Revolution. Growth Strategies*. <https://doi.org/10.1177/1946756712473437>
- Weiss, C. H. (1995). Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families. *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, (7), 65–92. <https://doi.org/10.1177/1356389003094007>
- Wikipedia. (2015). Beckstrom's law. Retrieved August 22, 2015, from https://en.wikipedia.org/wiki/Beckstrom's_law
- Winick, E. (2018). Every study we could find on what automation will do to jobs, in one chart. Retrieved August 19, 2019, from <https://www.technologyreview.com/s/610005/every-study-we-could-find-on-what-automation-will-do-to-jobs-in-one-chart/>
- Wisskirchen, G., Thibault, B., Bormann, B. U., Muntz, A., Niehaus, G., Soler, G. J., & Von Brauchitsch, B. (2017). Artificial Intelligence and Robotics and Their Impact on the Workplace. *IBA Global Employment Institute*, (April), 120.
- World Bank. (2013). *World Development Report 2013. World Development Report 2013*. The World Bank. <https://doi.org/10.1596/978-0-8213-9575-2>
- World Bank. (2019). *The changing nature of work*.
- Xue, Y., & Larson, R. C. (2015). STEM crisis or STEM surplus? Yes and yes. *Monthly Labor Review*, 2015(5). <https://doi.org/10.21916/mlr.2015.14>
- Youn, H., Bettencourt, L. M. A., Lobo, J., Strumsky, D., Samaniego, H., & West, G. B. (2016). Scaling and universality in urban economic diversification. *Journal of The Royal Society Interface*, 13(114), 20150937. <https://doi.org/10.1098/rsif.2015.0937>

NEDERLANDSE SAMENVATTING

Banen verdwijnen, banen worden gecreëerd, en banen veranderen op een geografisch ongelijkmatige manier. De arbeidsdynamiek is specifiek voor elke plaats. Dit proefschrift onderzoekt de manier waarop banen zich tot elkaar verhouden – gerelateerdheid – en hoe het de evolutie in de geografie van banen beïnvloedt. Ten eerste stellen we aan de hand van gegevens over de werkgelegenheid in meerdere landen vast dat grotere steden een meer dan proportioneel hogere dichtheid van gerelateerdheid tussen banen hebben. Ten tweede heeft gerelateerdheid de diversificatie van de arbeidsmarkt in steden in de Verenigde Staten bevordert, waardoor banen in steden worden samengebracht die complementair, gelijksoortig en/of synergetisch aan elkaar zijn – “magneeteffecten”. Ten derde verspreiden de effecten van automatisering zich via gerelateerdheid: de werkgelegenheidsgroei is groter voor banen die complementair maar niet gelijksoortig zijn aan lokale banen met een hoog risico op automatisering – “diffusie-effecten”. Ten slotte tonen we in een evaluatie van effecten van prikkels vanuit EU beleid in Portugese bedrijven aan dat beleid de kwaliteit van banen kan verbeteren door de dichte structuur van gerelateerdheid in innovatieve sectoren te stimuleren. Deze bevindingen helpen bij het ontwerpen van beleidsinstrumenten die de lokale verspreiding van negatieve effecten van automatisering neutraliseren en tegelijkertijd de positieve effecten bevorderen.

CURRICULUM VITAE

Teresa Farinha was born in Lisboa, Portugal. She studied in a music college from pre-school till university. During which she was introduced to the fascinating field of economics, and how scientific knowledge can help solve real problems. She holds a bachelor's degree in economics (2006) and a master's degree in monetary and financial economics (2010), both from Universidade de Lisboa. Parallel to her studies, she worked in audit, consulting, risk analysis, trade, as well as research in diverse organizations – research institutes, governmental agencies, NGO's – in Lisboa and in Manila, Philippines. During her Ph.D. in Utrecht University, she got a visiting student scholarship at MIT MediaLab, Massachusetts Institute of Technology, USA. She also lectured, supervised theses, and organized scientific events. Currently, she is a research fellow at United Nations University – MERIT, in Maastricht, The Netherlands.

Jobs disappear, jobs are created, and jobs change, in a geographically uneven way. Labour dynamics are particular to each place. Within each unique set of jobs in a city, workers influence each other in several ways. This doctoral thesis investigates the way jobs relate to each other – relatedness – and how it shapes the evolving geography of jobs. Our findings help design policy instruments that neutralize the local diffusion of negative impacts from automation, while promoting the positive ones.