Tasks, jobs and cities

Kok, Suzanne Jose

Tasks, Jobs and Cities

Suzanne Kok

Publisher: University of Groningen
Groningen, The Netherlands

Printed by: Ipskamp Drukkers B.V.

ISBN: 978-90-367-6451-3
eISBN: 978-90-367-6450-6

rijksuniversiteit
groningen

Tasks, Jobs and Cities

Proefschrift

ter verkrijging van het doctoraat in de
Economie en Bedrijfskunde
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
donderdag 31 oktober 2013
om 14.30 uur

door

Suzanne José Kok

geboren op 21 januari 1984
te ’s-Gravenhage

Promotores:	Prof. dr. S. Brakman
	Prof. dr. J.H. Garretsen
	Prof. dr. B. ter Weel
Beoordelingscommissie:	Prof. dr. J.P. Elhorst
	Prof. dr. A. de Grip
	Prof. dr. J.N. van Ommeren

Acknowledgements

Fortunately, I did not spend the last years on my own in an ivory tower. Here, I like to thank the key players of my PhD-years.

First and most of all, I thank my supervisors Bas ter Weel, Harry Garretsen and Steven Brakman. Bas is not only an inspiring and dedicated economist but also a great supervisor. He stimulated me to fully exploit every bit of the PhD-years and helped wherever he could. On the other side of the country, Harry and Steven always came up with new insights that made the travel worthwhile. Furthermore, I am thankful that my three captains hardly fought over the wheel. Thanks go out to the CPB Netherlands Bureau of Economic Policy Analysis and the SOM research school for offering me the opportunity to write this thesis.

Next, I am thankful that Paul Elhorst, Andries de Grip and Jos van Ommeren took the time to read the manuscript and provided useful comments. The different chapters furthermore improved thanks to suggestions of Yannis Ioannides, Jan Rouwendal, William Strange, Jasper de Jong, Andrea Jaeger and participants at several seminars and conferences.

I thank my colleagues in The Hague and Groningen for fruitful discussions and good coffees. A special thanks goes to my PhD-buddies Maaike, Tristan and of course roomy Sander. The down to earth guys of the 'dying S6' and the Bertha Babes of the 'Dames-lunches' brought additional laughter to my CPB-time.

In addition, I thank my friends, my parents, my Brabant family and the students of Chance2Study for the good life outside work. Jasper is the one who celebrates everyday life with me. I greatly enjoy his love, jokes and support and our adventures.

Thanks!
Suzanne Kok
August 2013, Den Haag

Contents

List of Tables

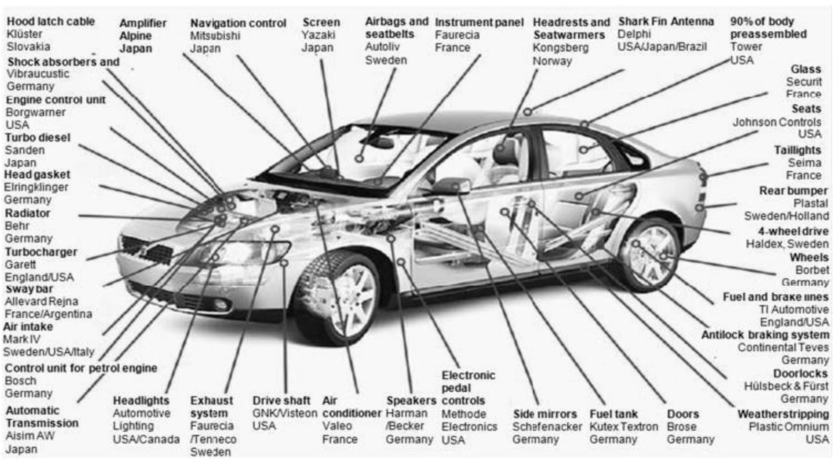
List of Figures

Chapter 1

Introduction

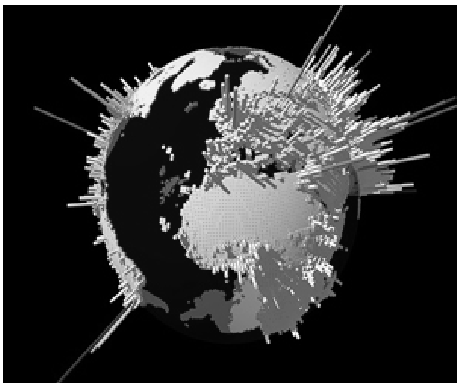
Cities remain a fascinating phenomenon. Why do 25 thousand people in Manhattan live on one square kilometre while the worldwide density could be 47 people per square kilometre? Why would you rent in Amsterdam while with a two and a half hour drive the rent price drops by a factor of 200? The improvements in information, communication and transport technologies eased the sprawl of people and economic activities during the last decades. Global supply chains emerged resulting in a relocating of separate production tasks across the world. The separate parts of the Volvo S40, for instance, are made in dozens of places (see Figure 1.1 below). Still, people and their activities cluster more and more together in the largest cities, as can be seen from the density spikes in Figure 1.2. Baldwin & Evenett (2012) talk about cities as 21st century 'factories'. What explains the facilitating role of cities in current production processes of millions of spatially separable tasks?

Figure 1.1. Global production chain Volvo S40



Source: Baldwin (2010)

Figure 1.2. Global population density



Source: Hill (2011)

Cities by definition bundle people and their economic activity. They exist because of proximity and scale benefits. The improvements in information, communication and transport technologies altered the bundling and unbundling of economic activity. Falling transport costs in combination with global technological trends such as computerisation made physical distance less relevant for some parts of production. This enabled the unbundling of production processes across workers, across firms and across locations. Some decades ago, a car was manufactured in one location. Today, the production of a car is a worldwide process. Some parts are best manufactured in China, others in Japan or Sweden. The changing spatial division of economic activity suggests altering location and city advantages. The geographically breaking-up of the production chain suggests that the worldwide division of labour takes place at a different level: the 'task' level instead of the industry level. Thus, at the level of 'repair', 'sewing' or 'design' instead of 'automobile industry'. New approaches to define these contents of work appear. The changing demand for tasks may vary within occupations as for instance computers substitute some tasks and complement other tasks of an occupation. This changing task demand indicates that a focus on occupations and industries does not capture the whole story anymore.

These observations indicate the relevance of a task approach in analysing the role of cities as 'factories of the 21st century'. This thesis describes the interactions between tasks, jobs and cities to create a better understanding of what kind of production cities continue to facilitate in the modern economy. The next pages introduce the theme, the outline of the main argument and the subjects of the thesis.

1.1 Bundling and unbundling of economic activity

Popular stories about the production of the Barbie doll, telecommunication chips and x-rays visualise current global production chains.[1] Although all continents participate in these chains, activity within these continents is spatially bundled. What drives the bundling and unbundling of economic activity?

Ages ago communities had to produce everything they consumed on their own. The extreme transport costs forced a geographical bundling of production and consumption. Spatial clustering of people generated possibilities to divide labour. Within a small town the butcher also needed to be the baker, while in a city these were two distinctive occupations. The spatial clustering of people induced scale benefits and enlarged the consumption possibilities. A larger local labour market meant a wider variety of consumption goods. A baker who devotes all his time to baking produces a wider variety of bread products than a baker who also has to cut the meat.

The development of cheaper ways to ship commodities gradually made trade more convenient. Production and consumption could be separated and communities no longer had to produce everything they wanted to consume. Together with the transport revolution, the emergence of mass-production and industrialisation resulted in a first wave of globalisation. Factories did not need to locate near their consumers and became footloose. Firms could now choose their location with respect to specific location and scale advantages. Cities, in their turn, started to produce products for which they had a comparative advantage to exploit specific location and scale advantages. Cities could focus, for instance, on the manufacturing of cars while importing computers. This resulted in a spatial division of labour across the world.

Cities exploit scale and location advantages. Not just aggregated activity is agglomerated but industries cluster in space as well. The spatial clustering of industries, such as the former auto-mobile industry in Detroit, seemed to generate agglomeration advantages.[2] These clusters attract firms in search of high productivity rates and ambitious workers in search of the best jobs. Many examples of clustering, such as the concentration of the film-industry in Hollywood, are or were pretty successful. Such successful examples induced many governments to strive for their own 'Hollywood'. A simple recipe on how to succeed as a city fails to appear so far.

[1] See Tempest (1996), Burrows (1995) and Pollak (2003).

[2] A broad and extensive literature discusses several agglomeration economies, for overviews see Glaeser & Gottlieb (2009) and Rosenthal & Strange (2004).

Successful cities vary substantially in their economic structure. Furthermore, specialisation increases lock-in risks and the sensitivity to a shock in the local industry. After the decline of the US auto-mobile industry for instance, the unemployment rate in Detroit surged.

The improving information, communication and transport technologies of the last decades initiated a new wave of unbundling. This 'second unbundling' is characterised by the fragmentation within the production process itself. Here, we focus on the bundling and unbundling of tasks in jobs (occupations) and in cities. First, these technological changes affect the bundling and unbundling of task packages of jobs. The necessity of performing several tasks by one worker disappeared by the easier and better coordination and communication possibilities altering the division of tasks across jobs. Additionally, computerisation has different effects on different tasks. Most tasks of a taxi driver are not replaced by a computer while many of the tasks of an office clerk are.

Second, economic activity is bundled and unbundled across space. Recent technological changes induced a fear for the relocation of 'our jobs' in rich countries to low wage countries such as China. Indeed, this relocation (offshoring) has been an influential economic force in the last 20 years (Feenstra, 2010). The estimates of how many jobs can potentially be offshored differ, but that more jobs become offshorable seems to be without discussion. Economic activity however still bundles in rich cities as well. Each part of the production process, each task, is produced at the most efficient place. For many tasks, this most efficient place continues to be a city in a rich country. The co-location of similar (or different) tasks, suppliers and consumers results in agglomeration advantages for the production of certain tasks. The economic focus of cities shifted from sectoral specialisation to functional specialisation. Location advantages vary across tasks and jobs. The design of computers and cars benefits, for instance, more from human interactions than the production of spare parts of computers and cars does. The co-location of suppliers is beneficial for the production of spare parts however.

It became possible to build global supply chains such as the one of the Barbie doll. Each part of the doll can be made by the most efficient worker, the most efficient firm and at the most efficient location. These trends affect the work content of jobs and of cities. The comparative advantages of workers and cities in rich countries shift. This has consequences for policies aiming to stimulate comparative advantages, e.g. educational policies, industrial policies and regional redistribution policies.

1.2 Main argument

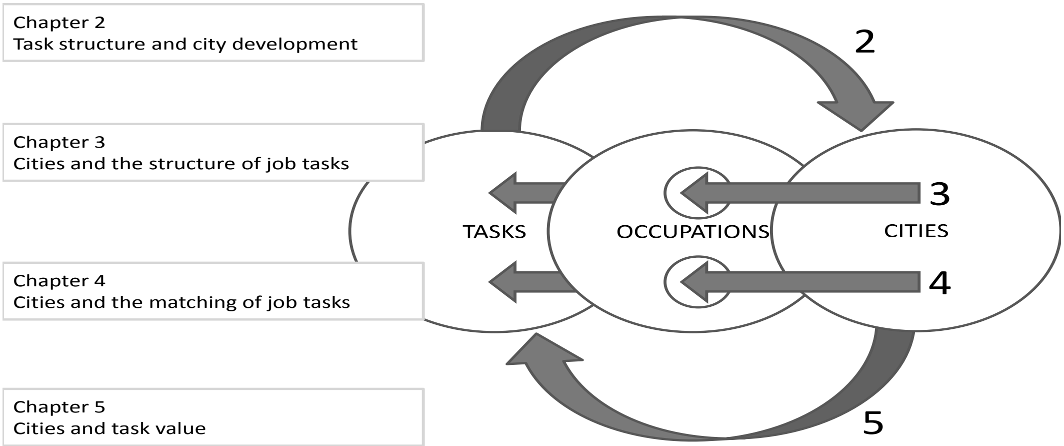
The search for the source of bundling of economic activity and people has a long and extensive history (Marshall, 1920; Jacobs, 1969; Jaffe et al., 1993). The rise of information, communication and transport technologies altered the division of labour substantially and with that the role of cities in the world economy. The changing impact of distance resulted in a vast stream of new studies about distance, cities and labour markets. According to Frances Cairncross (1997) and Thomas Friedman (2005) the ICT-revolution would eventually result in 'the death of distance' and a 'flat world'. The correlation between population density and productivity did however not weaken the last decades. The world does not seem to be a level playing field (Leamer, 2007). Bundling and unbundling of economic activity takes place across workers, firms and cities. This study applies a task approach to take into account the variations of tasks across workers and cities.[3] It is essential to draw the distinction to studies about the skill structure of cities, such as the work of Glaeser & Ressenger (2010). A city's task structure defines the work activities in the city that produce output. The skill structure of a city reflects the stock of capabilities the workers in the city apply to perform work activities. Since the assignment of skills to tasks changed over time, the analyses of cities' task structures differ from the ones of the skill structure (Autor, 2013). Considering the labour demand consequences this distinction between tasks and skills is relevant.

This study investigates what it is that cities facilitate within the current production process and focuses on what happens and does not happen within city jobs. By doing so, it relates to the field of urban economics. Therefore, the study does not take into account the relative spatial position of cities or rural areas. Our contribution lies in the task approach towards defining the role of cities. Specifically, this study analyses the interactions between tasks, jobs and cities. Figure 1.3 displays an overview of the considered interactions in this study.

The interactions between tasks, occupations and cities result in several research questions. A first focus is the connection between tasks and cities. How does the second unbundling affect city labour markets and what is the relation between task-structures and city development? Duranton & Puga (2005) and Baldwin (2010) argue that cities continue to benefit from similar agglomeration economies as before. The level at which these benefits take place changes. Where firms used to benefit from, for instance, the spatial clustering of similar industries, they currently be-

[3] See Acemoglu & Autor (2011) for an overview of this approach.

Figure 1.3. This thesis



nefit from the clustering of similar tasks or jobs. Agglomeration economies vary across tasks and jobs. Innovation seems to benefit heavily from clustering while bookkeepers do not perform their tasks more efficiently in close vicinity of many other bookkeepers. Autor & Dorn (forthcoming) show that the initial task structure of cities explains employment trends within cities. Cities with a large initial share of routine tasks polarised in terms of employment and wages. The distinction between non-routine and routine tasks does not explain spatial bundling and unbundling across cities however. Some routine tasks, such as cleaning, largely benefit from performance in close vicinity of certain other tasks. Agglomeration benefits likely vary across routine and non-routine tasks. The complementarity between agglomeration and modern technologies lies within the interconnection in the performance of tasks. Some tasks can be performed at isolated distance while other tasks require proximity to other tasks. This discussion results in the first question of this thesis:

1. *Does the connectivity between tasks explain employment development across cities?*

Tasks are not only interconnected within cities but also within jobs. The second considered connection is the one between cities and the task packages of jobs. Becker & Murphy (1992) argue that workers are more productive when their jobs contain fewer tasks. The more time a worker devotes to the performance of a certain task, the more task-specific skills he develops and the more efficient he becomes in the performance of that task. If the job of the worker contains many tasks, he has to divide his time across the tasks and has less time to 'learn by doing'. This makes the division of tasks across workers efficient. A doctor who has to perform many different kind of surgeries has to have many different surgical skills. When a doctor only removes appendices he will be better skilled for and more efficient in performing this task than when he also has to remove gallbladders. The division of labour is however bounded by coordination costs: the coordination of bundling the tasks of several people. When a patient has issues with both his appendix and his gallbladder and has two separate surgeons, they have to communicate intensively to decide upon the best strategy. Communication and coordination costs are lower when the patient has one surgeon who has surgeon skills for both the appendix and the gallbladder.

The thick labour markets of cities likely increase specialisation possibilities. The idea is quite simple: in a larger market the total number of tasks can be divided across more workers.[4] Hence, the number of tasks performed by each worker decreases with the size of the market. Duranton & Jayet (2011) extend this idea by showing that scarce tasks are performed more often in large markets. The large demand and supply in these markets make it possible to produce a wider variety of products. The assumed more extensive division of labour in cities provides workers the possibility to specialise in a subset of tasks and develop specific skills for this subset. This results in the following question:

2. Do workers in large cities specialise in a smaller subset of tasks and develop more specialised skills than workers located outside these cities?

Not only the number of tasks workers perform determines their efficiency, also the characteristics of these tasks do matter. People are most productive when they perform tasks that match their skills well. The assignment of heterogeneous workers to heterogeneous job tasks is a complex process however. This assignment faces frictions caused by imperfect information and an extensive amount of heterogeneous workers and jobs. Developments in information and communication technologies altered the assignment of skills to tasks.[5]

[4] This idea was already suggested by for instance Adam Smith (1776) and James Baumgardner (1988a). Empirical evidence is however scarce.

[5] See the work of Acemoglu & Autor (2011) and Autor (2013).

The thick labour market of large cities results in more choices for both job-seeking employees and for employee-seeking employers. The extensive choice leads to several labour market advantages. The chance to find a decent job or a decent new employee increases when there are more in the market. In case of a shock or a change in preferences, workers and employers can switch between jobs easier when located in a thick labour market. This reduces risk-aversion. Because it is easier to switch, turn-overs are higher in thick markets. 'Job-hopping' is associated with beneficial knowledge spillovers, but a high turnover also leads to a loss of firm-specific knowledge. A large market furthermore makes it possible to be more choosy who to pick. This fastidiousness takes place at both sides of the assignment and therefore results in better matches between heterogeneous workers and heterogeneous jobs. Several studies theoretically argue that large local markets should provide better assignments of workers to jobs (Helsley & Strange, 1990; Kim, 1990, 1991). Empirically, this literature is however rather unexploited [6] which results in a third question:

3. Does the thick labour market in large cities result in better matches of heterogeneous workers to heterogeneous job tasks?

Lastly, we consider the connection between city structures and the value of communication job tasks. The relatively smooth communication and coordination between people in cities seems to be a driving force behind current successful cities (Gaspar & Glaeser, 1998). Together people produce more knowledge and ideas than when working isolated, since people learn from each other by watching, observing and especially interacting.[7] Benefiting from watching, observing and especially interacting requires certain communication, social and emphatic skills. A very smart person cannot by definition interact easily with other persons. Clarification, patience and trust are examples of important communication skills. Since the main advantage of cities lies in their ability to combine economic activity and the knowledge of people, communication and coordination skills are key. Bacolod et al. (2009) show that the returns to certain skills, such as social skills, increase with city size.

As noted before, different types of (successful) city economies co-exist. This suggests variation in location advantages for the bundling of economic activity. Typically, city economies either focus on a certain activity or they perform a very wide

[6] Within a first empirical paper, Petrongolo & Pissarides (2006) find positive scale effects in both post-employment wage and reservations wages.
[7] Jaffe et al. (1993) show for example that distance bounds patent citations.

range of activities. In the literature this distinction is referred to as specialised and diversified cities (Duranton & Puga, 2000). These cities provide different agglomeration advantages for firms. Within specialised cities firms benefit from the co-location of similar, competing firms. This co-location lowers production costs and the focus of firms in these cities lies on producing their products as cheap as possible. Within diversified cities, firms benefit from the co-location of a wide variety of firms. The production costs are higher in these cities due to smaller sharing benefits of specific facilities and a specialised labour market. However, by watching, observing and interacting in a diversified environment firms benefit from a broad spectrum of knowledge and idea spillovers. This environment is especially favourable to young firms and products which are still in development. More mature firms and products flourish in specialised cities with low production costs (Duranton & Puga, 2001).

When we relate the co-existence of specialised and diversified cities to the importance of communication and coordination in today's cities, two hypotheses come up. First, knowledge and idea spillovers foster location advantages in both city types which suggests that communication is relevant in both types. Second, the communication in diversified environments is more important as firms are still in development. Furthermore, the communication across fields makes communication more complex than the communication between similar firms in specialised cities. This results in the following question:

4. Are communication tasks equally valuable in specialised and diversified cities?

1.3 Outline of the thesis

This thesis discusses the interactions between tasks, occupations and cities to provide insight in the facilitating role of cities in current production. The four relations displayed in Figure 1.3 and discussed in the previously introduced questions form a guideline for this study. This section presents the outline of the thesis.

Chapter 2 investigates the impact of cities' initial task structure on recent employment trends across US cities. In previous studies, these employment trends are mostly explained by differences in industrial structure or skill structure. The worldwide shift in the allocation of tasks indicate the relevance of a task approach. The chapter studies to what extent tasks benefit from the geographical presence of other tasks. Empirically, this so-called task connectivity is measured by the co-agglomeration of 41 job tasks from the Occupational Information Network (ONET) dataset. The connectivity between the performed tasks defines the economic structure of the 168 largest US cities. Our analyses show that cities with a one standard deviation higher 'connectivity' in 1990, experienced an additional employment growth of 30 to 45 percent of a standard deviation between 1990 and 2009. The interdependence of tasks seems to be an important determinant of recent employment trends across US cities. This result is observed within manufacturing, within services and across all skill groups.

Chapter 3 underlines the relevance of a task approach in analysing local labour markets by showing that task packages of jobs vary across space. This chapter is an empirical investigation of the theory of James Baumgardner (1988a) on the division of labour across space. It defines an extension of the model to frame the hypothesis that the division of labour is more extensive in cities than in towns. The model assumes that a more extensive division of labour results in additional development of skills as workers devote their time to a smaller subset of important job tasks. The chapter applies this idea empirically using the German labour force survey BIBB. Jobs in large German cities indeed contain different task packages compared to the the same jobs in small cities. Workers in large cities perform a smaller subset of tasks than their counterparts in towns. This focus allows them to develop more specific skills for their core tasks. As expected, jobs demand more cognitive skills when they are performed in a large city.

Not only the contents of jobs vary across space, also the skills of the workers do. Chapter 4 therefore investigates the connection between cities and the matching of worker skills to job tasks. The thick labour markets of large cities are expected to result in tighter matches between heterogeneous workers and heterogeneous jobs. The empirical evidence for this hypothesis is scarce. Chapter 4 compares the quality of assignment of heterogeneous workers to heterogeneous tasks between thick and thin Dutch labour markets. The framework of this chapter suggests that tighter labour matches drive spatial sorting of 'better' workers and 'better' jobs into thicker labour markets. These 'better' workers and jobs are more sensitive to bad matches as they have more to lose relatively to other workers and jobs. Empirically, these hypotheses are tested employing the Dutch LISS panel. As expected, the suitability of the skills of workers for their job is better in cities than in the Dutch countryside. When we control for the self-selection into jobs, the spatial variation remains. Workers with relatively many cognitive skills are over-represented in Dutch cities. The same holds for jobs with relatively many cognitive tasks. Additional analyses indicate that workers mainly sort into cities for job opportunities in the Netherlands. Workers in cities improve the suitability of their skills for their job faster than workers in the countryside. However, this does not explain the spatial differences in the matching quality. Lastly, chapter 4 presents positive wage returns of the better matching quality in cities. Thick labour markets in the Netherlands result in more productive matches between heterogeneous workers and heterogeneous tasks.

A task approach is a useful method to understand the underlying mechanisms of why some cities have fared well and others have not. However, different types of cities have fared well the last decades. Diversified and specialised city structures successfully co-exist (Duranton & Puga, 2000). The variation in city structures likely results in variation in the value of certain tasks in these cities. Therefore, Chapter 5 studies the relation between city structure and task value. Cities prosper because of the importance of proximity and human interactions in certain parts of production processes. Communication jobs tasks which stimulate human interactions are a clear example of an important task in cities. Chapter 5 analyses the value of communication tasks in both diversified and specialised cities. The chapter employs wage, employment and task information from the 168 largest US cities. It shows that performance of communication job tasks has positive wage returns in all US cities. As expected, these returns are higher in diversified cities where complex knowledge spillovers are relatively more important. The size of the wage returns decreases with the specialisation level of the city. We conclude that the importance of complex interactions in production decreases with the specialisation level of the city as well.

Lastly, chapter 6 combines the results of this thesis into concluding remarks. It summarises chapters 2 to 5 and discusses the interplay between tasks, jobs and cities. Since this thesis answers only a limited number of questions and since this research also raises new questions, the thesis ends with an agenda for further research.

CHAPTER 2

CITIES, TASKS AND SKILLS*

2.1 Introduction

The division of labour has changed over the past two decades. Technological change and especially rapid progress in information and communication technologies (ICT) has enabled a break-up of the production process, which has had implications for the organisation of work and the structure of employment (Bresnahan et al., 2002; Autor et al., 2003). ICT has changed the way individual tasks can be carried out and created new possibilities for communication between workers. Not only is this observed within and between firms, but also across space. In many cases, physical distance becomes less important for production because communication at distance can be as effective as communication in person (Bloom et al., 2009). At the same time cities flourish because of the increasing importance of human interactions in modern production processes (Glaeser & Maré, 2001). These trends have been accompanied by new approaches to relax the implicit equivalence between workers' skills and the tasks that have to be carried out at work. The core feature of these approaches is that workers apply their skills to tasks in exchange for wages. This distinction between skills and tasks becomes important when the assignment of skills to tasks is evolving with time, because the set of tasks demanded in the economy is altered by technological change that changes the need for proximity. Recent evidence on offshoring suggests that certain tasks have been more vulnerable to offshoring than others (Grossman & Rossi-Hansberg, 2008) and that a task-based approach to study these developments is worthwhile pursuing (Baldwin & Nicoud, 2010; Acemoglu & Autor, 2011).

This chapter uses such an approach to document and interpret recent employment trends in the 168 largest US metropolitan areas in the period 1990-2009. These

* This chapter is based on joint work with Bas ter Weel.

cities cover about 75 percent of total US employment in 2009. Employment trends across cities are often explained by differences in industrial structure (Ellison & Glaeser, 1997). Duranton & Puga (2001) and Desmet & Rossi-Hansberg (2009) argue that new industries cluster in expensive locations to benefit from knowledge spillovers, while more mature industries relocate to less expensive places because production processes have become more standardised. The main benefit of a task-based approach is that it allows us to analyse how employment growth across cities is caused by interactions among job tasks and agglomeration forces. Understanding this mechanism is potentially important in explaining why some cities have fared well while others have been in decline.

We first present a simple framework in which we show how the decision to trade tasks alters the division of labour. Occupations are defined as bundles of tasks. The main focus of the framework is to understand how tasks are connected with each other. Connectivity explains to what extent tasks are benefiting from the presence of other tasks and to what extent tasks are clustering. Empirically, our measure of task connectivity measures the importance of proximity or co-agglomeration for 41 job tasks defined in the Occupational Information Network (ONET) database. This survey classifies all occupations in terms of the importance of job tasks.

We analyse employment growth in the period 1990-2009. We construct a database of the 168 largest US cities in which we pair representative data on job task requirements from the ONET database with samples of employed workers from the Current Population Survey and Census to form a consistent panel of industry, occupations and spatial task input.

Our main results can be summarised as follows. Our measure of task connectivity explains a significant part of the changes in employment in US cities over the last two decades. We find that a one standard deviation increase in task connectivity increases employment by 30 to 45 percent of a standard deviation. Cities with a larger share of connected tasks have grown faster relative to other cities, conditional on initial employment and location characteristics. Other measures of the task composition of cities, such as the spatial concentration of tasks, do not explain growth patterns. The result is robust to the inclusion of differences in the structure of employment or industries (Glaeser et al., 1992; Ellison & Glaeser, 1997; Glaeser & Kerr, 2009), the rise in the importance of social skills (Bacolod et al., 2009), the routinisation and computerisation of some parts of employment (Autor et al., 2003). In addition, the connectivity between tasks is important in both manufacturing and service sectors and for all skill groups.

This chapter is related to a relatively new and growing body of empirical research documenting and interpreting changes in the structure of employment and wages using a task-based approach in which worker skills are allocated to job tasks. Contributions to this way of analysing trends have been made by Autor et al. (2003), Autor et al. (2006), Borghans & Ter Weel (2006), Goos & Manning (2007), Goos et al. (2009), Firpo et al. (2009) and Criscualo & Garicano (2010). They show across a variety of data sources that certain types of occupations seem to be disappearing in terms of employment shares and/or seem to be paying lower wages over time, while others grow and obtain wage growth. Duranton & Puga (2005) focus on a related issue by distinguishing sectoral and functional specialisation of employment. Acemoglu & Autor (2011) review these international trends and argue that a task-based approach is helpful when the assignment of worker skills to job tasks is evolving with time, either because shifts in market prices command reallocation of skills to tasks or because the set of tasks demanded in the economy is changed by technological developments, trade, or offshoring. We add to these approaches a spatial dimension because reallocation of skills to tasks changes the division of tasks across space too.[1] This helps to understand the employment developments across different types of cities.

By addressing the spatial dimension of employment our work is related to the recent contributions of Glaeser & Maré (2001), Bacolod et al. (2009), Bacolod et al. (2010), Autor & Dorn (forthcoming) and Florida et al. (2012). They document trends in regional employment and show that the structure of employment reveals path dependence. In addition, some tasks seem to be associated with employment growth, while others predict declines. Especially human capital seems to be important for growth. We use human capital too and find that it is an important determinant of employment growth across cities. We add to this that the structure of employment in terms of task combinations seems to be even more important. Our arguments and findings are related to the empirical work on the division of labour across space. Duranton & Jayet (2011) show, based on occupations, that the distribution of workers across occupations in dense urban areas is different relative to more rural areas. We show that the connectivity of tasks is positively correlated to employment growth.

Finally, the importance of cities in gluing tasks together is also used in the approaches developed in Jensen & Kletzer (2005) and Akcomak et al. (2011). They

[1] Rosenthal & Strange (2004) and Glaeser & Ressenger (2010) extensively review the literature in urban economics.

develop and apply measures of task connectivity similar to the ones we use here. Charlot & Duranton (2004) and Bacolod et al. (2009) emphasise the importance of communication and social skills in city employment. Our measure includes this mechanism but also emphasises the combination of tasks associated with employment changes. The research by Duranton & Puga (2001), Duranton & Puga (2005) and Desmet & Rossi-Hansberg (2009) points to the complementary relationship between cities and technological change in explaining changes in employment structure. We implicitly use this argument to explain why some tasks can be placed at distance.

This chapter proceeds as follows. In the next section the theoretical background is presented which results in a measure of task connectivity and predictions for city size. Section 2.3 documents the most salient details of the data sources and presents the empirical strategy. In Section 2.4 the main estimation results are shown. Section 2.5 discusses other measures of task-composition and Section 2.6 applies the analyses to several sub-samples. Section 2.7 concludes.

2.2 Theoretical background

Before documenting the impact of a cities' initial task structure on employment growth, this section lays the background for our indicator on the task structure of cities. Within the light of the recent rapid development of ICT, this framework suggests what tasks will be carried out in close vicinity and what types of tasks will be performed at distance. This has repercussions for employment growth in and outside cities.[2]

2.2.1 Basic setting

Firms decide upon the division of tasks across workers and across space. Production in large cities is more expensive, but also comes with a set of well-known positive agglomeration forces. Firms are assumed to be small relative to the market and take wages (w) as given. The market price (p) of the output in this economy is also given.

Human capital is multidimensional. Workers produce output by performing tasks. The performance of different tasks requires a different set of skills. Changes in technology and labour supply determine what tasks are performed by what

[2] The focus of this chapter lies on the impact on city employment. Although interesting, rural areas are outside the scope of this paper.

types of workers. The connectivity between different tasks determines what tasks are performed in what location. Given the state of technology, firms take the connectivity between different tasks (tc) as given.

Workers are able to perform many different tasks. The production function of the firm describes the time needed to carry out these tasks. Connectivity determines the time requirements d to carry out the tasks.

To illustrate the mechanism of location choices we have in mind, we assume, for convenience, that all occupations consist of two tasks, task a (t_a) and task b (t_b). The firm maximises profits per unit of production:

$$max_{tc} \quad p - w(tc)\big(d_a(tc) + d_b(tc)\big). \tag{2.1}$$

For an individual worker the total time needed to produce one unit of output equals:

$$d(tc) = d_a(tc) + d_b(tc). \tag{2.2}$$

The time needed to produce one unit of output depends on task connectivity where we assume that more connected tasks can be carried out faster because they more naturally combine into one occupation.

Assume that t_b can be placed at distance and that t_a needs to be performed inside the firm. When t_b is performed outside the firm it takes Δ_b instead of d_b to perform t_b. The performance of t_a also changes because the firm has to coordinate with another firm or plant in a different location or even a different country. At the same time, the outside firm or plant is more efficient in performing t_b, which makes production more efficient. It now takes Δ_a to carry out this task, with $\Delta_a \neq d_a$. Total time to produce one unit of output now equals:

$$\Delta = \Delta_a + \Delta_b. \tag{2.3}$$

The decision to place t_b at distance depends on the costs per unit of output. There are two inputs in these costs: the difference in wage costs $\big(w(d_a + d_b)$ vs. $w(\Delta_a + \Delta_b)\big)$ and the cost advantage of producing at a different location, which we label c. The break-even point at which $w(d_a + d_b) = w\big((\Delta_a + \Delta_b) + c(\Delta_a + \Delta_b)\big)$ equals:

$$x = w\Big(\frac{d_a + d_b}{\Delta_a + \Delta_b} - 1\Big). \tag{2.4}$$

If $x < 0$, t_b will be performed at distance because the actual costs of producing in a different location are below the break-even costs. The term in brackets in equation (2.4) represents the time gain of dividing tasks across locations. It depends on (i) the character of the tasks to be carried out and on (ii) the connectivity with other tasks. $\mu_t(tc) = -(\frac{(\partial d_t(tc)/\partial tc)}{d_t(tc)})$ is the time for task t saved by a marginal increase in tc. t_a is a more connected task relative to t_b if the time saved to perform this task is larger, i.e. $\mu_a(tc) > \mu_b(tc)$.

Differences in connectivity across tasks have employment effects. Specifically, places with with a larger share of highly connected tasks will face employment growth and places with a larger share of only loosely connected tasks will face declines in employment. Many of the tasks historically performed by production and administrative workers have become automated or have been offshored. This does not make these tasks obsolete because they are now performed much more efficiently in other places (or by computers). Since these tasks are necessary to produce output, the more efficient performance is beneficial for the more connected tasks in the cities at home (Grossman & Rossi-Hansberg, 2008).

2.2.2 What is connectivity?

Whether or not tasks will be placed at distance depends on three facets of the division of labour. First, it depends on the time lost with the coordination of this task relative to the gains of the division of tasks across locations. This balance has been changing over the last decades as a result of technological change. Improved communication technologies reduce the time lost communicating when placing tasks at distance (Duranton & Jayet, 2011). In addition, technological change affects the organisation of work. The division of production time might change, which changes the decision on the division of tasks across workers and locations (Borghans & Ter Weel, 2006; Garicano & Rossi-Hansberg, 2006). Finally, worker skills could complement or substitute for computer technology. Some tasks could be taken over by computer technology, which also changes the performance of other tasks (Borghans & Ter Weel, 2004). Lower coordination costs induce a further division of tasks across locations where it is most cost effective to carry out the work.

Second, it depends on the nature of the tasks. Some tasks are non-tradable and cannot be done at distance at reasonable costs, e.g. cleaning the offices in the headquarters. The difference between $(d_a + d_b)$ and $(\Delta_a + \Delta_b)$ is infinite in such cases. Hence, in all cities we observe the presence of a certain number of basic tasks that have to be carried out in close vicinity. This is a similar argument as the one noted in Autor et al. (1998), who find that computerisation has a detrimental effect on low-skilled workers, but not at the very low end because some low-end service occupations are unaffected by this type of technological change.

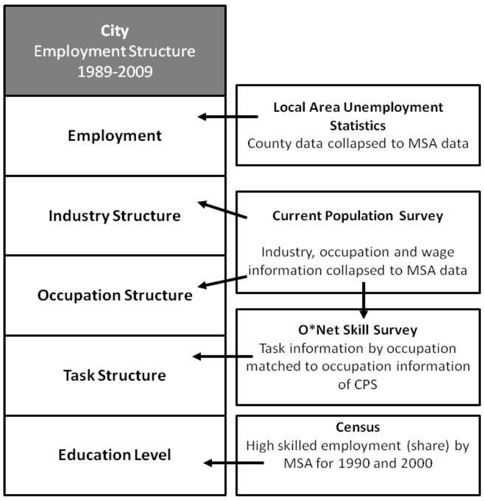
Third, tasks are connected in cities because of the existence of agglomeration forces. Coordination costs in terms of sharing inputs and transmitting information and knowledge are lower when tasks are performed closely together (Duranton & Puga, 2004). Tasks for which input sharing and information and knowledge transmission are important complement other tasks and connect in space. This seems especially true for tasks that demand higher levels of skill (Glaeser & Ressenger, 2010), tasks that require more coordination and face-to-face interactions (Gaspar & Glaeser, 1998; Blum & Goldfarb, 2006) and knowledge tasks (Von Hippel, 1994). Bacolod et al. (2009) and Florida et al. (2012) show that urban wage premiums tend to be higher for analytical and social tasks and lower for physical and technical tasks. Charlot & Duranton (2004) argue that larger and more educated cities require workers to communicate more. They find support for this hypothesis in a sample of French firms and show that workers who communicate more earn higher wages. Agglomeration disadvantages such as congestion costs limit the size of the city and cause smaller cities to have lower rent costs.

2.3 Data

We combine the information from several data sources to construct a database on the division of tasks in the 168 largest US cities in the period 1990-2009. The tasks in the database are broadly defined and could be performed in all occupations and industries. The construction of the database from the several sources is visualised in Figure 2.1.

The main indicators for the cities' division of labour are collected from the Current Population Survey (CPS). For about 140,000 individuals we obtain information about occupation, industry and city of residence (defined as Metropolitan Statistical Area (MSA)) in each year between 1990 and 2009. We distinguish 326 three-digit occupations and 142 three-digit industries. Cities are defined as MSAs, as the classification of MSAs is based on the nature of their economic activity. In 2009 MSAs were responsible for more than 85 percent of the employment, income, and production of products and services in the United States (Bureau of Economic Analysis, 2009). The MSA definitions, in terms of borders, change over time. This complicates analyses of employment developments of cities. To be able to analyse a consistent sample of cities and a sound match between several data sources, MSAs are defined as combined counties following the 1990 definition. The borders of counties are consistent over time. A sample of 168 MSAs is obtained.

Figure 2.1. Database construction



Information about job tasks is collected from the ONET Database. ONET provides information about the importance of abilities, interest, knowledge, skills, work activities, work context and work values within occupations. The work activities represent the job tasks of the worker. ONET distinguishes 41 broadly defined work activities. All tasks could be performed within all 326 occupations and are therefore not industry-related. Examples are thinking creatively, scheduling work and activities and processing information. For each occupation the importance of the 41 tasks is provided by ONET on a scale from 1 (not important at all) to 5 (extremely important). The importance of the tasks by occupation are matched to the occupations observed in the CPS. Aggregating the task information at the city level generates the division of tasks by city over time (1990-2009). Table A.1 in Appendix A.3 lists all tasks and presents information on type, employment shares and connectivity.

ONET categories the work activities into four groups: information input, mental processes, work output and interacting with others. The second column of Table 2.1 presents an example of a task within each group, columns (1) to (4) show the average importance of the four task-groups within four broad occupational groups. The average task importance varies between 2.24 (work output for clerical and sale occupations) and 3.58 (information input for production and operator occupations). Information input tasks define where and how the information and data are gained that are needed to perform the job. These tasks obtain high importance levels in all occupational groups. Mental processes tasks indicate what processing, planning, problem-solving, decision-making and innovating activities are performed in the occupation. These tasks are especially important within professional, managerial and technical occupations. The standard errors of the importance of these tasks are very low within the broad occupational groups. The work output tasks refer to 'what physical activities are performed, what equipment and vehicles are operated/controlled and what complex/technical activities are accomplished as job outputs'. The production and operators occupations obtain the highest importance of these tasks. Lastly, the importance of interacting with others is the most important within professional, managerial and technical occupations.

The last two rows of Table 2.1 present the employment shares of the occupations in respectively 1990 and 2009, while the last two columns present these for the task-groups. Professional, managerial and technical occupations obtain both the highest employment share in 1990 and the highest employment growth between 1990 and 2009. Also service occupations grow in terms of employment share while the shares of clerical, sales, production and operators occupations decline. These findings are consistent with the findings of Acemoglu & Autor (2011). Information input is the largest task-group in terms of employment while interacting with others and mental processes experience the largest growth. Remarkably, the changes between occupational groups are larger than the changes between task-groups.

The division of tasks across US cities is constructed from the matching of task information to occupations. We have to assume that the task structure of jobs does not differ by city characteristics. This is a strong assumption but a necessary one because we only observe the task content of occupations once. Bacolod et al. (2009) and Autor et al. (2003) face the same problems. For example, a car mechanic in Detroit conducts the same tasks relative to a car mechanic in New York. The extent of the market might however affect the task package of workers and generate specialisation possibilities (Baumgardner, 1988a). If this is the case, these differences are caused by the extent of the city suggesting that all the tasks are still performed

Table 2.1. Task importance by broad occupational groups

		Occupational groups				Employment share	
		Professional, managerial, technical	Clerical, sales	Production, operators	Service	1990	2009
Task group	Task example	(1)	(2)	(3)	(4)	(5)	(6)
Information input	Getting information	3.52	3.09	3.58	3.37	28.02	27.89
		(0.63)	(0.73)	(0.34)	(0.41)		
Mental processes	Processing information	3.64	3.13	3.03	2.99	26.73	27.05
		(0.29)	(0.36)	(0.30)	(0.33)		
Work output	Handling and moving objects	2.44	2.24	3.04	2.51	21.04	20.46
		(0.85)	(0.89)	(0.57)	(0.61)		
Interaction with others	Assisting and caring for others	3.12	2.80	2.60	2.88	24.21	24.69
		(0.46)	(0.56)	(0.43)	(0.44)		
Employment share 1990		29.73	30.31	24.55	15.42		
Employment share 2009		36.18	25.52	20.95	18.15		

Note: the task groups refer to the ONET classification. Importance is measured on a scale from 1 (not important at all) to 5 (extremely important). A cell shows the average importance, and the deviation herein, of the tasks of a task group within a broad occupation group. The four broad occupational groups are defined as in Acemoglu & Autor (2011).

within the city. This would not affect our measurement of connectivity of tasks within the city. In addition, the ONET data is based on data collected in 1998 (released in 2001). This means that we only have of a cross-section of task data at our disposal, which implies that the time variation in the division of tasks is based on the employment development of individual occupations. To deal with this issue, the task structure of cities in the initial year (1990) is used to document and interpret employment changes. This is similar to the approach taken in Autor et al. (2003), who use the Dictionary of Occupational Titles of 1977 to explain employment changes from 1963 onwards. We discuss the consequences of this approach in Section 2.6, where we examine the robustness of our approach and estimates in more detail.

Next, employment figures for cities over the period 1990 to 2009 are collected from the Local Area Unemployment Statistics from the Bureau of Labor Statistics (BLS). Lastly, a city's share of high-skilled inhabitants is gathered from the Census Decennial Database. High-skilled workers are defined as those workers with at least a bachelor's degree.

Data Appendix A discusses the data sources and provides insight in the construction of the classification of cities, industries, occupations and tasks. Data Appendix A.3 also includes a list of all variables, their aggregation level and their sources.

2.3.1 Descriptive statistics

The database we use for the empirical analysis contains information on the division of labour and other characteristics of the 168 largest US cities. Table 2.2 presents the summary statistics of the core variables used in the empirical analysis. Table A.3 in Appendix A.3 presents the correlation coefficients. Cities vary in terms of characteristics such as size, skill level and economic structure. Figure 2.2 shows the development of the division of the four task categories over time. The importance of tasks is measured as its share in the city's total task importance of all 41 tasks for all workers. We have set 1990 to zero. Relative to 1990 the employment share of mental processes tasks and interacting with others tasks has risen. This is consistent with the observations of Borghans et al. (2006), Borghans et al. (2008) and Bacolod et al. (2009) that interpersonal skills gain importance. The employment shares of information input tasks and especially work output tasks decreased during the period. The shock in 2001 results from a change in definition of Metropolitan Areas.

Over time employment in this sample of cities grows. Urbanisation seems to go

Table 2.2. Summary statistics

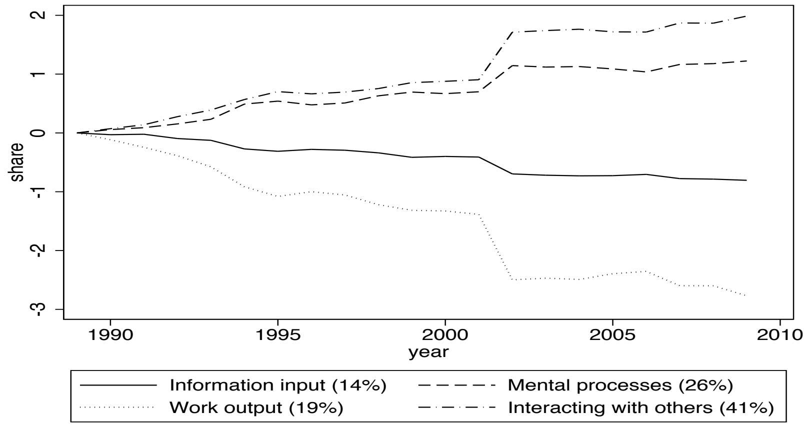
	Mean	SD	Min	Max
Employment 1990-2009	0.00	1.00	-1.85	3.60
Employment	0.00	1.00	-1.74	2.89
Connectivity	0.00	1.00	-1.87	2.77
Industrial specialisation	0.00	1.00	-2.09	2.90
Labour suitability	0.00	1.00	-2.29	1.75
Social skills	0.00	1.00	-4.17	3.14
Computer use	0.00	1.00	-2.50	3.24
Share high skilled	0.00	1.00	-1.78	4.02
Rent	0.00	1.00	-0.99	4.92
January temperature	0.00	1.00	-1.65	2.72
July temperature	0.00	1.00	-1.79	1.34
North-east	0.12	0.33	0.00	1.00
Midwest	0.25	0.43	0.00	1.00
South	0.41	0.49	0.00	1.00
West	0.23	0.42	0.00	1.00

Note: n=168 cities. All variables are measured in 1990. Variables are measured as described in Table A.2 in Appendix A.3.

along with increased city size: in 1990 about 65 percent of the US population lived in one of the 168 largest cities; in 2009 this share has risen to almost 75 percent. In addition, the relatively larger cities in our sample of 168 are growing faster than the relatively smaller cities. The rank size of cities is fairly stable with Los Angeles, New York and Chicago being the top 3 (more than 4,000,000 employees in 2009). At the bottom the same cities turn up in both 1990 and 2009. Table 2.3 lists the five largest and five smallest cities in our list of 168 cities in 1990 and 2009. The next columns list the five fastest growing cities and the five slowest growing or shrinking cities in the period 1990-2009 both in absolute numbers of employees and in percentages. Phoenix, Atlanta, Houston, Washington and Las Vegas are the fastest growers, adding over 500,000 employees between 1990 and 2009. On the other hand, Detroit is shrinking in both absolute and percentage terms relatively fast.

The skill level of the largest US cities varies too. Boulder-Longmont, Washington and San-Francisco form the top 3 of high-skilled cities over the whole time period. In these cities more than 40 percent of the workforce is high skilled, which holds for only 10 to 12 percent of the workforce in the lowest educated cities. The average share of high-skilled workers in cities increases from 20 to 24 percent in the sample period. Computer use (in terms of average importance in occupations)

Figure 2.2. Division of employment of six broad task groups over time[a]

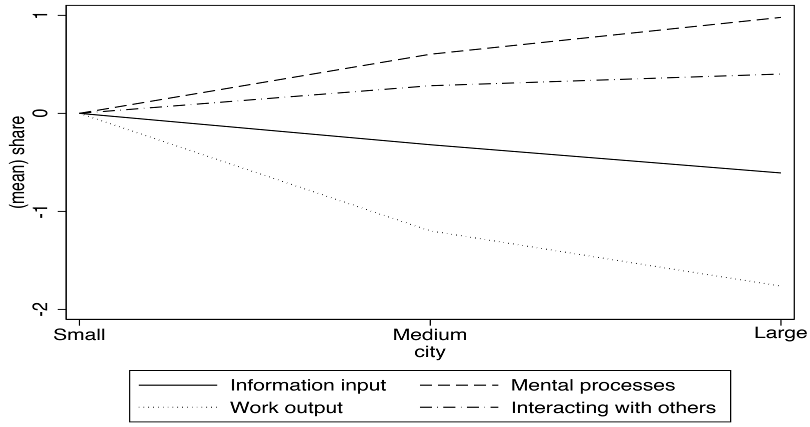


[a] The development is normalised to the employment share in 1990. The actual employment share in 1990 is in brackets. Changing city definitions cause the shock in 2001.

varies across US cities as well. In 1990 computer use was valued the most in the occupations in Huntsville, San Jose and Washington (all with an average importance above 2 on a scale from 0 to 4). As we only have cross-section information on the importance of computer use, the change of computer use over time is only based on the change in the division of labour across occupations. The average importance of computer use increases slightly from 1.82 to 1.85 which indicates that occupations for which computer use is relatively important in 1990 increase slightly in terms of employment share.

Finally, Figure 2.3 shows the division of tasks across city sizes. We define three size classes. Small cities employ less than 250,000 workers, medium-sized cities have a working population in between 250,000 and 1 million, and large cities employ over 1 million workers. Relative to small cities, in larger cities mental processes and interaction with others tasks seem to be more important, while work output and information input tasks are less important.

Figure 2.3. Division of tasks across city sizes (1990) [a]



[a] The employment share of tasks is normalised to zero at the employment share in small cities. Small sized cities have less than 250,000 employees, medium cities between 250,000 and 1,000,000 employees and large cities more than 1,000,000 employees.

Table 2.3. The largest, smallest, fastest growing and shrinking MSAs

Size (number of workers)		Growth	
1990	2009	Employment (number of workers)	Percentage
Largest MSAs		Fastest growers	
Los Angeles	Los Angeles	Phoenix-Mesa	Las Vegas
(4,259,705)	(4,328,589)	(814,075)	(11.42)
New York	New York	Atlanta	McAllen-Edinburg-Mission
(3,745,220)	(4,256,376)	(792,870)	(107.07)
Chicago	Chicago	Houston	Provo-Orem
(3,645,767)	(4,000,905)	(630,134)	(85.66)
Boston	Boston,	Washington	Fayetteville-Springdale-Rogers
(2,910,471)	(3,101,796)	(606,593)	(85.38)
Philadelphia	Philadelphia	Las Vegas	Austin-San Marcos
(2,355,639)	(2,454,509)	(524,178)	(81.99)
Smallest MSAs		Slowest growers	
Pueblo	Florence	Detroit	Hickory-Morgantown
(48,728)	(60,580)	(-178,313)	(-12.19)
Florence	Monroe	New Orleans	Benton Harbor
(58,064)	(66,048)	(-50,632)	(-10.52)
Waterloo-Cedar Falls	Jackson	San Jose	Binghamton
(58,862)	(66,162)	(-37,472)	(-9.37)
Fort Walton Beach	Pueblo	Dayton-Springfield	Detroit
(62,143)	(67,660)	(-28,604)	(-9.15)
Monroe	Benton Harbor	Newark	New Orleans
(62,704)	(67,730)	(-21,371)	(-9.08)

2.3.2 Measuring connectivity

To bring the theoretical approach to the data, we need to measure the extent to which tasks are connected to each other. To do so, we follow Akcomak et al. (2011)

and construct a measure of task connectivity based on correlations between observed patterns of task combinations across different cities. It measures the probability of the presence of a task if another task is also present in that city. To empirically measure connectivity between tasks we proxy employment shares by information about the importance of job tasks within the occupations. The occupation of a worker provides information about the importance of 41 tasks scaled from 1 (not important at all) to 5 (extremely important). The total time devoted to worker tasks in city l is measured by the sum of all task scores of all workers in city l. The employment share of a task in a city is equal to the share of task scores of that task in the total task score of the city ($\tilde{E}_l$). Task connectivity for task t is constructed as follows:

$$tc_t = \sum_{t'=1}^{t'=41} c\big(\tilde{E}_{t,l} \mid \tilde{E}_{t',l}\big) \quad for \quad t' \neq t. \tag{2.5}$$

The measure is a task specific indicator. It is based on a task-city matrix of the employment shares of tasks within cities. In equation (2.5) the term $c\big(\tilde{E}_{t,l} \mid \tilde{E}_{t',l}\big)$ represents the correlation between the estimated employment shares of task t and task t' in cities. We use this correlation as a measure of the extent to which task t and task t' are connected with each other in cities. Or, in terms of the agglomeration literature, the extent to which they co-agglomerate (Ellison et al., 2010). The higher the value of this measure, the more task t and task t' are found to be performed together in space. The sum of the connectivity with all other tasks generates the spatial connectivity of task t. The higher tc_t, the more task t is connected in space to other tasks and the more expensive it becomes to place this task at distance.

The connectivity measure provides the highest levels of task connectivity for tasks such as provide consultation and advice to others and interpreting the meaning of information for others. These tasks are relatively strongly correlated with other tasks in space and face the lowest probability to be placed at distance. By contrast, tasks such as handling and moving objects and repairing and maintaining mechanical equipment have the lowest level of task connectivity. These tasks could be done relatively easy at distance from other tasks.

Table 2.4 shows the task connectivity measure for a sample with five cities and eight tasks. Although the differences between employment shares in these tasks are rather small, there is a spatial pattern in task connectivity. Spatially, the tasks getting information, processing information, scheduling work and activities and developing and building teams co-agglomerate. The same holds for handling and moving objects together with controlling machines and processes. The higher the employment share of the first four tasks, the lower the share of the second group of tasks. Cities either obtain a relatively high share of information input tasks and interacting with others tasks or a relatively high share of work output tasks. The work output tasks obtain a negative connectivity as they are loosely connected to the performance of the other tasks. Especially the task developing and building teams depends on the co-location of several other tasks. Data Appendix A.3 presents a list of the employment shares of all 41 tasks and their levels of task connectivity. At the level of single tasks most tasks of the group interacting with others obtain high connectivity levels, which is consistent with the analysis of Bacolod et al. (2009). The task connectivity of city employment is defined as follows:

$$C_l = \sum_{t=1}^{t=41} tc_t * \tilde{E}_{t,l}. \tag{2.6}$$

The task connectivity level of the city reflects the average connectivity of the (estimated) employment of tasks. The last column in the example of Table 2.4 presents the connectivity of the task employment of the five cities for eight tasks. The performed tasks in Boston are the most connected while in Los Angeles and Detroit the tasks are relatively loosely connected.

Table 2.4. Example of task connectivity of eight tasks and five cities

	Information input		Mental processes		Work output		Interacting with others		Connectivity
	Getting information	Monitor, processes, materials, or surroundings	Processing information	Scheduling work and activities	Handling and moving objects	Controlling machines and processes	Assisting and caring for others	Developing and building teams	
Boston	18.88	15.17	15.48	13.67	12.65	11.37	12.78	13.27	0.72
Dallas	18.87	15.14	15.37	13.68	12.72	11.51	12.71	13.25	0.43
Detroit	18.58	15.15	14.98	13.21	13.37	12.05	12.65	13.05	-1.56
Los Angeles	18.80	15.15	15.19	13.50	13.01	11.75	12.60	13.16	-0.42
New York	18.95	15.08	15.35	13.72	12.65	11.23	13.01	13.21	0.83
Task connectivity	2.06	-0.04	2.33	2.16	-2.23	-2.02	0.94	2.33	

Note: cells represent task-city employment shares (defined in Section 2.3). Task connectivity is measured as in equation (2.5) for these eight groups in these five cities. Connectivity measures the connectivity of the city employment as defined in equation (2.6). Both measures are standardised with a mean of zero and a standard deviation of one.

2.4 Results

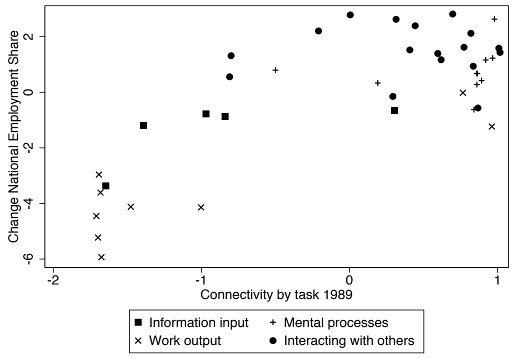
Because our database contains only 168 observations we first display simple graphical analyses of the task connectivity of cities and several bi-variate patterns in the data. We continue by adding regression analyses for our task connectivity measure in Section 2.4.2. Next, in Section 2.4.3 we add several co-variates to the analysis. The sensitivity of these results is tested in Section 2.5 which discusses various other measures of the task structure of cities and Section 2.6 which shows the results for different samples.

2.4.1 Graphical analyses

Figure 2.4 plots the (standardised) measures of task connectivity for all 41 tasks against changes in employment shares of these tasks in the period 1990-2009. Task connectivity is defined at the task level indicating the spatial correlation between the performance of tasks (see equation (2.5)). Each dot represents a task. The figure displays a positive correlation between task connectivity and subsequent employment change, which suggests that more connected tasks have gained in terms of employment shares over the last two decades. The correlation coefficient equals 0.75 and is significant at the one percent level. The different markers in Figure 2.4 represent the four different types of tasks as defined by ONET. Task connectivity is relatively high among the different interacting with others tasks and mental processes tasks. Among most work output tasks the connectivity is low, exceptions are interacting with computers and documenting/recording information. Information input tasks are more scattered. Table A.1 in Appendix A.3 presents the task connectivity for all 41 tasks.

Figure 2.5 provides information about the characteristics of cities and consists of five panels. The horizontal axis measures the standardised task connectivity in 1990 for the 168 cities in our sample and the vertical axis the standardised log of employment in 1990. Here, task connectivity is measured at the city level. It reflects the spatial correlation between the performed tasks within city employment as defined in equation (2.6). The dots in all five panels are cities, the markers define several city characteristics. Panel A presents a scatter plot of the correlation between task connectivity and city size. The correlation coefficient (standard error) between the two variables equals 0.88 (0.00). Florence, Visalia-Tulare-Porterville, Johnstown, Fort Wayne and Pueblo are the cities with the lowest task connectivity. Boston, New York, Chicago, Washington and Los Angeles obtain the highest connectivity

Figure 2.4. Task connectivity and change in employment share [a]



[a] Dots represent the 41 tasks. The correlation is 0.75 (0.00) and significant at the 1 percent level. The task connectivity measure is calculated following equation (2.6). The values are standardised with a mean of zero and a standard deviation of one. The correlations differ by task group. For the information input tasks the correlation is 0.66 (0.23), for mental process tasks 0.11 (0.77), for work output tasks 0.83 (0.01) and for interacting with others tasks -0.02 (0.95).

in their task employment. In the four remaining panels we split the sample of 168 cities according to different characteristics.

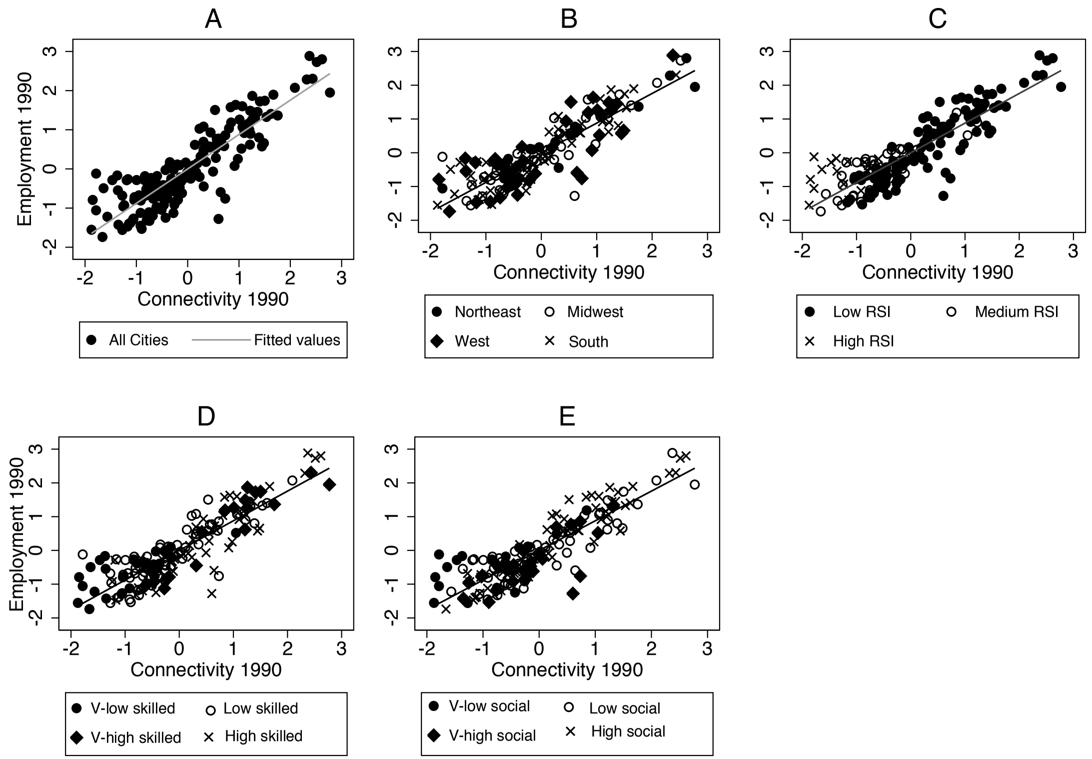
Panel B splits the sample into different regions. We have defined four regions: the North-east, the Midwest, the South and the West. The figure does not return a clear pattern; cities with relatively high shares of connected tasks are not spatially concentrated in the United States.

Differences in the industrial structure of cities partly explain the development of cities (Glaeser et al., 1992; Henderson et al., 1995). A useful measure to account for such differences is the relative specialisation index (RSI). The level of specialisation measures the over-representation of an industry within a city relative to other cities. We define the RSI index using the employment shares E for industry j and city l:

$$RSI_l = max\big(log(E_{j,l}) - log(E_j)\big). \tag{2.7}$$

RSI_l measures industry j's employment share in the city ($E_{j,l}$) relative to the share of the industry in national employment (E_j). A high specialisation level indicates that the city economy is relatively focussed towards a certain industry, such as the focus of Detroit on the car industry. The correlation between task connectivity and RSI_l equals -0.59 (0.00). In Panel C we again present the correlation between task connectivity and employment, but characterise cities by different categories of relative specialisation. We have split the sample into three categories using the stand-

Figure 2.5. The relationship between task connectivity and employment in 1990 [a]



[a] Dots represent the 168 cities. Table A.2 in Appendix A.3 displays measures and sources of all variables and Table A.3 presents the correlations. Equation (2.6) defines the measure for the task connectivity of cities. Both employment and task connectivity are standardised with a mean of zero and a standard deviation of one. The regional division represents the Census regional division. RSI is defined by the standardised value of equation (2.7). A low RSI represents a negative score, medium RSI a score between 0 and 1 and high RSI a score above 1. Skill categories are defined by the deviations from the mean. Very low-skilled cities have a high-skilled employment share below 14.6 (one standard deviation from the mean of 20.2 percent). Low-skilled cities have a share between 14.6 and 20.2, high-skilled cities between 20.2 and 25.8 and very high-skilled cities above 25.8. The social categories are obtained with the same strategy. Cities with a very low-social share have a share of social tasks below 12.7, low-social cities a share between 12.7 and 13.02, high-social cities between 13.02 and 13.32 and high-social cities above 13.32.

ardised RSI: not specialised cities (a negative deviation from the mean), medium specialised cities (a small positive deviation from the mean) and highly specialised cities (more than one standard deviation above the mean). The picture suggests that the highly specialised cities are the ones with the lowest level of task connectivity. This seems plausible, since specialisation means a strong division of labour with less tasks being carried out at home and more tasks being outsourced to other places. For the two measures of lower levels of specialisation there is no clear pattern in the data in relation to task connectivity.

The structure of human capital in cities explains another major part of the development of cities (Glaeser & Maré, 2001; Glaeser & Ressenger, 2010; Moretti, 2004; Berry & Glaeser, 2005; Venables, 2011). Glaeser & Ressenger (2010) document that mainly cities with a relatively high-skilled population benefit from agglomeration economies. Connected tasks turn out to be more likely to be performed by relatively high-skilled workers. The importance of connected tasks for performing a job ranges (on a standardised scale) from 0.018 for high-school graduates to 0.125 for workers with at least a bachelor degree. In Panel D of Figure 2.5 we have split the sample of cities according to skill level. There are four categories defined based on the deviations from the mean: very low-skilled cities (less than 14.6 percent of the employees is skilled), low-skilled cities (between 14.6 and 20.2 percent is skilled), high-skilled cities (between 20.2 and 25.8 percent is skilled) and very high-skilled cities (more than 25.8 percent is skilled). The picture shows that cities with a more highly skilled workforce obtain a higher level of task connectivity. The correlation between the share of high-skilled workers and task connectivity equals 0.48 (0.00).

Finally, Panel E addresses the importance of social skills. Recent work by Charlot & Duranton (2004), Bacolod et al. (2009) and Florida et al. (2012) suggests that people skills are important in explaining the success of cities. The existence and wealth of dense areas indicates that interaction is valuable. Social or people skills ease interaction and are therefore more valued in larger cities (Bacolod et al., 2009). In terms of our analysis this could imply that our measure of connectivity picks up social skills. We define social (or people) skills by the share of the ONET social skills in city employment. The task connectivity of city employment slightly correlates with the share of social skills (0.17 (0.03)). Panel E of Figure 2.5 shows no very clear-cut pattern when discriminating between the importance of social skills across cities to explain the correlation between task connectivity and employment in 1990. It seems that task connectivity is not only picking up the effect of social skills on employment.

Cities with a highly connected task structure tend to be larger, less specialised, more skilled and perform more social skills than cities with a low task connectivity. In the next subsection we distinguish between different city characteristics and their impact on growth.

2.4.2 Regression results

We estimate a number of specifications in which we explain changes in employment across our sample of cities ($\Delta E_{90-09,l}$) by our connectivity measure in the initial year ($C_{90,l}$), location characteristics (L_l) and a set of covariates in the initial year ($X_{90,l}$). The equation we estimate is:

$$\Delta E_{90-09,l} = \alpha_0 + \alpha_1 E_{90,l} + \alpha_2 C_{90,l} + \alpha_3 L_l + \alpha_4 X_{90,l} + \epsilon_l, \tag{2.8}$$

where l is an index for cities, α_0 is a constant term, $E_{90,l}$ is the initial employment and ϵ_l an error term with the usual assumptions. The summary statistics of the variables are shown in Table 2.2.

Table 2.5 presents the results of estimating a number of straightforward regression models. We estimate the determinants of the employment growth of cities between 1990 and 2009. We find that a one standard deviation increase in task connectivity increases employment by 30 to 45 percent of a standard deviation. We include initial employment (in logs) in all models. This always returns negative and significant coefficients, which suggests a tendency towards convergence in city size in our sample. In the estimates presented in column (1) of Table 2.5 we show the effect of task connectivity on employment growth. The coefficient is positive and significant. The interpretation of the coefficient is that a one standard deviation increase in connectivity increases the growth of the employment by about 43 percent of a standard deviation or about 144,000 employees.

The second column of Table 2.5 includes common controls for location characteristics. Three main trends determined the growth of cities the last decades. First, cities with a high level of human capital grew faster than relatively low-skilled cities (Glaeser & Ressenger, 2010; Eeckhout et al., 2010). Second, workers were attracted to the warmer, drier places in the US. The rise of the 'Sunbelt' is associated with capital accumulation (Caselli & Coleman, 2001), improvements in the political institutions and local policies (Besley et al., 2010) and consumption amenities (Mueser & Graves, 1995; Rappaport, 2007). And lastly, public transport routes became less important for city development (Glaeser & Shapiro, 2003). As a counterforce, density in cities often results in congestion and higher costs of living and especially housing (Moretti, 2013). To capture these main trends we add the city's share of high-skilled workers, housing prices, January and July temperature and regional dummies to the regressions in column (2). Consistent with the results obtained by Glaeser & Ressenger (2010) and Eeckhout et al. (2010), cities with a one standard deviation higher share of high-skilled workers grow about 18 percent of a standard deviation faster. The cost of housing decreases the growth of cities: a one standard deviation higher housing price results in about 45 percent of a standard deviation lower employment growth. The coefficient of July temperature is significant and positive while January temperature does not affect employment growth. Given temperature, the western part of the US experienced the highest growth. Adding our measure of task connectivity to an estimation with only the location variables increases the adjusted R-square from 0.418 to 0.432. The task connectivity of the employment in the city seems to have an additional and sizeable impact on employment growth.

2.4.3 Other city-structure indicators

Next, we add various other city-structure indicators to the analyses. Columns (3) to (8) in Table 2.5 present the results and the relations between these indicators and task connectivity and employment growth. We also visualise these relationships in Figures 2.6 and 2.7.

First, our results might be driven by differences in industrial structure of the city (e.g., Glaeser et al. (1992), Henderson et al. (1995)). Besides the previously used relative specialisation index we define the local industrial structure by the labour pool suitability as in Glaeser & Kerr (2009). The labour pool suitability index measures the quality of the city's employment in terms of it's industrial structure. The Glaeser-Kerr index for city l is defined as follows:

$$GK_l = -\sum_j E_j \left(\sum_o |E_{j,o} - (\sum_j E_{j,l} E_{j,o})| \right). \tag{2.9}$$

The index measures the occupational-relatedness of industries in the city or 'labour pool suitability'. The availability of employment by occupation is measured by the industry structure of the city ($\sum_j E_{j,l} E_{j,o}$). This measure is compared with the national employment share of the occupation in the industry. Hence, $E_{j,o} - (\sum_j E_{j,l} E_{j,o})$ defines the absolute difference between the national employment share

Table 2.5. Regression results

	Employment growth 1990-2009							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employment	-0.464***	-0.286**	-0.339**	-0.284*	-0.264*	-0.257*	-0.297**	-0.332**
	[0.147]	[0.143]	[0.146]	[0.149]	[0.144]	[0.143]	[0.145]	[0.150]
Connectivity	0.425***	0.375**	0.318**	0.387**	0.344**	0.384**	0.400**	0.440**
	[0.143]	[0.158]	[0.146]	[0.193]	[0.156]	[0.156]	[0.172]	[0.213]
Industrial specialisation			-0.194**					-0.219***
			[0.076]					[0.076]
Labour suitability				-0.014				-0.037
				[0.148]				[0.155]
Social skills					0.046			-0.017
					[0.065]			[0.070]
Routine tasks						-0.077		-0.121
						[0.069]		[0.079]
Computer use							-0.042	-0.116
							[0.082]	[0.088]
High skilled		0.179*	0.162	0.177*	0.184*	0.163	0.200*	0.185*
		[0.103]	[0.099]	[0.104]	[0.101]	[0.103]	[0.107]	[0.102]
Rent		-0.450***	-0.465***	-0.451***	-0.449***	-0.464***	-0.452***	-0.500***
		[0.087]	[0.086]	[0.088]	[0.087]	[0.090]	[0.087]	[0.089]
January temperature		-0.117	-0.191	-0.117	-0.121	-0.108	-0.119	-0.190
		[0.146]	[0.145]	[0.147]	[0.147]	[0.147]	[0.147]	[0.147]
July temperature		0.391***	0.440***	0.391***	0.389***	0.369***	0.392***	0.413***
		[0.128]	[0.131]	[0.129]	[0.127]	[0.127]	[0.129]	[0.130]
North-east		-0.295	-0.293	-0.296	-0.270	-0.284	-0.290	-0.278
		[0.230]	[0.239]	[0.231]	[0.237]	[0.233]	[0.229]	[0.241]
Midwest		-0.535**	-0.546***	-0.537**	-0.533**	-0.565***	-0.532**	-0.592***
		[0.208]	[0.207]	[0.208]	[0.208]	[0.208]	[0.207]	[0.206]
West		1.439***	1.518***	1.441***	1.433***	1.402***	1.441***	1.481***
		[0.252]	[0.249]	[0.253]	[0.247]	[0.253]	[0.252]	[0.249]
Constant	-0.000	-0.193	-0.209	-0.192	-0.174	-0.179	-0.194	-0.201
	[0.076]	[0.130]	[0.127]	[0.130]	[0.129]	[0.130]	[0.130]	[0.126]
Observations	168	168	168	168	168	168	168	168
Adjusted R-squared	0.039	0.432	0.451	0.428	0.430	0.433	0.429	0.449

Note: variables defined as in Table A.2 in Appendix A.3, Table 2.2 displays summary statistics of these variables. All variables are standardised with a mean of zero and a standard deviation of one. There are three regional dummies, region 'South' is the reference group. Robust standard errors are in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

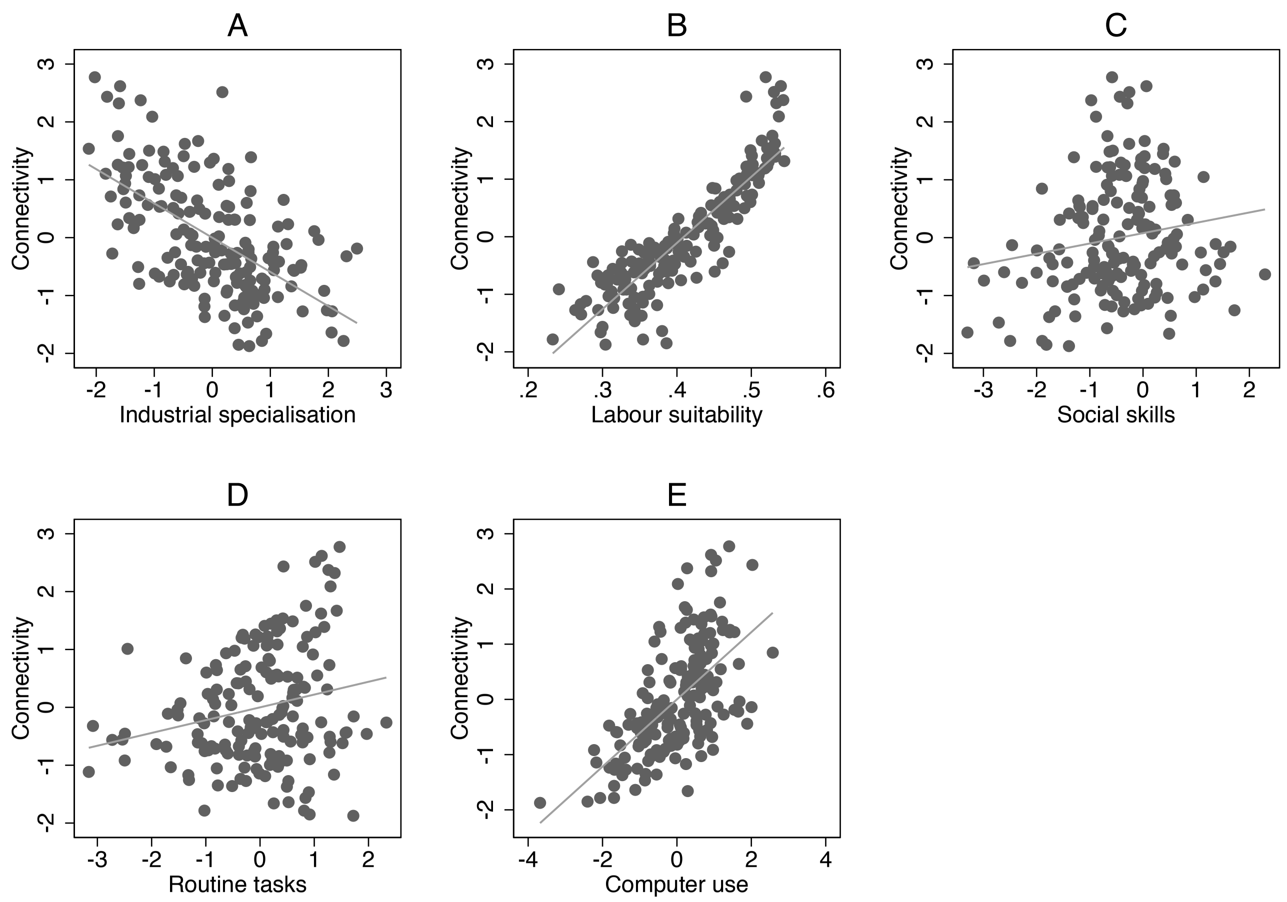
of an occupation in an industry and the local availability of employment given the industrial structure. Aggregated at the city-industry level this measure shows the suitability of the overall city employment for that certain industry. This is calculated for all industries and weighted by the importance of the industry in city employment ($\sum E_j$).

Panels A and B in Figure 2.6 show that the connectivity of the employment in the city correlates strongly with the industrial specialisation level (-0.59, significant at the 1 percent level) and with the labour pool stability measure of labour suitability (0.89, significant at the 1 percent level). It could be the case that our measure picks up the impact of spatial variation in industrial structure on employment growth. The correlation between the indicators for industrial structure do not correlate with employment growth (see Panels A and B in Figure 2.7). Column (3) adds the city's industrial specialisation level to our baseline regression, while column (4) includes the labour pool suitability of the industrial structure. The coefficient for industrial specialisation is negative and statistically significant while the labour suitability does not have a significant impact. Both indexes do not affect the significance or size of the connectivity coefficient. The decrease of the adjusted R-square indicates that these indexes do not add explanatory value concerning employment growth. When we exclude task connectivity from the regressions the coefficient of labour suitability becomes significant, while the RSI coefficient remains significant in explaining employment growth in this period.

Column (5) in Table 2.5 adds the importance of social skills. Bacolod et al. (2009) show that the presence of social skills positively influences employment. In terms of our analysis this could imply that our measure of connectivity indirectly measures social skills. Indeed, there is a positive and significant correlation between the relative importance of social skills and the connectivity of the performed tasks (see Panel C in Figure 2.6, 0.17 (0.03)). Panel C in Figure 2.7 shows a positive correlation (0.20 (0.01)) between employment growth and social skills. When we control for size and local characteristics, the coefficient of social skills becomes insignificant. This suggests that task connectivity is not picking up the effect of social skills on employment growth.

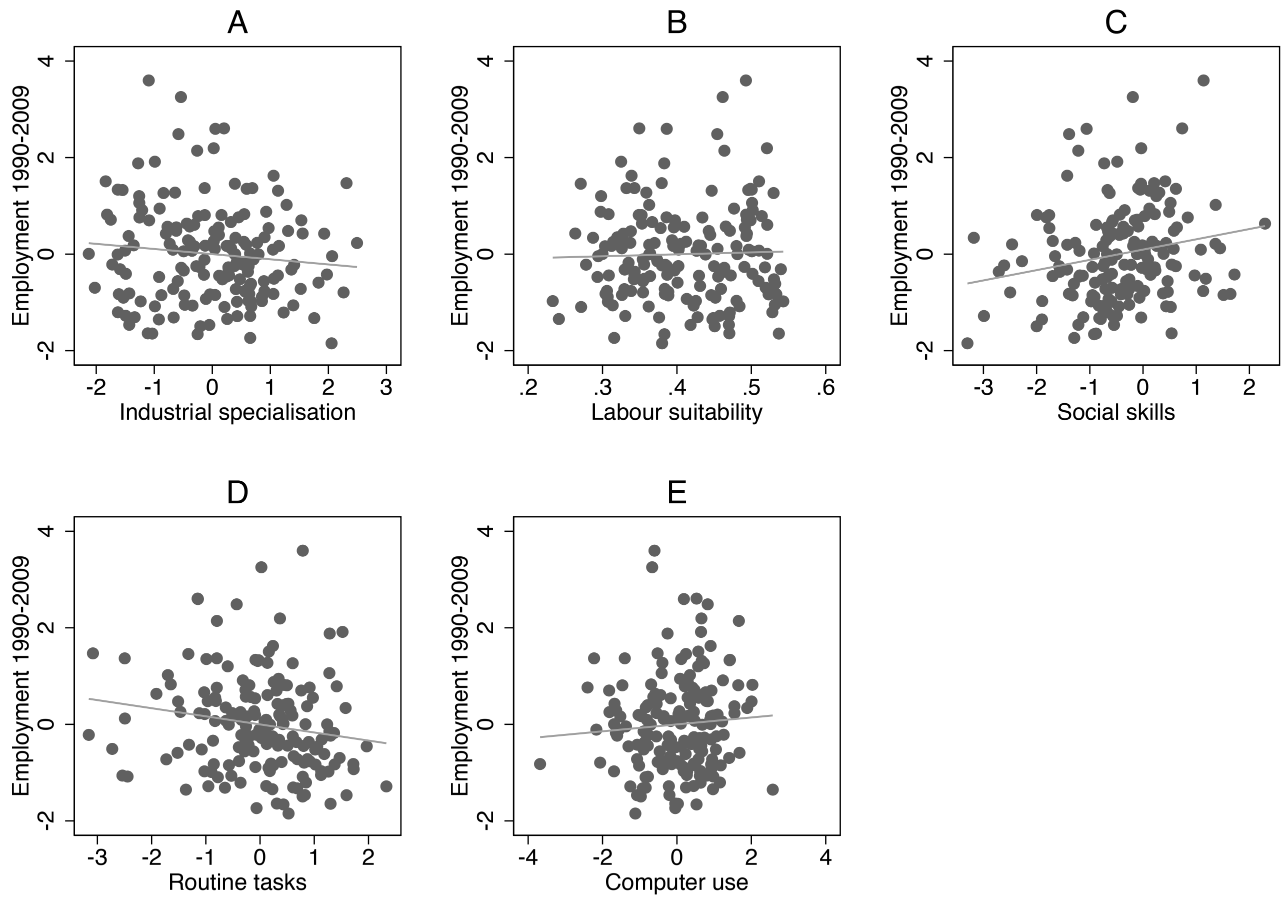
Finally, we address the importance of routine and non-routine job tasks and the use of computers. Tasks that are connected seem to require more interactions. Communication technologies make these interactions easier and less costly (e.g., Gaspar & Glaeser (1998), Blum & Goldfarb (2006)). Autor et al. (2003) have carefully introduced the notion of routine and non-routine job tasks. Their analysis

Figure 2.6. The relationship between connectivity and other factors in 1990 [a]



[a] Dots represent the 168 cities of the sample. The correlations are displayed in the correlation matrix (Table A.3 in Appendix A.3). Table A.2 displays measures and sources of all variables. All variables are standardised with a mean of zero and a standard deviation of one. The group of outlier cities in Panel B with a relatively higher connectivity given their labour suitability includes the largest metropolitan areas of the US (New York, Chicago, etc.).

Figure 2.7. The relationship between employment growth and other factors 1990 [a]



[a] Dots represent the 168 cities of the sample. The correlations are displayed in the correlation matrix (Table A.3). Table A.2 in Appendix A.3 displays measures and sources of all variables. All variables are standardised with a mean of zero and a standard deviation of one.

focuses on changes in the importance of job tasks to explain changes in wages and employment in the United States. The definitions of routine and non-routine tasks used in the analysis are based on the complementarity and substitutability of job tasks and computer technology. Routine tasks are substituted and likely to lose in terms of employment and wages, while non-routine tasks are complemented by computers. The latter set of tasks gains in terms of labour-market prospects. Autor & Dorn (forthcoming) add a spatial dimension and show that cities with employment specialisation in routine-intensive occupations in the 1960s experience employment and wage polarization after 1980. A possible concern with our results could be that non-routine tasks and tasks that require more computer use are more connected relative to routine tasks. We define the importance of routiness and the importance of computer use in cities. The routiness variable is defined as the ratio of the importance of routine tasks relative to the importance of non-routine tasks in city employment. Routine and non-routine tasks are defined as in Autor et al. (2003). Task importance by occupation from the DOT is matched to the CPS data in the same way the task data from ONET is matched. The importance of routine and non-routine tasks in US cities is defined as their average importance measured via occupation distributions. Computer use is defined as the ONET task 'using computers and computer systems (including hardware and software) to program, write software, set up functions, enter data, or process information.' This way of using computers does not reflect all types of uses, but forms a relatively good approximation for the analysis of clustering tasks together or placing some of them at distance (for a discussion of computer measures in analyses such as ours, see Katz (2000)). Indeed, Figure 2.6 shows that the share of connected tasks correlates with both the share of routine tasks (0.31 (0.00)) and the importance of computer use (0.61 (0.00)). Columns (6) and (7) of Table 2.5 present the results of a regression model in which we explain changes in employment between 1990 and 2009 by the importance of respectively routiness and computer use. The insignificant coefficients of both indicators suggest that this measure of routineness does not add explanatory power to our estimates. The effect of task connectivity remains significant.

Lastly, column (8) includes all covariates in one regression. Both the significance and the point estimate of the connectivity coefficient remain similar, the size of the point estimate even increases a bit.

2.5 Alternative measures of task composition

The estimates documented in Table 2.5 suggest that task connectivity is correlated with employment growth across our sample of cities. We now analyse whether the connectivity between tasks is the appropriate measure for the task-composition of cities. Section 2.5.1 defines two alternative measures of task connectivity. Section 2.5.2 shows estimates with the employment shares of the task groups and defines three other indicators that could capture task connectivity: the labour suitability of tasks and the specialisation and diversity level of the task structure.

2.5.1 Measures of task connectivity

Table 2.6 presents the results. Column (1) displays the baseline results with our measure of task connectivity, which is copied from Table 2.5, column (2). Second, we construct the spatial connectivity between required job skills. ONET defines skills as 'Developed capacities that facilitate learning or the more rapid acquisition of knowledge'. Examples are speaking, writing, programming and repairing. 46 separate skills are distinguished. We measure the connectivity between these 46 skills in the same way as our task connectivity measure. The connectivity between skills refers to the importance of human capital in cities (Glaeser & Ressenger, 2010). Column (2) presents the results of an analysis with this indicator of skill connectivity instead of our common indicator. The coefficient of connectivity between worker skills is insignificant. The connectivity between worker skills does not explain employment growth of cities. If we include both the connectivity between tasks and the connectivity between skills the coefficient of task connectivity is not affected. This suggests that worker tasks seem to capture the concept of task connectivity better than required skills.

Ellison & Glaeser (1997) and Ellison et al. (2010) use an indicator to define the co-agglomeration of industries. Here, we apply their indicator at the task level. The co-agglomeration index for city l is defined as:

$$CA_l = \sum_{t=1}^{t=41} \tilde{E}_{t,l} \left(\frac{\sum_{l=1}^{l=168} (\tilde{E}_{t,l} - \bar{E}_l)(\tilde{E}_{t',l} - \bar{E}_l)}{1 - \sum_{l=1}^{l=168} \bar{E}_l^2} \right). \tag{2.10}$$

$\tilde{E}_{t,l}$ refers to the estimated employment share of task t in city l. $\bar{E}_l$ refers to the average employment share of tasks in city l. The fraction on the right-hand-side calculates the co-agglomeration of task t. The numerator in the fraction calculates the over-representation of task t in city l relative to the over-representation of task

t'. The denominator controls for city size. The left part of the right-hand-side generates the average co-agglomeration of the city by multiplying task employment by task co-agglomeration.

In contrast with our connectivity measure, the co-agglomeration index includes information about the diversity of the city's employment. Task connectivity and co-agglomeration strongly correlate (0.63 (0.00)). However, when co-agglomeration is included in the analysis instead of task connectivity the task composition has no significant impact on employment growth (see column (3)). Including both measures does not change the results. The co-agglomeration index is originally used to measure the co-agglomeration of industries. The insignificant coefficient of this index suggests that spatial concentration is less important at the task level.

2.5.2 Measures of task composition

We next consider the effect of the four task groups separately to investigate whether employment growth is driven by one particular set of tasks. First, we define the city's task composition by the employment share of the four task groups. Columns (4) to (7) of Table 2.6 present the estimates in which the employment shares of the four tasks groups are included instead of the city's task connectivity. The city's employment share of information input obtains a negative coefficient (significant at the 10 percent level). A one standard deviation larger employment share of one of these task groups results in about 14 percent of a standard deviation lower employment growth. The coefficients of the share of work output and mental processes tasks are insignificant (column (6) and (7)). Lastly, the employment share of interacting with others has a positive impact on employment growth. The coefficient is smaller than the one of task connectivity and is significant at the 10 percent level only. Table A.4 in Appendix A.3 shows the estimates of regressions in which cross-terms between task groups are included. None of the cross-terms between task groups is statistically significant.

Next, we define the task structure of the city by constructing the relative specialisation index, the Hirschman-Herfindahl index and the Glaeser-Kerr index at the task level. Duranton & Puga (2004) indicate three microfoundations for the efficiency mechanism of cities; increasing the possibilities to share, match and learn. Spatial concentration of industries enhances possibilities to share facilities and suppliers, match employees to employers and learn from similar workers and firms. Empirical evidence of these mechanism is substantial (for an overview of the literature, see Glaeser & Gottlieb (2009)). Here, we test whether these mechanisms also

Table 2.6. Measures of task composition

	Employment growth 1990-2009									
	Connectivity	Skills	Co-agglomeration	Information input	Work output	Mental processes	Interacting with others	Task specialisation	HHI	Labour suitability
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Employment	-0.286**	0.011	-0.212	0.014	0.014	0.029	0.016	-0.004	0.076	0.025
	[0.143]	[0.067]	[0.157]	[0.067]	[0.067]	[0.070]	[0.066]	[0.076]	[0.090]	[0.065]
Measure of connectivity	0.375**	0.095	0.256	-0.136*	-0.075	-0.045	0.131*	-0.064	-0.083	0.014
	[0.158]	[0.082]	[0.160]	[0.071]	[0.082]	[0.079]	[0.076]	[0.071]	[0.061]	[0.063]
High skilled	0.179*	0.207*	0.263***	0.187*	0.221**	0.301***	0.201**	0.238***	0.286***	0.271***
	[0.103]	[0.109]	[0.085]	[0.099]	[0.108]	[0.098]	[0.099]	[0.087]	[0.082]	[0.084]
Rent	-0.450***	-0.481***	-0.469***	-0.465***	-0.485***	-0.488***	-0.473***	-0.470***	-0.525***	-0.491***
	[0.087]	[0.089]	[0.087]	[0.088]	[0.089]	[0.088]	[0.087]	[0.087]	[0.092]	[0.090]
January temperature	-0.117	-0.095	-0.095	-0.110	-0.085	-0.095	-0.101	-0.094	-0.087	-0.085
	[0.146]	[0.149]	[0.146]	[0.149]	[0.147]	[0.149]	[0.147]	[0.147]	[0.150]	[0.151]
July temperature	0.391***	0.361***	0.391***	0.364***	0.352***	0.363***	0.350***	0.357***	0.373***	0.362***
	[0.128]	[0.125]	[0.131]	[0.125]	[0.123]	[0.126]	[0.122]	[0.125]	[0.125]	[0.125]
North-east	-0.295	-0.242	-0.261	-0.243	-0.241	-0.262	-0.226	-0.273	-0.215	-0.248
	[0.230]	[0.222]	[0.220]	[0.227]	[0.221]	[0.214]	[0.225]	[0.220]	[0.207]	[0.227]
Midwest	-0.535**	-0.512**	-0.492**	-0.531**	-0.512**	-0.521**	-0.521**	-0.533***	-0.488**	-0.502**
	[0.208]	[0.203]	[0.200]	[0.206]	[0.202]	[0.200]	[0.202]	[0.201]	[0.194]	[0.220]
West	1.439***	1.401***	1.443***	1.377***	1.398***	1.410***	1.369***	1.392***	1.475***	1.423***
	[0.252]	[0.246]	[0.255]	[0.248]	[0.245]	[0.249]	[0.236]	[0.252]	[0.256]	[0.245]
Constant	-0.193	-0.199	-0.210	-0.187	-0.199	-0.197	-0.190	-0.187	-0.228*	-0.207
	[0.130]	[0.128]	[0.128]	[0.130]	[0.128]	[0.126]	[0.127]	[0.128]	[0.122]	[0.131]
Observations	168	168	168	168	168	168	168	168	168	168
Adjusted R-squared	0.432	0.409	0.415	0.416	0.407	0.405	0.416	0.406	0.408	0.404

Note: variables defined as in Table A.2 in Appendix A.3, Table 2.2 displays summary statistics of these variables. All variables are standardised with a mean of zero and a standard deviation of one. There are three regional dummies, region 'South' is the reference group. In column (2) the connectivity between ONET skills is used instead of the connectivity between ONET work activities. Co-agglomeration refers to the index defined in equation (2.10). The share of the task group in columns (4) to (7) represents the city's standardised employment share in the respective task group in 1990. Task-RSI defines the relative task specialisation and is measured as in equation (2.7). HHI refers to the inverse Hirschman-Herfindahl-Index and is defined in equation (2.11). Labour suitability is measured by the Glaeser-Kerr index using tasks instead of occupations (see equation (2.9)). Robust standard errors are in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

exist at the task level using indirect measures for the benefits of sharing, matching and learning.

First, the spatial concentration of tasks could ease the possibilities to share facilities and suppliers for these tasks. Column (8) in Table 2.6 presents the results of an analysis including the regional specialisation index at the task level. The index measures the over-representation of a task within the city relative to the importance of the task in national employment. The coefficient is insignificant. The spatial concentration of our 41 tasks does not seem to explain employment growth.

As Jacobs (1969) suggested, learning might be especially beneficial under the cross-fertilisation with workers with different task packages. The idea is that the combination of workers with different experiences and skills results into radical new ideas. To apply this idea at the task level, we also consider the impact of a diverse task composition. The inverse Hirschman-Herfindahl index measures the diversity of tasks in the city employment:

$$HHI_l = \frac{1}{\sum_t \tilde{E}^2_{t,l}} \quad , \tag{2.11}$$

where $\tilde{E}_{t,l}$ represents the estimated employment share of task t in city l. The lower the index, the more dominant a certain task is in city employment. A high value indicates a diverse composition of employment in tasks. The inverse Hirschman-Herfindahl index is included in the analysis in column (9). The coefficient shows an insignificant effect of the index.

Lastly, the matching possibilities of workers with similar task packages is measured using the labour suitability measure of Glaeser and Kerr. Instead of measuring the occupational suitability of industries, the index (defined in equation (2.9)) now measures the task suitability of occupations. Hence, the index values the quality of the task packages of workers given the occupational structure of the city. The index for the suitability of the labour pool for tasks is included in column (10). The coefficient is insignificant.

The three alternative indicators for task connectivity do not seem to explain employment growth. Including the measures together with our measure of task connectivity does not change the results: the coefficient of task connectivity remains positive and significant. We conclude that the spatial connectivity between tasks correlates more strongly with city growth than the level of specialisation, diversity and labour suitability of tasks.

2.6 Alternative samples of occupations, workers and cities

We continue by testing whether our findings are robust across different samples of occupations, workers and cities. First, our estimates result from spatial variation in employment shares; they are not based on variation in the importance of tasks within occupations. We test the impact of this static measure of task importance and construct a sample which only considers the most important tasks within occupations. The analysis focuses on the main tasks within occupations, assuming that the main job tasks do not vary across space. Another possible concern is that the division of labour has changed because of the introduction of ICT. This technology has created new communication possibilities, which could have changed task connectivity. Section 2.6.2 presents estimates of our connectivity measure using two separate samples of computer intensive and computer extensive occupations. Third, in Section 2.6.3 we address the issue of the possible differences in tasks performance between cities that are relatively manufacturing and services intensive. Fourth, in Section 2.6.4 we deal with the question whether our results are driven by the importance of interactions between high-skilled workers. Finally, we deal with possible biases in our results caused by a few successful metropolitan areas such as New York City and Los Angeles. These cities belong to the largest, most connected and fastest growing cities in our sample. In Section 2.6.5 we present estimates in which we exclude these cities from the sample. Table 2.7 shows the regression results of these tests.

2.6.1 Spatial variation within occupations

Our analysis exploits spatial variation in occupational composition to measure variation in task input. The reason is that we only observe national task inputs. This approach suffers from the problem that it assumes that tasks carried out within occupations are static. Baumgardner (1988a) and Duranton & Jayet (2011) suggest that this is unlikely to be true. A car mechanic in New York might carry out a different task package than a car mechanic in Detroit. Bacolod et al. (2009) also point at this caveat in their analysis.

To deal with this issue, we conduct an additional analysis using only the 'core' tasks of an occupation. Task connectivity is calculated across the most important tasks. The assumption is that the task composition of occupations varies over space but that the 'core' tasks do not vary. For example, the task packages of a car mechanic vary between cities but the task 'repairing' will be an important task in all car mechanic jobs. The distribution of tasks across US cities is now defined by the tasks within occupations with an importance above the mean of all 41 tasks in that same occupation. Column (1) of Table 2.7 shows the results of a regression analysis with task connectivity defined for the most important tasks only (instead of all 41 tasks). The coefficient of task connectivity in explaining changes in employment growth drops, but the coefficient remains significant at the 10 percent level.

2.6.2 Computer intensity

Job tasks that need to be performed in close vicinity are likely to require more face-to-face interactions. These interactions are affected by computers. The use of computers either complements or substitutes face-to-face interactions (Ioannides et al., 2008). Acemoglu & Autor (2011) indicate a crucial distinction between the employment development of computer intensive and computer extensive occupations. In Section 2.4.3 we have shown that the importance of computer use and routine tasks is unlikely to explain the impact of task connectivity on employment growth. Here, we extend this analysis and focus on the role of computer intensive occupations. Column (2) shows estimates for the correlation between the connectivity of a city's computer intensive occupations and employment growth. For computer intensive occupations the importance of computer use is at least one standard deviation larger than the average importance. The task connectivity between tasks of computer intensive occupations has a positive and significant impact on employment growth. Column (3) presents the estimates for all other occupations. The coefficient is positive and insignificant. Especially the connectivity between tasks performed in computer intensive occupations seems to relate to employment growth. This is in line with the literature on the employment effects of computerisation (see Acemoglu & Autor (2011)).

2.6.3 Idea-producing versus product-producing cities

The changing economy and especially the de-industrialisation of the US economy was beneficial to cities, such as New York, but detrimental to others, such as Detroit. Glaeser & Ponzetto (2010) show that improvements in transport and communication technologies increased the returns to ideas. Idea-producing cities, such as New York and Boston, are favoured by this trend while product-producing places, such as Detroit, are hurt. Here, we test whether task connectivity is beneficial for

Table 2.7. Additional samples

	Employment growth 1990-2009								
	Most important tasks (1)	Computer intensive occupations (2)	Not computer intensive occupations (3)	Manufacturing sectors (4)	Services sectors (5)	High-skilled workers (6)	Medium-skilled workers (7)	Low-skilled workers (8)	Without largest cities (9)
Employment	-0.129	-0.259*	-0.206	-0.262*	-0.302**	-0.390***	-0.245*	-0.277*	-0.284*
	[0.138]	[0.147]	[0.150]	[0.139]	[0.143]	[0.149]	[0.135]	[0.141]	[0.144]
Connectivity	0.235*	0.334**	0.251	0.334**	0.389**	0.503***	0.322**	0.356**	0.372**
	[0.140]	[0.157]	[0.158]	[0.149]	[0.160]	[0.166]	[0.147]	[0.158]	[0.158]
High skilled	0.247*	0.221**	0.272***	0.224**	0.213**	0.157	0.220**	0.226**	0.170
	[0.138]	[0.095]	[0.081]	[0.092]	[0.092]	[0.098]	[0.093]	[0.091]	[0.104]
Rent	-0.468***	-0.463***	-0.475***	-0.462***	-0.469***	-0.467***	-0.475***	-0.469***	-0.488***
	[0.116]	[0.085]	[0.084]	[0.083]	[0.079]	[0.077]	[0.080]	[0.081]	[0.079]
January temperature	-0.125	-0.096	-0.081	-0.100	-0.085	-0.057	-0.082	-0.090	-0.089
	[0.200]	[0.143]	[0.143]	[0.143]	[0.139]	[0.134]	[0.141]	[0.140]	[0.145]
July temperature	0.187	0.399***	0.377***	0.396***	0.389***	0.389***	0.382***	0.385***	0.417***
	[0.148]	[0.128]	[0.125]	[0.125]	[0.124]	[0.124]	[0.125]	[0.124]	[0.130]
North-east	-0.817**	-0.263	-0.256	-0.263	-0.263	-0.255	-0.264	-0.276	-0.210
	[0.369]	[0.220]	[0.219]	[0.224]	[0.223]	[0.217]	[0.221]	[0.226]	[0.229]
Midwest	-0.741**	-0.520**	-0.496**	-0.512**	-0.486**	-0.451**	-0.494**	-0.509**	-0.436**
	[0.339]	[0.205]	[0.200]	[0.200]	[0.197]	[0.195]	[0.198]	[0.199]	[0.205]
West	0.887***	1.440***	1.397***	1.431***	1.422***	1.454***	1.423***	1.425***	1.579***
	[0.319]	[0.251]	[0.239]	[0.243]	[0.237]	[0.243]	[0.242]	[0.238]	[0.250]
Constant	0.141	-0.202	-0.200	-0.201	-0.206*	-0.224*	-0.205*	-0.199	-0.269**
	[0.274]	[0.126]	[0.127]	[0.126]	[0.124]	[0.123]	[0.124]	[0.126]	[0.126]
Observations	86	168	168	168	168	168	168	168	163
Adjusted R-squared	0.447	0.429	0.417	0.428	0.438	0.446	0.426	0.433	0.447

Note: column (1) displays the baseline regression of Table 2.5. In column (2) the measure of connectivity only includes tasks which obtain a national employment share of more than 0.02 which results in a sample of about 75% of all tasks. Occupations for which the importance of computer use is more than one standard deviation above average are defined as 'computer intensive' (column (3)). High-skilled workers obtained at least a bachelor degree while low-skilled workers obtained at most a high-school degree, medium-skilled workers continued studying after high-school but did not obtain a bachelor degree. Lastly, column (9) excludes cities which are more than two standard deviations larger than the mean: Detroit, Philadelphia, Washington, Chicago, New York and Los Angeles. Table A.2 in Appendix A.3 presents the measurement and sources of the variables.

idea-producing cities, product-producing places or both.

Column (4) of Table 2.7 shows estimates for a sample of manufacturing sectors only. The correlation between task connectivity and employment growth is somewhat smaller for these sectors but still substantial and statically significant. Next, column (5) presents the estimates for a sample of only service sectors. For service sectors, the impact of task connectivity is stronger than for manufacturing sectors. Hence, changes in the employment of both product-producing and idea-producing cities seem to be partly explained by our measure of task connectivity.

2.6.4 Worker skills

We continue by addressing the importance of the complementary between skills and cities. High-skilled workers tend to sort into larger cities and this sorting explains spatial wage and employment differences (Combes et al., 2008; Glaeser et al., 2012). The relation between skills and cities seems to be complementary (Glaeser & Ressenger, 2010; Elvery, 2010). Urban density particularly stimulates human capital spillovers (see Rosenthal & Strange (2008)) and human capital accumulates more quickly in urban areas (see Glaeser & Maré (2001)). Large cities are however characterised by relatively fat tails and their inhabitants are more likely to be high and low-skilled workers, while medium-skilled workers seem to sort into smaller cities (Eeckhout et al., 2010). New York and Detroit house both the best workers of the country, with degrees from the best universities, and the lowest-skilled of the nation. A possible concern with our results is that they might be driven by the strong connectivity between the tasks of high-skilled workers.

We analyse whether our findings hold for several groups of workers. Column (6) in Table 2.7 shows the estimates for a sample of high-skilled workers who obtained at least a bachelor degree. Second, columns (7) and (8) show the estimates for samples of medium- and low-skilled workers. In all three samples the coefficient for task connectivity is positive and significant. As expected, task connectivity of high-skilled workers has a stronger impact on employment growth than task connectivity of low-skilled workers. An increase of one standard deviation in connectivity rises the employment growth by about 50 percent of a standard deviation in the sample of high-skilled workers and by 36 percent of a standard deviation in the sample of low-skilled workers. In line with the work of Eeckhout et al. (2010), the connectivity between tasks of medium-skilled workers is only moderately correlated with employment growth.

2.6.5 Without the main metropolitan cities

Finally, we test whether some large metropolitan cities dominate our results. The largest cities in our sample of 168 cities are the cities with the highest shares of high-skilled people, the strongest connectivity between the performed tasks and the highest growth given size. Column (8) excludes cities with more than two standard deviations employment above the mean. These are Detroit, Philadelphia, Washington, Chicago, New York and Los Angeles. The coefficient of task connectivity hardly decreases and remains statistically significant. The adjusted R-square increases a bit, which seems to be caused by a stronger impact of location characteristics, such as rents and July temperature in this sample.

2.7 Conclusion

This chapter is concerned with measuring and interpreting changes in employment across 168 US cities in the period 1990-2009. Within this period (characterised by rapid technological change) not only the division of labour between and within occupation was affected, but also the division across space. Our analysis provides a task-based approach, which allows us to investigate the underlying relations between technology and employment shifts.

Our framework relies upon the idea that employment grows when job tasks need to be performed in close vicinity and human interactions are important. The importance of vicinity and human interactions for tasks can lead to clustering of tasks or spreading to other places, which we measure by task connectivity. The extent to which tasks are spatially connected indicates whether they require face-to-face contacts or whether they could be done at distance at reasonable costs. To analyse employment effects of changes in the division of tasks, we develop an empirical measure of task connectivity based on the correlation between several tasks in cities.

Our estimates suggest that differences in task connectivity contribute to explaining changes in employment structure across US cities. In particular we show that changes in employment across US cities can for some part be explained by our measure of task connectivity. Higher task connectivity at the city level implies less room for placing tasks at distance. When tasks are glued to the location (and to other tasks) cities are more likely to grow relative to cities with lower levels of task connectivity. We find that a one standard deviation increase in task connectivity increases employment by 30 to 45 percent of a standard deviation. The coefficient of task connectivity is not affected by the inclusion of several other city characteristics. Furthermore, the spatial connectivity between tasks seems to be more effective than the spatial concentration of certain tasks and the labour suitability of the task composition. We investigate the robustness of our work by considering the effects of computerisation of work, the de-industrialisation trend, the sorting of workers and the impact of outstanding successful cities. We also investigate the limitations of our cross-section of task data and find that our results seem to be robust.

This chapter adds to the literature in labour economics and urban economics by offering a measure to explain employment changes across space. This complements the literature in labour economics focusing on changes in the task composition of work (for a review, see Acemoglu & Autor (2011)) and to the literature in urban economics explaining changes in employment in cities (for a review, see Glaeser & Gottlieb (2009)). Future work should consider deepening of the exact anatomy of task connectivity for explaining the success and decline of cities.

CHAPTER 3

TOWN AND CITY JOBS:

YOUR JOB IS DIFFERENT IN ANOTHER LOCATION

3.1 Introduction

A doctor in a small rural town is responsible for all kinds of treatments. Whether you have a heart attack or giving birth, he is the person to go to. In large cities there are thousands of doctors, with hundreds of different specialities. If you have a heart attack you definitely go to another doctor then when you are giving birth. Big cities provide more career opportunities than small towns. In the big city you have more chances to become a 'true' expert, work on more complex cases and learn from your peers. These examples stress the complexity of job contents and the variation by the extent of the market. Both the demand for a certain activity and the supply of skills vary with the extent of the market. Life is different in large cities, workers are different, local industries are different, but to what extent does the content of jobs vary across city size?

Back in 1988 James Baumgardner (1988a) modelled the idea of Adam Smith that the division of labour is bound by the extent of the market. Cooperation in a larger local market results in a more efficient division of labour. Workers segregate into subsets of different activities. In a town with two doctors, the doctors can divide the medical activities and specialise in only half the activities. Duranton & Jayet (2011) translate the model of Baumgardner such that scarce occupations are more likely to be performed in larger cities which they back-up with empirical evidence for France. On the level of job activities, the empirical literature tends to focus on particular industries and case-studies (Baumgardner, 1988b; Garicano & Hubbard, 2009).

The sorting of workers themselves, the ambitious doctor who would rather work in the capital than in a rural town, is a central issue in urban economics (Glaeser & Maré, 2001; Eeckhout et al., 2010; Combes et al., 2008; Venables, 2011). Given this central issue, there is remarkably little empirical work on the skill requirement for jobs across space. Most research uses education, occupation and industry information or just worker fixed effects to analyse the mechanism behind the productivity in cities. Only modest attention is paid to the fact that jobs might differ across cities and the fact that a more efficient division of labour across jobs or different skill requirements might affect the mechanism. Ignored variation within occupation and industries between cities hampers adequate analyses on the mechanism behind agglomeration economies.

In this chapter we take a step towards unravelling the efficiency of cities by analysing the variation in job content across cities. Most datasets hinder such an attempt as they lack spatial variation in job content. We exploit the German survey of the working population, which includes job activities for individuals across German cities. Our main result is that the specialisation level of jobs and the required level of cognitive skills increase with city size.

To conceptually guide our empirical analyses, we first set out a theoretical background. The basic setting of our framework relies on the model of Baumgardner (1988a). The production of a good consists of the performance of a continuum of tasks. The more time a worker devotes to the performance of a task, the more specific skills he develops for this task. Workers are more productive when they focus on a smaller subset of tasks and there are increasing returns to worker input. Local workers cooperate which results in a more efficient division of labour in larger markets. Hence, workers in large cities are more specialised than workers in small cities and develop more specific skills for their job tasks.

Second, we test the predictions of our theoretical set up using the German survey of the working population on qualification and working conditions (the BIBB survey). In contrast to most information on job tasks, the dataset includes individual task data next to a very broad set of other personal and work characteristics. For each worker in the dataset we obtain information on job tasks, occupation, industry, demographic characteristics, education, location and so forth. We construct two measures for job content. The first measure is the number of subtasks (performed 'sometimes' or 'rarely'). The number of tasks a worker performs sometimes or rarely serves as a measure for the time devoted to the core tasks of his job. The fewer tasks a worker performs sometimes (subtasks), the more time he has to focus on the main job tasks and the more specialised he is. The second measure specifies the importance of skill development in the job. Respondents indicate the importance of several cognitive skills for the performance of their job. The demanded cognitive skills reflect the importance of task specific knowledge for performing the job.

As documented by Duranton & Jayet (2011), scarce occupations are observed more often in large cities than in smaller ones. To control for this unequal spatial distribution of jobs and their task packages we include job fixed effects. We find that workers in large cities on average perform 7 percent of a standard deviation fewer subtasks than workers in small cities. The same job consists of more subtasks when it is performed in a small city (less than 20,000 inhabitants) compared to a large city (more than 100,000 inhabitants). Jobs in larger cities also demand 8 percent of a standard deviation more cognitive skills than the same jobs in small cities. The higher specialisation level of workers in large cities explains part of the higher requirement of cognitive skills. The sorting of more capable workers into large cities likely explains further spatial variation in the demand for cognitive skills. Furthermore, these sorting patterns are likely to affect the spatial variation in specialisation levels of jobs as well. We do not distinguish the causes of these spatial variations. The results are however robust over several sub-samples, for different measures, at different spatial units and to the inclusion of several co-variates.

Our model relates to theory about the division of labour and the extent of the market. This literature is largely based on the framework of Baumgardner (1988a). The specialisation of workers into certain job tasks increases with market size. Duranton & Jayet (2011) argue that larger markets allow workers to perform more efficient. Another strand in the literature (Becker & Murphy, 1992) argues that the extent of the market is irrelevant for the division of labour. They state that the costs of coordination between workers overrules the costs of transportation of tasks. In this chapter, we empirically examine whether the extent of the local market, hence the city size, is relevant for the division of labour.

Empirically, this field is left rather untouched. The empirical work tends to focus on case-studies. For example, Baumgardner (1988b) and Garicano & Hubbard (2009) study the division of labour across market sizes for doctors and lawyers. Other analyses focus solely on variation *between* jobs and not variation *within* jobs. Duranton & Jayet (2011) show that scarce occupations are more often observed in large French cities, while Bacolod et al. (2009) show that the allocation of cognitive skills only slightly varies across city sizes. Combes et al. (2012) find that much of the skill differences, measured by worker fixed effects, across French cities can be explained by differences between occupations rather than within occupations. We add to previous empirical work by analysing spatial variation of cognitive skills within and between occupations. Our dataset makes it possible to analyse the variation in job content instead of controlling for worker skills by using fixed effects.

Lastly, our work relates to the empirical work on job contents and especially the task-based approach in analysing employment pioneered by Autor et al. (2003). The spatial dimension of this strand can be found in the work of, among others, Autor & Dorn (forthcoming) and Bacolod et al. (2010). Autor & Handel (forthcoming) demonstrate that measures at the individual level offer substantial additional explanatory power relative to occupation level data from datasets such as Occupational Information Network (ONET). Earlier work with the German surveys is done by, among others, Spitz-Oener (2006), Gathmann & Schnberg (2010) and Dustmann et al. (2009).

The rest of the chapter is structured as follows. The next section sets out a simple framework to justify our empirical analyses. Section 3.3 provides insight in the database construction, the main variables and some descriptive statistics. The empirical strategy is discussed in Section 3.4. Section 3.5 presents the results on the spatial variation in job content. In Section 3.6, further sensitivity analyses are presented. Section 3.7 concludes.

3.2 Spatial variation in job content

This section sets out a framework for the division of labour across cities. The framework draws on the work of Baumgardner (1988a). Workers are more productive when they focus on fewer tasks and the division of labour is efficient. The extent of the market increases possibilities for division of labour.

3.2.1 Tasks

As in Adam Smith's pin factory, a very large number of tasks (activities) are combined to produce one good. The set of tasks to produce a good is presented by a segment T of length 1 and indexed by $t \in [0, 1]$. The model considers an economy with one product. All tasks need to be performed to produce one unit of this product. The market consists of I workers, indexed i. Each worker i is endowed with limited time E_i, which is all spent on performing tasks:

$$E_i = \int_{t_i=\delta_1}^{t_i=\delta_n} x_{i,t} dt. \tag{3.1}$$

In this time a worker performs a subset of tasks ($t_i = \delta_1 \rightarrow t_i = \delta_n$, we label this subset with δ_i). He uses time input $x_{i,t}$ for each task t. Hence, a larger subset of tasks implies that the worker has less input x_t per task. Following Baumgardner, we assume symmetry in the production technology and demand across all tasks on the segment. As a result of the symmetry, identification and relative positions on the segment do not affect the model.[1]

Worker i uses skills $S_{i,t}$ to perform a task t. The worker develops these task specific skills during the performance of the task. Hence, his task-specific skills depend on the amount of time he spends on task t:

$$S_{i,t} = cx_{i,t}, \tag{3.2}$$

where c is the general human capital each worker is endowed with from the start. $x_{i,t}$ refers to the time worker i spends on task t. The more time a worker devotes to the production of one specific task, the more specific skills for producing this task he develops, see Becker & Murphy (1992). The more a worker specialises in one task, the more efficient he becomes in producing that specific task. For instance, a doctor who only performs heart surgeries will learn more about that surgery than a doctor who also removes appendices. A heart surgery specialist will be more efficient than a general surgeon in performing a heart surgery. The task-specific worker skills determine the time it takes to produce the task:

$$x_t = \frac{a}{S_{i,t}} = \frac{a}{cx_{i,t}}, \tag{3.3}$$

where a defines the fixed amount of time for the performance of task t. The variable amount of time needed to produce the task depends on the task-specific worker skills. The time needed to perform a task is endogenous. The more time a worker devotes to a certain task, the less time producing an extra unit takes. The production of task t by worker i is determined by both the time the worker devotes to the

[1] The consequences of this assumption for our empirical strategy are discussed in Section 3.4.

performance of the task and the amount of time it takes to produce one unit:

$$q_{i,t} = \frac{x_{i,t}}{x_t} = \frac{x_{i,t}}{a/S_{i,t}} = \frac{cx_{i,t}^2}{a}. \tag{3.4}$$

This indicates that there are increasing returns to worker input for a task. As the worker has limited time endowment, there are increasing returns to worker input. The total output of worker i consists of the sum of the output of all the tasks he performs:

$$q_i = \int_{t_i=\delta_1}^{t_i=\delta_n} q_{i,t} dt. \tag{3.5}$$

A worker is most productive when he spends all his time on performing one task. However, to produce one good all tasks on the segment T should be performed. Hence, if only one worker spends time on producing the good he has to perform all the tasks. The output of the good is as follows:

$$Q = \int_0^1 q_t dt, \tag{3.6}$$

where $q_t = \int_{i=1}^{i=I} q_{i,t} di$. q_t refers to the output of task t generated by all workers in the market.[2]

3.2.2 Extent of the local labour market

All workers I in the local market cooperate in the production of the good. The division of tasks follows from the maximisation of the output Q. Substituting equations (3.2) to (3.4) into the output function (equation(3.6)) it follows that:

$$Q = \int_0^1 \int_{i=1}^{i=I} \frac{cx_{i,t}^2}{a} di, \tag{3.7}$$

subject to equation (3.1). There are increasing returns to worker task input $x_{i,t}$. Worker productivity decreases with the subset of tasks they perform:

$$\frac{\partial q_i}{\partial \delta_i} < 0. \tag{3.8}$$

[2] For simplicity the model ignores comparative advantages of workers in certain tasks.

Full specialisation into one task may however be hindered by the production function (3.6) which states that all tasks of the segment T need to be performed to produce one good. The extent of specialisation depends on the size of the market (I). The segment of tasks T with $t \in [0,1]$ is divided over all workers in the market:

$$\delta_i = \frac{T}{I}. \tag{3.9}$$

Thus, the subset of tasks of each worker (δ_i) decreases with the number of workers in the market (I). When workers cooperate, they divide the tasks, benefit from the increasing returns to individual input and become more productive. The continuum nature of the task segment induces endless specialisation benefits of increasing market size. Clearly, in reality coordination costs limit the division of tasks (Becker & Murphy, 1992). Section 3.4 discusses the consequences of coordination costs.

Figures 3.1 and 3.2 illustrate this mechanism with, for simplicity, a discrete example. In both figures 7 tasks need to be performed to produce 1 good (T = 7). Each worker is endowed with $E_i = 7$ time units. For simplicity we assume $a = 2$ and $c = 2$. In Figure 3.1 only 1 worker is available to produce the 7 tasks (I=1). Therefore $\delta_i = 7$ and $x_{i,t} = E_i/\delta_i = 1$. With his input of 7 time units he generates an output of 1 for each task ($\frac{cx_{i,t}^2}{a} = \frac{2*1}{2}$) and 7 in total. Hence, one good is produced. Next, 7 workers operate in the market in Figure 3.2. As they cooperate, they divide the tasks and benefit from the increasing returns to input: $\delta_i = 7/7$ and $E_i/\delta_i = 7$, each worker performs 1 task 7 times. The output by worker is then 49 ($\frac{2*49}{2}$). Hence, 49 goods are produced with the 7 workers. The workers in the market with 7 workers each specialise in 1 task, develop specialistic skills for this task and become more efficient in producing this task. The division of labour between the workers rises the output by worker from 7 to 49 tasks. Figure 3.3 shows the relation between number of workers and produced tasks by worker for this example.

3.2.3 Empirical predictions

In summary, the fewer tasks a worker performs, the more efficient he is in the performance of these tasks. The possibilities to divide the tasks over workers increase with the size of the local population I. Larger cities house more specialised workers and this creates more possibilities for workers to develop task-specific skills.

The model discusses an economy with one good. Cities produce many (inter-

Figure 3.1. Market with 1 worker

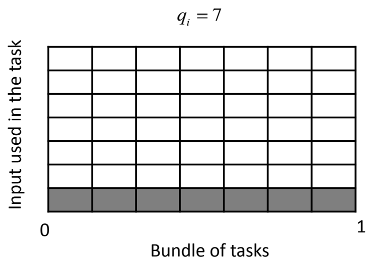
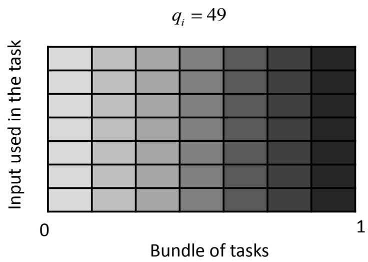
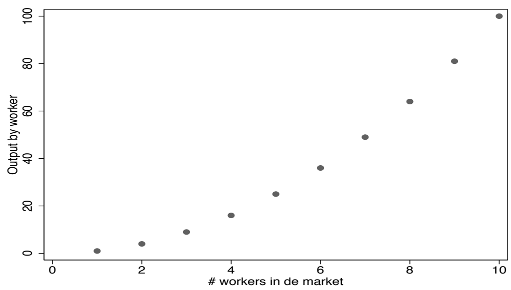


Figure 3.2. Market with 7 workers



Each block represents a time unit, each worker is endowed with 7 time units. For simplicity $a = 2$ and $c = 2$. The different shades of grey indicate the time units of the different workers. The horizontal axis divides the good into 7 tasks, the vertical axis represents the time units a worker spends on each task. q_i is measured by equation (3.5).

Figure 3.3. Extent of the market and output by worker



The figure visualises the relation between the number of workers in the market and the output by worker. For simplicity we assume $a = 2$, $c = 2$, endowment $E_i = 7$ and the good consists of 7 tasks.

mediate) goods and the distributions of industries varies across cities. The benefits from specialisation vary between goods. To control for the different specialisation benefits and task packages between goods, we estimate the spatial variation in tasks subsets (δ_i) of jobs. The model results into two empirical predictions about the spatial variation in the subset of tasks, or task packages, of jobs:

1. The jobs of workers in large cities contain smaller subsets of tasks compared to the same jobs in small cities.

2. Workers in large cities have more task-specific skills compared to workers with the same jobs in small cities.

3.3 Data, indicators and descriptive statistics

3.3.1 Data

The empirical predictions demand individual task data for workers across cities of different size. This analysis relies on the survey of the working population in Germany carried out by the German Federal Institute for Vocational Training (BIBB) and the Federal Institute for Occupational Safety and Health (BAuA).[3] The BIBB is a survey among a representative sample of Germans. The survey aims to measure qualifications, career history and detailed job characteristics of the German labour force. In contrast to most task datasets, the BIBB dataset includes individual information on tasks, occupations and locations. We employ the 2006 wave as this is the most recent wave. This wave consists of information about 20,000 Germans. Here, we focus on the definitions and construction of our main variables: tasks and local markets. For more information about the survey and the dataset we refer to the work of Rohrbach-Schmidt (n.d.).

Several questions in the BIBB relate to the content of occupations. For the empirical analyses we employ information on job tasks, job characteristics, required cognitive and specialised skills and task requirements. Examples of all these content measurements and the number of different tasks appearing in the BIBB are displayed in Table 3.1. The full list of tasks is displayed in Appendix B. For each task, the survey examines the frequency of appearance as a measure of the importance in the job. As the scaling varies between the questions, we construct three possibilities: (a) the task is a core task (appears ’always’ or ’often’), (b) the task is a subtask (appears ’sometimes’ or ’rarely’) or (c) the task is not performed by the worker. Most studies on job tasks include similar types of job tasks and measurement. Autor et al. (2003) employ the Dictionary of Occupational Titles (DOT) while most recent studies, such as Bacolod et al. (2009) and Goos et al. (2009) employ the successor of the DOT, the Occupational Information Network (ONET) dataset. Spitz-Oener (2006), Gathmann & Schnberg (2010) and Dustmann et al. (2009) (among others) employ the BIBB surveys.

The disadvantage of estimations for the whole economy is that job tasks vary between industries. As the survey includes tasks that could occur in each occupation and each industry, many (more specific tasks) are missing. Hence, the range of individual tasks in an occupation could be smaller than in real life.[4] For an extens-

[3] Hereafter we refer to this dataset as the BIBB dataset.
[4] Section 3.4 discusses whether this biases our estimates.

Table 3.1. Task definitions in the BIBB Survey

Variable	Examples	# of tasks
Job tasks	Manufacturing, organising	16
Job characteristics	Having to react to and solving unforeseeable problems	9
	Making tough choice on your own responsibility	
Cognitive skills	Manual / craft skills, technical skills	12
Specialised skills	Book-keeping, fiscal	8
Task requirements	Have to work under great deadline pressure	12
	Working very quickly	

Table 3.2. Observations seven city categories

Category	Inhabitants	Employees in Germany	
		in data	weighted
1	1–1,999	1,030	1,907,418
2	2,000–4,999	1,460	2,651,726
3	5,000–19,999	3,982	6,701,806
4	20,000–49,999	2,761	4,570,964
5	50,000–99,999	1,340	2,199,916
6	100,000–499,999	2,693	3,972,548
7	500,000–...	2,404	3,621,968

ive discussion about the disadvantages of task information we refer to the work of Autor & Handel (forthcoming), Acemoglu & Autor (2011) and Autor (2013).

The dataset contains information on the size of the city of residence. For the descriptive data we exploit the variation between the seven different categories. Table 3.2 presents the (weighted) number of observations in the dataset for the seven size categories. In the analyses we consider three city sizes: small (less than 20,000 inhabitants), medium (between 20,000 and 100,000 inhabitants) and large cities with more than 100,000 inhabitants.

3.3.2 Measuring job content

A job is defined as a three-digit occupation and two-digit industry combination. Throughout the paper the term 'job' refers to an occupation within an industry. Examples of jobs are a protective service worker within the veterinary sector and a machinery worker within the manufacture of motor vehicles, trailers and semi-trailers sector.

The theoretical model results in predictions about two forms of job contents: the number of job tasks and demand for skills. The number of tasks a worker performs defines his specialisation level and indicates the time he has to devote to his core job tasks. Within the BIBB, the questions about job tasks largely refer to demanded skills. To avoid measuring the level of demanded skills instead of the level of specialisation we focus on the number of less relevant tasks a worker performs. The more irrelevant tasks a worker has to perform, the less time he devotes to his core tasks. For instance, a scientist who also needs to organise meetings has less time to focus on his core task, namely doing research. The level of specialisation is measured by the number of tasks a worker performs sometimes or rarely (subtasks). The fewer subtasks a worker performs, the more time he has to focus on his core tasks and the more specialised he is.

Second, we define the job content by the demanded skills. Workers who focus more on their core tasks have more time and incentive to develop task-specific knowledge and skills. As explained before, we only observe broad tasks. Therefore, we measure the importance of cognitive and non-routine tasks which we assume to be applied by workers for the development of task-specific skills. We include seven cognitive skills: research, adapt to unforeseen problems, mathematical skills, technical skills, solving new problems, process optimising and do things you have not learned before. The number of cognitive skills which is crucial for the job performance proxies the development of task-specific skills. Section 3.6.1 tests the sensitivity of the results towards the choice of indicators for job content.

3.3.3 Descriptive statistics

Tables 3.3 to 3.5 present the most salient descriptive statistics for our data. On average, workers perform 15 subtasks and 18 core tasks out of the range of 58 possible tasks. First, Table 3.3 shows the variation within specialisation levels and demanded job skills across different subgroups of workers. Variables are standardised to make comparison across occupations easier. Workers above the age of 50 perform fewer subtasks than younger workers. Work experience enhances specific knowledge and with that a more specialised task package. Women have more generalist jobs than men and native speakers more than non-natives. Furthermore, the number of subtasks increases with education level. The average job requires 1.8 cognitive skills. The last two columns in Table 3.3 present the standardised values across different groups of workers. The jobs of younger workers require more cognitive skills than the ones of older workers. Logically, the demand for cognitive skills increases with education level. Females and non-native speakers indicate that their job demands fewer cognitive skills than respectively males and native speakers in-

Table 3.3. Descriptive statistics

	Number of subtasks		Demanded cognitive skills	
	Mean	SD	Mean	SD
Age groups				
Younger than 35 years	0.02	0.96	0.08	1.01
35-50 years	0.04	1.00	-0.02	1.00
Above 50 years	-0.11	1.04	-0.03	1.00
Gender				
Male	-0.11	0.99	0.13	1.00
Female	0.10	1.00	-0.12	0.99
Educational groups				
Unskilled	-0.66	1.13	-0.02	0.93
Low skilled	-0.20	1.02	0.06	1.00
Medium skilled	-0.03	1.02	-0.13	0.96
High skilled	0.13	0.92	0.21	1.03
Origin				
Native speaker	0.01	1.00	0.00	1.00
Non-native speaker	-0.14	1.02	-0.04	1.00

Note: n = 15,670. Variables are standardised with a mean of zero and a standard deviation of one. The classifications and the definitions of variables are displayed in Table B.2 in Appendix B. There are 58 subtasks defined. On average workers perform 15.61 subtasks with a standard deviation of 5.79. We distinguish seven cognitive skills, on average a worker indicates that his job demands 1.76 cognitive skills with a standard deviation of 1.09.

dicate.

Table 3.4 presents the variation in specialisation level and demanded skills across broad occupational groups. Management occupations are the most generalist occupations while elementary occupations are the most specialised ones. The variation within the group of elementary occupations is however relatively high. Professional occupations require the most cognitive skills and agricultural and fishery occupations the fewest cognitive skills. The variation across industries is smaller than the variation across occupations (Table 3.5). The wholesale trade sector is the least specialised of all sectors and the other services sector the most specialised. The administration and support sector demands the most cognitive skills and the wholesale trade the least.

Next, Figures 3.4 to 3.7 present the spatial distribution of jobs regarding their specialisation level and demanded skills. Figure 3.4 shows the kernel distribution of the performed number of subtasks in small, medium and large cities. The distribution of the number of subtasks performed shows an inverted u-shape distri-

Table 3.4. Summary statistics - occupational groups

	Number of subtasks		Required cognitive skills	
	Mean	SD	Mean	SD
1. Managers	0.36	0.97	0.11	1.01
2. Professionals	0.03	0.84	0.38	1.03
3. Technicians	0.10	0.94	0.05	1.00
4. Clerks	0.10	1.04	-0.15	0.91
5. Service workers	-0.10	1.03	-0.03	0.95
7. Craft and trade workers	0.02	1.03	-0.42	0.92
8. Operators and assemblers	-0.40	1.03	-0.34	0.85
9. Elementary	-0.71	1.71	-0.03	0.80

Note: n = 15,670. Variables are standardised with a mean of zero and a standard deviation of one. Table B.2 in Appendix B displays the definitions of the variables. Occupations are defined by one-digit ISCO 1988 codes. Skilled agricultural and fishery are dropped because their location depends on natural resources.

bution. Workers in large cities perform slightly fewer subtasks than workers in medium and small cities. The differences are however only modest. Figure 3.5 presents the same distributions for a sample of high-skilled workers. The distributions are rather similar but the spatial variation is somewhat larger. Figures 3.6 and 3.7 show the same exercise for the number of demanded cognitive skills. Most workers indicate that their job demands a maximum of two cognitive skills. The share of workers who perform more cognitive skills is larger in large cities than in medium and small cities.

Lastly, we test whether scarce occupations are more likely to be performed in large cities and replicate the analysis of Duranton & Jayet (2011) for German cities. Cities produce many (intermediate) products with various demand thresholds. Duranton & Jayet (2011) show that scarce occupations are more often found in large cities. Using a logit approach, we estimate the probability that a job (occupation–industry combination) is performed in a city. Within each industry, the occupations are classified into four categories: occupations with a very high scarcity level, a high, a low and a very low scarcity level. The scarcity level represents the national employment of that occupation within a certain industry. Appendix B describes the full estimation method. Table 3.6 presents the results. Scarce occupations appear more often in large cities than in small cities. This relation increases both with the scarcity level of the occupation and the size of the city. Thus, to measure the spatial division of specialisation, we should control for the division of jobs across cities.

Table 3.5. Summary statistics - industrial groups

	Number of subtasks		Cognitive skills	
	Mean	SD	Mean	SD
3. Manufacturing	0.00	1.03	-0.17	0.97
4. Electricity and gas	0.06	1.05	-0.19	0.88
5. Water supply	-0.05	0.98	-0.30	0.90
6. Construction	0.16	1.06	-0.28	0.95
7. Wholesale trade	0.21	1.05	-0.47	0.91
8. Transport	0.03	1.08	-0.16	0.94
9. Accommodation and food	0.06	1.2	-0.09	0.09
10. Information and communication	-0.15	1.05	-0.16	0.89
11. Financial	0.15	0.97	-0.01	0.99
13. Professional, scientific and technical activities	-0.04	0.98	0.11	0.97
15. Administration and support	-0.04	0.85	0.55	1.03
16. Education	-0.04	0.92	0.18	1.00
18. Arts, entertainment	0.05	1.01	0.18	1.03
19. Other services	-0.37	1.2	-0.02	0.99
20. Household	-0.27	1.19	-0.03	1.00
21. International organisations	0.18	0.96	-0.45	1.07

Note: n = 15,670. Variables are standardised with a mean of zero and a standard deviation of one. Table B.2 in Appendix B displays the definitions of the variables. Industries are defined by one-digit NACE codes. Agriculture, forestry, fishing, mining and quarrying industries are dropped as the location of these industries depends on natural resources.

Table 3.6. Logit estimation results for all occupations - six city categories

	Scarcity			
City size	Very High	High	Low	Very Low
Less than 5,000 inhabitants	-1.285***	-1.118***	-0.927***	0
	[0.209]	[0.201]	[0.190]	
5,000 - 20,000 inhabitants	-1.591***	-1.594***	-1.102***	0
	[0.207]	[0.205]	[0.199]	
20,000 - 50,000 inhabitants	-1.450***	-1.156***	-1.008***	0
	[0.207]	[0.198]	[0.189]	
50,000 - 100,000 inhabitants	-0.356	-0.296	-0.402**	0
	[0.219]	[0.209]	[0.197]	
100,000 - 500,000 inhabitants	-0.810***	-0.744***	-0.628***	0
	[0.200]	[0.196]	[0.187]	
More than 500,000 inhabitants	0	0	0	0

Note: n = 15,670. The estimation method is explained in Appendix B. Scarcity levels refer to the quartiles of scarcity level of occupations by industry. For each industry the occupations with the least (most) employment are defined as occupations with a very high (very low) scarcity level.

Figure 3.4. Distribution of number of performed subtasks

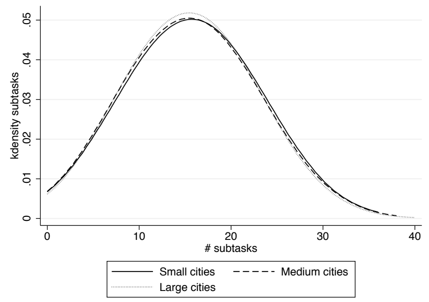


Figure 3.5. Distribution of number of performed subtasks - high-skilled workers

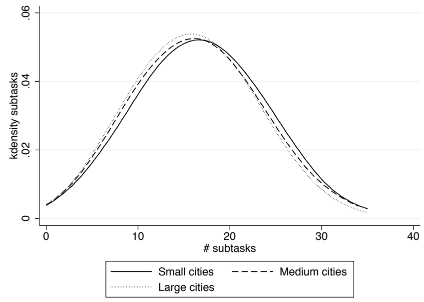


Figure 3.6. Distribution of demanded cognitive skills

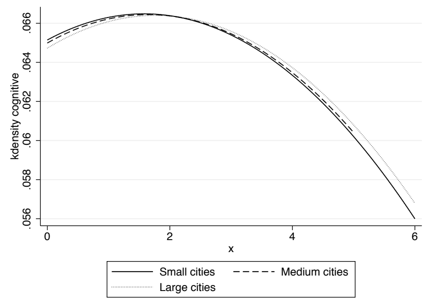
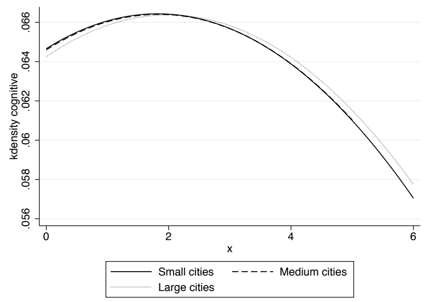


Figure 3.7. Distribution of demanded cognitive skills - high-skilled workers



3.4 Empirical strategy

Our empirical analyses consider the relation between the content of a job and the size of the city where the job is performed:

$$C_{i,o,j,l} = \alpha_1 + \alpha_2 C_{o,j} + \alpha_3 I_l + \alpha_4 S_i + \alpha_5 (I_l \cdot S_i) + \alpha_6 V_i + \epsilon_{i,o,j,l}. \tag{3.10}$$

$C_{i,o,j,l}$ refers to the job content, either the specialisation level or the demanded cognitive skills, of worker i with occupation o in industry j and city l. $C_{o,j}$ are job fixed effects controlling for the average job content. Furthermore, we control for the education level of worker i (S_i) and several other factors, such as age and gender (V_i). I_l refers to the main variables of interest; dummy variables indicating whether worker i lives in a small, a medium or a large city. To avoid underestimation of the standard errors we cluster them at the job level (Moulton, 1990). The observations are weighted by the national employment of the job.

A concern with this empirical strategy is the possible impact of measurement error. First, the task packages of individuals are determined in such a way that each worker could perform all tasks. The underestimation of the range of job tasks likely affects the estimations. The more individual job tasks exist, the more spatial variation is possible. Therefore, we expect this measurement error to create a underestimation of the spatial variation. Both the model in Section 3.2 and the estimation as described in equation (3.10) do not take into account the relative position and importance of job tasks within the production process. The benefits of specialisation and the demand for cognitive skills likely vary with the task package of jobs. We assume that we take this variation into account by including job fixed effects. Section 3.6 displays the spatial variation of job contents across separate broad occupational groups. Furthermore, Section 3.6 shows additional analyses with different indicators for specialisation and demanded skills.

Second, we observe the location of residence of the worker. No information is available about the working location. We assume that the worker lives and works in the same city. Most commuting workers in Germany commute from a small city to a larger city (Patuelli et al., 2010). We expect workers in larger cities to perform fewer subtasks and to possess more cognitive skills. If a worker lives in a small city but works in a large city, the fewer subtasks / more cognitive skills are classified under the small city but should be related to the large city. Again, this suggests that we are more likely to underestimate the spatial variation than to overestimate the actual spatial variation. In Section 3.6 we test several different spatial units to see whether the results are sensitive to measures of city size.

Third, skilled workers sort into large cities. Workers with high observed and unobserved abilities sort into larger cities for better education, career possibilities, spouse markets and amenities (Glaeser & Maré, 2001; Berry & Glaeser, 2005; Combes et al., 2008; Venables, 2011). These sorting patterns have consequences for the demand for cognitive skills in cities. Higher demand for cognitive skills in cities might reflect specialisation benefits but it might also reflect sorting patterns of more able workers. Our data limit us to proxy worker skills with education levels. Section 3.6 presents the relation between specialisation and cognitive skills and separate analyses for educational groups. Again, workers with strong unobserved skills might be more specialised which results in a relation between specialisation level of and required cognitive skills for the job. We cannot rule out the role of sorting of workers into jobs and locations. The explanation of spatial variation in job contents is therefore left for further research, here we focus on examining whether there is spatial variation in job contents.

Fourth, Becker & Murphy (1992) indicate that especially coordination costs affect the division of labour. When workers divide the complementary tasks of the production of a good, they need to coordinate the production process. Even if workers fully cooperate and do not compete to some extent, information about the tasks will be lost within the coordination process. Tacit knowledge about tasks is difficult to transfer across different workers. Coordination costs hinder workers to perform a unique subset of tasks and fully exploit the increasing returns. The model in Section 3.2 considers a continuum of tasks, resulting in a unique subset of tasks for each worker in the market. There is no overlap in the worker's subset of tasks. In reality most workers do not perform a unique subset of tasks. Moreover, as our dataset only contains 58 tasks, unique subsets are not possible. The demand for and supply of scarce tasks and products rise with local population (Duranton & Jayet, 2011). This suggests that possibilities to benefit from specialisation rise with population as well.

Lastly, characteristics of cities such as the share of high-skilled workers, the industrial structure and the amount of amenities likely affect the demand for certain tasks and with that the division of labour (Baumgardner, 1988b). Unfortunately, the only location information in the BIBB is a categorised size variable. We do not observe the name of the location or any other characteristics besides size. To test the impact of industrial structure, analyses in Section 3.6 make the distinction between manufacturing and service sectors.

3.5 Job contents across cities

3.5.1 Specialisation level

We start our empirical analysis by examining whether the division of labour is bound by the extent of the market. The subset of tasks is different for each job. The estimation of a simple regression explaining the number of subtasks by the job results in an adjusted R-squared of 0.26. The correlation between number of subtasks performed by a worker and the average number of subtasks of the job is 0.45 (significant at the 1 percent level). We confirm the notion of Autor & Handel (forthcoming) that measures of task composition at the occupational level obtain substantial measurement error.

As expected, workers in large cities perform fewer subtasks and are more specialised given their job (column (1) in Table 3.7). We distinguish three size categories of local population: small (less than 20,000 inhabitant), medium (between 20,000 and 100,000 inhabitants) and large cities (more than 100,000 inhabitants). The level of specialisation of a certain job increases linearly with city size. The variation in specialisation across city size is significant, but only modest in terms of size. Workers in large cities perform 5 percent of a standard deviation fewer subtasks than workers in small cities.

Table 3.7. The level of specialisation is higher in cities

	Number of subtasks			
	(1)	(2)	(3)	(4)
Medium city	-0.043**	-0.045**	-0.048**	-0.047**
	[0.019]	[0.019]	[0.021]	[0.021]
Large city	-0.053***	-0.065***	-0.062***	-0.070***
	[0.020]	[0.020]	[0.023]	[0.023]
Unskilled		-0.218***		-0.282***
		[0.079]		[0.092]
Medium skilled		0.103***		0.109**
		[0.040]		[0.045]
High skilled		0.167***		0.254***
		[0.041]		[0.050]
Age		-0.002**		-0.003**
		[0.001]		[0.001]
Female		-0.115***		-0.198***
		[0.016]		[0.027]
Native speaker		0.085**		0.104***
		[0.034]		[0.038]
Job average	1.000***	0.957***		
	[0.001]	[0.008]		
Constant	0.030***	0.003	0.026**	0.010
	[0.010]	[0.064]	[0.012]	[0.078]
Job fixed effects			YES	YES
Observations	15,670	15,670	15,670	15,670
Adjusted R-squared	0.193	0.202	0.001	0.019

Note: individual data. Table B.2 in Appendix B displays the definitions of the variables. Clustered standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Other personal characteristics affect the level of specialisation as well. The spatial distribution of skilled workers is unequal. High-skilled workers are overrepresented in large cities. If high-skilled workers perform on average more (fewer) subtasks this will underestimate (overestimate) our results. The same reasoning holds for the spatial distribution of young workers, females and non-native speakers. The regression in column (2) includes information on the education level and demographic characteristics of the worker. High-skilled workers perform on average more tasks than low-skilled workers, probably caused by higher productivity levels. Workers with a college or university degree perform 17 percent of a standard deviation more subtasks than workers with only a high school degree. Controlling for the education level of the workers increases the impact of the city size dummies. The unequal spatial distribution of high-skilled workers underestimated the variation across cities in column (1). The number of subtasks varies significantly across other subgroups as well. Older workers perform fewer tasks than younger workers, females perform fewer tasks than males and native speakers more than non-native speakers. Likely, variation in the trade-off between coordination costs and efficiency benefits for specialisation causes these variations (Becker & Murphy, 1992).

To firm-up our results, the estimates in column (3) and (4) include job fixed effects. The estimates in these columns measure the spatial variation within jobs. The coefficients of the medium-sized and large-sized city dummies remain significant and negative. The size of the coefficients increases slightly. Workers in a large city perform 7 percent of a standard deviation fewer subtasks than workers with the same job in a small city. The explanatory power of the estimation is very low which suggests that job codes do not explain a large part of the variations of job contents.

3.5.2 Demanded cognitive skills

Workers who specialise in a smaller subset of tasks gain benefits from increasing returns to scale. The more time a worker spends on a certain task, the more skills he develops to perform this task. Specialists tend to perform more complex and cognitive tasks than workers who perform more tasks. The specialisation level of jobs increases with city size, as indicated in the previous section, so we expect the demand of cognitive skills to increase with city size as well.

The estimates in column (1) in Table 3.8 show that jobs demand more cognitive skills when they are performed in large cities than when they are performed in small cities. Workers in large (medium) cities indicate that their job requires 8 percent (4.5 percent) of a standard deviation more cognitive skills than workers with the same job in small cities. The regression in column (2) includes additional demographic and education information of the workers. Females indicate that their job demands more cognitive skills than males with the same job do. The jobs of young

Table 3.8. Jobs demand more cognitive skills in cities

	Demanded cognitive skills			
	(1)	(2)	(3)	(4)
Medium city	0.045**	0.039**	0.051**	0.043**
	[0.019]	[0.019]	[0.021]	[0.021]
Large city	0.081***	0.067***	0.093***	0.079***
	[0.021]	[0.021]	[0.024]	[0.024]
Unskilled		0.035		0.029
		[0.055]		[0.065]
Medium skilled		-0.102**		-0.127**
		[0.044]		[0.051]
High skilled		-0.034		-0.046
		[0.045]		[0.054]
Age		-0.006***		-0.007***
		[0.001]		[0.001]
Female		0.126***		0.195***
		[0.022]		[0.033]
Native speaker		-0.019		-0.022
		[0.034]		[0.039]
Job average	0.994***	0.960***		
	[0.002]	[0.009]		
Constant	-0.040***	0.246***	0.051***	0.350***
	[0.010]	[0.064]	[0.011]	[0.076]
Job fixed effects			YES	YES
Observations	15,670	15,670	15,670	15,670
Adjusted R-squared	0.202	0.210	0.002	0.015

Note: individual data. Table B.2 in Appendix B displays the definitions of the variables. Clustered standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

and non-native workers require more cognitive skills than the same jobs of older and native workers. The coefficients for city size are not affected by the inclusion of these additional factors.

Lastly, columns (3) and (4) include job fixed effects and focuses on the estimation of the spatial variation within jobs. Again, the explanatory power of the regressions drops. The city size coefficients slightly increase in size and remain significant and negative. Workers in large cities indicate that their job demands 8 percent of a standard deviation more cognitive skills than workers with the same job and characteristics in small cities.

In summary, the specialisation level and demanded cognitive skills of jobs increase with city population. The spatial variation in job contents is significant but modest. As discussed in the previous section, the set up of the survey likely causes underestimation of the variation in job content. We find significant spatial variation in the content of jobs and expect the actual variation to be larger. It should be noted that in most countries, the largest cities house the best workers (as discussed in Section 3.4). We expect the sorting of the more skilled workers to affect the skill demand as well.

3.6 Further analyses

Previous estimates include several assumptions which we test in this section with five robustness checks. A first concern of the previous estimates is the impact of measurement error. We start by testing the sensitivity of the indicators for specialisation and cognitive skills (Section 3.6.1). Second, we deal with the regional measurement error in Section 3.6.2 and run estimations with different spatial units. Third, the possible impact of sorting of more able workers is discussed in Section 3.6.3. Section 3.6.4 provides separate estimates for the manufacturing and service sectors and for eight broad occupational groups to indicate whether spatial variation in job contents is present across the whole economy. Lastly, we test the hypothesis that learning and experience could affect the results. We only show estimates including job fixed effects.[5]

3.6.1 Indicators for specialisation and cognitive skills

Measuring task packages of jobs is challenging. As described in Section 3.3, we use a broad interpretation of 'tasks' and define specific skills and activities as tasks as well. Although common in the literature, this is an arbitrary choice. We test the sensitivity of our results towards the choice of included tasks and construct alternative measures of our indicators. Our alternative measure for the job's specialisation level only includes task information and does not include required cognitive and specific skills any more (see Table 3.1 for an overview of the included tasks). The alternative measure of the demanded cognitive skills includes only the information on the question about the importance of required cognitive skills and not the other tasks such as 'doing research'. Columns (1) and (2) in Table 3.9 present the results. Our results are not sensitive to the measurement of our indicators. However, the demand for cognitive skills is only significantly larger in large cities when we control for other factors as well. Our dataset likely underestimates the range

[5] OLS-estimates show similar results and are available upon request.

of job tasks which likely results in a underestimation of the spatial variation in job tasks as well. The estimates in Table 3.9 indicate that measurement error does not drive our results.

3.6.2 Spatial units

The empirical analyses define a city as the local labour market. Spatial units are chosen for convenience and nothing guarantees that the city is indeed the correct aggregation level for local demand and supply of tasks. We classified the cities into three categories: small, medium and large cities. Column (3) of Table 3.9 presents estimates for the level of specialisation in which we distinguish seven city size categories. The number of subtasks a worker performs diminishes with the size of the city of residence. Column (4) presents estimates with the same city categories for the cognitive skill demand. The importance of cognitive skills increases with city size.

Column (5) and (6) present the results when applying alternative spatial units. Instead of measuring the size of the city we measure the population density of the region. In a region with a high population density it is easier to cooperate and to learn from your peers than in a region with a low population density. Therefore, population density of 16 German regions ('Bundesländer') provides us information on the possibilities to cooperate and divide the tasks among workers. Column (5) presents the results with this measure of size (again classified into three categories: low, medium and high density) for estimations of the level of specialisation. Workers in high density areas perform fewer subtasks and are more specialised than workers in low density areas. Jobs in dense areas also demand more cognitive skills (column (6)).

3.6.3 Sorting of more skilled workers

In most countries the largest cities house the 'best' workers. Combes et al. (2008) show that the sorting of these better workers into large cities is only partly captured by educational differences. Unobserved skills, such as cognitive skills, play a key role in spatial wage disparities. Concerning our results it might be the case that these better workers who sort into large cities are more specialised and have more cognitive skills. In other words: our results may be driven by characteristics of workers who sort into large cities instead of job characteristics based on market efficiency. We test this hypothesis in two ways. First, we analyse the relation

Table 3.9. Further analyses

	Alternative measures		Alternative spatial units				Skills and specialisation	
			Seven city sizes		Inhabitants per square km			
	Specialisation	Cognitive	Specialisation	Cognitive	Specialisation	Cognitive	Cognitive	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Medium city	-0.047**	0.035*			-0.057***	0.064***		0.024
	[0.021]	[0.019]			[0.019]	[0.021]		[0.020]
Large city	-0.047**	0.045**			-0.077**	0.081**		0.050**
	[0.020]	[0.020]			[0.036]	[0.039]		[0.021]
5,000–20,000 inhabitants			-0.055*	0.067**				
			[0.031]	[0.030]				
20,000–50,000 inhabitants			-0.060*	0.071**				
			[0.031]	[0.032]				
50,000–100,000 inhabitants			-0.124***	0.116***				
			[0.037]	[0.039]				
100,000–500,000 inhabitants			-0.094**	0.097***				
			[0.037]	[0.029]				
500,000– .. inhabitants			-0.117***	0.149***				
			[0.036]	[0.037]				
Subtasks							-0.335***	-0.411***
							[0.014]	[0.012]
Other factors	YES	YES	YES	YES	YES	YES		YES
Constant	0.061	0.023	0.050	0.301***	-0.001	0.361***	0.004	0.354***
	[0.074]	[0.077]	[0.084]	[0.084]	[0.077]	[0.077]	[0.007]	[0.065]
Job fixed effects	YES	YES	YES	YES	YES	YES		YES
Observations	15,670	15,670	1,591	1,591	1,591	1,591	15,670	15,670
Adjusted R-squared	0.007	0.048	0.019	0.016	0.019	0.015	0.307	0.174

Note: individual data. Subtasks in column (1) refer to the number of job tasks, job characteristics and task demand which appears ’sometimes’ or ’rarely’ in the worker’s job. Cognitive skills in column (2) are measured by the number of cognitive skills that are ’always’ or ’often’ demanded in the worker’s job. City size is defined by dummy variables, sizes are defined as in Table 3.2 for columns (3) and (4). In columns (5) and (6) size refers to the density level of the region (’Bundesländer’): size 2 refers to medium population density, size 3 to high population density. Size 1 is the reference group and refers to low population density. Table B.2 in Appendix B displays the definitions of the variables. Regressions include controls for education, gender, age and whether the worker’s mother tongue is German. Standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

between cognitive skill demand and specialisation level of workers. Such a relation does not need to be causal. Therefore, we estimate spatial variation in job contents for separate educational groups as an additional sensitivity check.

If the cognitive skill demand depends on the job's specialisation level this should be visible in the estimation of cognitive skills. In column (7) in Table 3.9 we include the job's specialisation level into the estimation of demanded cognitive skills. There is a strong and negative relation between the demanded cognitive skills of a job and the performance of subtasks. Workers who focus more on core tasks indicate that their job requires more cognitive skills. This relation remains significant when we control for other factors (column (8)). The coefficients of the size dummies slightly change when we include specialisation level of the worker. If we control for the worker's specialisation level, the spatial variation of the cognitive skill demand increases. The coefficient of medium cities however becomes insignificant with the inclusion of the number of subtasks.

Workers with relatively many unobserved skills may both have more cognitive skills and be more specialised. The results in column (7) and (8) of Table 3.9 therefore could reflect the higher specialisation level of more capable people. We assume that the sorting of more capable people into cities is partly captured by analysing the sorting of observed skills. If we find spatial variation in job contents of other skill groups this suggests that the spatial variation captures more than sorting of the most capable workers. Table 3.10 presents separate estimations for four educational groups: unskilled, low-skilled, middle-skilled and high-skilled workers. Columns (1) to (4) show that workers within all educational groups perform significantly fewer subtasks when they are located in large cities. Columns (5) to (8) show that workers within all skill groups, except the unskilled, indicate that their job demands more cognitive skills when they are located in a large city. We conclude that the division of job tasks is beneficial for all skill groups. The fact that also unskilled and low-skilled workers specialise more in large cities suggests that our results are not solely driven by sorting patterns.

3.6.4 Variation across industry and occupational groups

The production processes of jobs vary across sectors and occupational groups. For instance, the local market for the demand and supply for service products may be much more local (smaller) than the market for manufacturing products. Furthermore, the importance of tacit knowledge within service sectors causes coordination costs to be higher in service sectors than in manufacturing sectors. The relatively

Table 3.10. By educational group

	Subtasks				Cognitive skills			
Skill level	No	Low	Medium	High	No	Low	Medium	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Medium city	-0.004	-0.056**	-0.045**	-0.045**	-0.003	0.056**	0.044**	0.044**
	[0.036]	[0.024]	[0.021]	[0.021]	[0.038]	[0.025]	[0.021]	[0.021]
Large city	-0.079**	-0.062**	-0.064***	-0.064***	0.021	0.098***	0.085***	0.085***
	[0.038]	[0.027]	[0.023]	[0.023]	[0.035]	[0.027]	[0.024]	[0.024]
Age	-0.006***	-0.002	-0.003**	-0.003**	-0.005***	-0.007***	-0.007***	-0.007***
	[0.002]	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]	[0.001]	[0.001]
Female	-0.270***	-0.197***	-0.209***	-0.209***	0.187***	0.191***	0.191***	0.191***
	[0.053]	[0.031]	[0.028]	[0.028]	[0.049]	[0.037]	[0.032]	[0.032]
Native speaker	0.116*	0.111**	0.110***	0.110***	-0.014	-0.045	-0.036	-0.036
	[0.060]	[0.045]	[0.037]	[0.037]	[0.054]	[0.047]	[0.039]	[0.039]
Constant	0.352***	0.103	0.151**	0.151**	-0.083	0.350***	0.267***	0.267***
	[0.095]	[0.075]	[0.062]	[0.062]	[0.076]	[0.072]	[0.059]	[0.059]
Job fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	511	608	9,045	5,506	511	608	9,045	5,506
Adjusted R-squared	0.014	0.010	0.010	0.010	0.009	0.015	0.013	0.013

Note: individual data. Table B.2 in Appendix B displays the definitions of the variables. Standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

low importance of tacit knowledge enables manufacturing firms to split up their production process easier between workers and even across space (Glaeser & Ponzetto, 2010). Therefore, the spatial unit of interest might also vary between manufacturing and services. Similarly, we expect benefits from specialisation and cognitive skill demand to vary across occupational groups. These sectoral differences likely result in different spatial patterns.

Table 3.11 presents separate estimates for manufacturing and services to assess whether our results hold for both type of industries. Columns (1) and (2) show that the specialisation level of jobs in manufacturing and of jobs in services are higher in larger cities than in small cities. Workers in the manufacturing (services) perform about 9 percent (7 percent) of a standard deviation fewer subtasks when they are located in a large city. Columns (3) and (4) present the same exercise for the requirement of cognitive skills. Within the sample of manufacturing sectors, the city size coefficient becomes insignificant while the one in the service sector sample remains positive and significant. This result suggests that benefits from specialisation in large cities occur in both industry types. Specialisation does not lead to a higher demand for cognitive skills in the manufacturing sector. This result could be caused by the focus of manufacturing on product-producing while services rely more on cognitive intense idea-producing (Glaeser & Ponzetto, 2010).

Table 3.12 presents separate estimates for eight broad occupational groups (one-digit). As expected, not all occupational groups experience spatial variation in their job contents. Within the samples of professional, service and craft occupations the coefficient of a large city is significant and negative. These occupational groups seem to benefit most from specialisation possibilities in cities. Noticeable is the negative and significant coefficient for medium-sized cities within the technical occupations. This confirms the theory of Duranton & Puga (2001) and Desmet & Rossi-Hansberg (2009) that medium-sizes cities focus on technical specialisation. Columns (9) to (16) present the same estimates for the cognitive skill demand. A similar spatial variation pattern is found. Professional and service occupations require more cognitive skills when they are performed in large cities, while technical occupations require more skills in medium-sized cities.

3.6.5 Learning and experience

The task packages of workers vary with age (Autor & Dorn, 2009). During their career, workers specialise and become experts on a subset of core tasks (Lazear, 2009). Experience probably leads to more expert knowledge and with that to more

Table 3.11. Manufacturing and service sectors

	Subtasks		Cognitive skills	
	Manufacturing	Services	Manufacturing	Services
	(1)	(2)	(3)	(4)
Medium city	-0.008	-0.057**	-0.005	0.055**
	[0.036]	[0.024]	[0.039]	[0.025]
Large city	-0.088**	-0.067**	0.018	0.092***
	[0.036]	[0.026]	[0.036]	[0.027]
Unskilled	-0.007***	-0.002	-0.005***	-0.007***
	[0.002]	[0.001]	[0.002]	[0.001]
Medium skilled	-0.241***	-0.189***	0.189***	0.195***
	[0.050]	[0.030]	[0.048]	[0.037]
High skilled	0.084	0.110**	0.001	-0.032
	[0.063]	[0.047]	[0.054]	[0.047]
Age	-0.155	-0.308**	0.121	0.003
	[0.101]	[0.122]	[0.082]	[0.083]
Female	0.219***	0.086*	-0.057	-0.144**
	[0.071]	[0.052]	[0.075]	[0.060]
Native speaker	0.487***	0.206***	0.003	-0.058
	[0.085]	[0.057]	[0.081]	[0.064]
Constant	0.122	-0.017	-0.051	0.445***
	[0.127]	[0.092]	[0.109]	[0.093]
Job fixed effects	YES	YES	YES	YES
Observations	6,583	9,087	6,593	9,087
Adjusted R-squared	0.032	0.017	0.010	0.016

Note: individual data. Table B.2 in Appendix B displays the definitions of the variables. Standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Table 3.12. Occupational groups

	Subtasks								Cognitive skills							
	Managers	Professionals	Technicians	Clerks	Service workers	Craft workers	Operators	Elementary occupations	Managers	Professionals	Technicians	Clerks	Service workers	Craft workers	Operators	Elementary occupations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Medium city	-0.014	-0.061	-0.088**	0.022	-0.011	-0.012	-0.035	-0.182	0.097	0.005	0.081**	-0.018	0.008	0.058	-0.036	0.160**
	[0.089]	[0.052]	[0.036]	[0.064]	[0.040]	[0.060]	[0.104]	[0.126]	[0.125]	[0.027]	[0.038]	[0.034]	[0.041]	[0.058]	[0.060]	[0.076]
Large city	-0.068	-0.092***	-0.048	-0.040	-0.149*	-0.110*	-0.011	-0.002	0.143	0.083*	0.039	0.133**	0.121***	-0.024	0.023	0.175
	[0.069]	[0.026]	[0.035]	[0.086]	[0.076]	[0.062]	[0.149]	[0.093]	[0.115]	[0.045]	[0.031]	[0.054]	[0.040]	[0.062]	[0.121]	[0.104]
Unskilled	-0.007	0.000	-0.002	-0.002	-0.006*	-0.005*	-0.010***	-0.012**	0.009***	-0.012***	-0.010***	-0.001	-0.001	-0.007***	-0.004	-0.004
	[0.004]	[0.002]	[0.002]	[0.003]	[0.003]	[0.003]	[0.003]	[0.005]	[0.003]	[0.002]	[0.001]	[0.002]	[0.003]	[0.003]	[0.005]	[0.004]
Medium skilled	-0.357***	-0.077	-0.163***	-0.228***	-0.445***	-0.289***	-0.281**	-0.844***	0.168*	0.214***	0.117**	0.157***	0.247***	0.291***	0.342***	0.322***
	[0.060]	[0.067]	[0.041]	[0.055]	[0.063]	[0.068]	[0.125]	[0.130]	[0.087]	[0.060]	[0.047]	[0.055]	[0.057]	[0.078]	[0.084]	[0.058]
High skilled	-0.055	0.071	0.097*	0.136	0.067	0.277**	0.038	0.317**	0.200	0.156	-0.061	-0.131	-0.131*	0.090	-0.170*	-0.046
	[0.248]	[0.096]	[0.053]	[0.124]	[0.057]	[0.109]	[0.120]	[0.149]	[0.202]	[0.100]	[0.073]	[0.133]	[0.069]	[0.133]	[0.084]	[0.124]
Age	0.061	-0.283	-0.495***	0.038	-0.644***	-0.107	-0.432	-0.253	-0.834**	-0.208	0.082	-0.053	0.228**	0.191	0.156	0.109
	[0.497]	[0.519]	[0.110]	[0.120]	[0.231]	[0.177]	[0.385]	[0.214]	[0.318]	[0.654]	[0.111]	[0.172]	[0.090]	[0.230]	[0.275]	[0.153]
Female	0.007	0.149	0.130	0.341***	0.007	0.408**	-0.183	0.114	0.002	-0.087	-0.113	-0.258**	-0.014	-0.083	-0.134	-0.026
	[0.141]	[0.099]	[0.088]	[0.088]	[0.122]	[0.159]	[0.238]	[0.131]	[0.132]	[0.185]	[0.075]	[0.119]	[0.117]	[0.203]	[0.144]	[0.146]
Native speaker	0.061	0.106	0.233***	0.606***	0.252	0.943***	0.119	0.295	0.019	-0.003	0.017	-0.198	0.042	-0.004	-0.207	-0.119
	[0.166]	[0.112]	[0.082]	[0.137]	[0.178]	[0.152]	[0.292]	[0.202]	[0.165]	[0.181]	[0.082]	[0.119]	[0.130]	[0.224]	[0.237]	[0.175]
Constant	0.807**	-0.079	0.045	-0.156	0.516*	-0.420*	0.187	-0.034	-0.516*	0.648***	0.549***	0.123	-0.062	-0.192	0.054	-0.022
	[0.375]	[0.089]	[0.159]	[0.189]	[0.258]	[0.241]	[0.404]	[0.221]	[0.268]	[0.165]	[0.146]	[0.154]	[0.120]	[0.252]	[0.206]	[0.291]
Job fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	869	3,462	4,136	1,817	1,458	2,112	991	798	869	3,462	4,136	1,817	1,458	2,112	991	798
Adjusted R-squared	0.032	0.006	0.019	0.033	0.063	0.073	0.032	0.154	0.030	0.029	0.016	0.019	0.017	0.018	0.045	0.055

Note: individual data. Occupations are defined by one-digit ISCO 1988 codes. Table B.2 in Appendix B displays the definitions of the variables. Standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Table 3.13. Age groups

	Subtasks			Cognitive skills		
Age	35^-	35-50	50^+	35^-	35-50	50^+
	(1)	(2)	(3)	(4)	(5)	(6)
Medium city	-0.029	-0.055*	-0.066	0.080	-0.005	0.119**
	[0.049]	[0.030]	[0.043]	[0.052]	[0.029]	[0.049]
Large city	-0.071*	-0.081***	-0.049	0.173***	0.058*	0.030
	[0.040]	[0.031]	[0.045]	[0.053]	[0.030]	[0.048]
Unskilled	-0.011*	0.000	-0.007	0.009	-0.010***	-0.007
	[0.007]	[0.003]	[0.005]	[0.007]	[0.003]	[0.005]
Medium skilled	-0.237***	-0.191***	-0.181***	0.224***	0.160***	0.256***
	[0.041]	[0.035]	[0.057]	[0.058]	[0.041]	[0.081]
High skilled	0.075	0.080	0.103	-0.008	-0.048	0.025
	[0.067]	[0.050]	[0.124]	[0.061]	[0.058]	[0.126]
Age	-0.020	-0.261**	-0.295	-0.064	0.032	0.197
	[0.215]	[0.119]	[0.191]	[0.175]	[0.103]	[0.145]
Female	0.188**	0.118*	0.083	-0.148	-0.106	-0.065
	[0.084]	[0.072]	[0.118]	[0.111]	[0.084]	[0.123]
Native speaker	0.297***	0.249***	0.292**	-0.088	-0.037	0.051
	[0.086]	[0.076]	[0.128]	[0.111]	[0.081]	[0.127]
Constant	0.233	-0.076	0.145	-0.183	0.520***	0.213
	[0.205]	[0.130]	[0.341]	[0.219]	[0.139]	[0.356]
Job fixed effects	YES	YES	YES	YES	YES	YES
Observations	3,600	8,518	3,552	3,600	8,518	3,552
Adjusted R-squared	0.019	0.016	0.022	0.016	0.009	0.018

Note: individual data. Table B.2 in Appendix B displays the definitions of the variables. Standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

specialisation. Here, we test whether the found results hold for all age groups.

Table 3.13 presents results for separate age groups. For all age groups, the number of performed subtasks decreases with city size (columns (1) to (3)). Within the group of workers above 50 years this relation is not significant. Columns (4) to (6) present separate estimates for the cognitive skill demand for the three age groups. Again, our main findings hold for all the age groups but the large city coefficient for workers above 50 years is insignificant. Workers of all ages below 50 indicate that their job consists of fewer subtasks and demands more cognitive skills when they are located in a large city. The spatial variation in cognitive skill demand is the strongest for young workers in the beginning of their career.

3.7 Concluding remarks

This chapter shows that a job contains a different task package in a large city compared to the same job in a small city. Our theoretical model suggests that the spatial variation in job contents is the result of a stronger division of labour in large cities. The empirical analyses indicate that both the specialisation level of jobs and the demand for cognitive skills rise with city size.

Most research ignores the possible spatial variation in job contents. Our indicators rely on very broad tasks and measure spatial variation of job contents and might underestimate the variation. The fact that we do find spatial variation in job contents despite this possible underestimation suggests a substantial spatial variation.

Regional inequality is a hot policy topic. We take a step towards unravelling the inequality in wages and productivity. Further steps, especially in more adequately separating sorting and productivity effects, is an important challenge for research.

Chapter 4

Matching worker skills to job tasks

Sorting into cities for better careers

4.1 Introduction

The matching of workers to jobs is better in thick labour markets than in thin ones. The benefits of thick labour markets first gained attention with the work of Alfred Marshall (1920). A thick labour market is associated with both a better chance of a job match and better match quality. An extensive literature has studied whether the chances of a job match rise with market size. The empirical evidence is, however, ambiguous, see (Petrongolo & Pissarides, 2001). Both workers and employers likely raise their match standards when they have more choice. This results in constant returns to scale for the matching chance and increasing returns to scale in match quality (Petrongolo & Pissarides, 2006). Empirical work on the quality of matches is scarce, mainly because it is hard to define quality (Rosenthal & Strange, 2004). In a first attempt, Petrongolo & Pissarides (2006) proxy the quality of the match using wages.

The present chapter compares the quality of matches between thick city labour markets and thin ones in the Netherlands. The extent to which the skills of workers suit their job tasks is used to define job match quality. Heterogeneity in both worker skills and job tasks is considered in match quality, in addition to commonly used education level and occupation codes. This chapter thus extends the work of Petrongolo & Pissarides (2006) by applying a more detailed measure of match quality. We find that the quality of the match is indeed significantly better in Dutch cities than in the Dutch countryside. The better career prospects induce better workers and more complex jobs to gravitate to cities.

To assign skills to tasks across labour markets, we propose a model in the spirit of Burdett & Coles (1997) and Gautier et al. (2010). The model considers heterogeneous workers searching for a job and employers holding heterogeneous jobs for which they are seeking workers. Workers seek the most complex and subsequently best-paying jobs they can obtain. Employers seek the most skilled workers willing to accept the job, since more skilled workers are more productive. The ’distance’ between worker skills and job complexity determines the quality of the match: the smaller the distance, the better the match. Workers and jobs are divided into quality segments. Hence, the maximum difference between worker skills and job complexity is the difference between the least (most) skilled worker in a segment and the most (least) complex job in the same segment. Workers and employers choose a location to work/operate in before they start their search. The economy has two locations: a scarcely populated countryside and a densely populated city. The density of the city results in a better match quality but also in higher rents. Because of these better matching qualities, the expected utility of the matches depends more on the quality of workers and jobs in cities than the more ’random’ assignments in the countryside. Relatively more skilled workers and more complex jobs sort into the city, since they have higher opportunity costs. The advantages of better matches soon exceed the disadvantage of higher rents in the city.

Empirically, we employ the Longitudinal Internet Studies for the Social Sciences (LISS) panel of 3,000 Dutch individuals. The panel contains information about the suitability of skills for a person’s job and additional information about personalities, job tasks, and the usual demographic, occupational, and educational variables. In contrast with the commonly used Occupational Information Network (ONET) and Dictionary of Occupational Titles (DOT) datasets of job tasks, the LISS panel contains person-level instead of occupation-level information. As indicated by Autor & Handel (forthcoming), the within occupation differences in task packages are substantial, which makes our dataset relevant. Each respondent indicates the suitability of his or her job skills, the importance of 33 broad job tasks within the job, and statements about personality. The indicated suitability is used as an estimate of the quality of the match between the worker’s skills and job tasks. Information about preferences, as in preferring complex problems to simple problems, proxies for the investment a person has made in developing skills, given his or her education. We assume, for instance, that workers who prefer complex over simple problems invest more in their cognitive skills than workers with the same education who prefer simple problems. The importance of certain job tasks, given the occupation, defines the job’s complexity. In line with the work of Heckman et al. (2006), Borghans et al. (2006), and Bacolod et al. (2009), we can decompose skills and tasks into cognitive and social worker skills and job tasks. We define the quality of a match as the inverse gap between cognitive (social) skills and cognitive (social) job tasks.

Our results can be summarised as follows. The skills of workers in Dutch cities suit their job tasks better than the average suitability in the Dutch countryside. In addition, spatial variation in match quality exists within occupations. Given the occupation, the match of skills to tasks is 14 percent of a standard deviation better in cities than in the Dutch countryside. The spatial patterns for industrial occupations resemble that of service occupations but are less extensive. Regions outside the Randstad area show stronger spatial variation than those within the Randstad area, which operate more as a single regional labour market. As expected, more skilled workers sort into cities. Additional analyses suggest that work location choice for more skilled workers is mainly based on job opportunities. Learning mechanisms raise the skills of workers in cities only slightly more than in the countryside, but this does not explain the variation in match quality. Lastly, we show that better match quality is associated with higher wages. Thick labour markets in the Netherlands have advantages in terms of more productive matches.

Labour demand and supply matching is one of the three microfoundations of urban agglomeration economies, suggested by Duranton & Puga (2004), and a commonly cited source for agglomeration externalities. The frameworks of Helsley & Strange (1990), Kim (1990), and Kim (1991) generate externalities whereby the expected match quality increases with the size of the local market. The model of Duranton & Puga (2004) extends this mechanism by showing that the stronger competition for labour in cities results in additional agglomeration economies. Wheeler (2001) suggests that lower search costs in cities result in better matches, greater output per worker, more wage inequality, and higher expected returns to worker skills. Venables (2011) finds that the better match quality derives from the city's signalling function and crowding costs. The empirical evidence for these models is scarce. Petrongolo & Pissarides (2006) find positive scale effects in both post-employment and reservations wages. This study contributes to this work by analysing the spatial variation in the match between worker skills and job tasks. In a different field but using the same underlying mechanism, Costa & Kahn (2001) find that the overrepresentation of power couples in cities can be explained by better dual career possibilities with better chances and better match quality. Gautier et al. (2010) show that more attractive singles sort into cities for better matches.

The rest of the chapter is organised as follows. Section 4.2 proposes a matching model to guide empirical analyses about spatial variation in match quality. The strategy of these empirical analyses is presented in Section 4.3. Section 4.4 discusses the results of the empirical analyses. Section 4.5 presents some additional analyses to rule out other mechanisms and provides some back-of-the-envelope calculations of wage returns of match quality. Section 4.6 offers some concluding remarks.

4.2 Model

We consider a labour market in the spirit of Burdett & Coles (1997) and Gautier et al. (2010).[1] In the labour market, heterogeneous workers are assigned to heterogeneous jobs. Skill level characterises workers while complexity level characterises jobs. More skilled workers have a comparative advantage in more complex jobs. Skill and complexity level are indexed continuously: the shorter the distance between worker skills and job complexity, the better the quality of the match. Employers holding a vacancy seek the most skilled worker who wants the job, while workers search for the most complex job they can get. Our economy consists of two locations: the city, with a high density of agents, and the countryside, with a low density.[2] Workers (employers) decide where to work (operate) before they enter the market. Working (living) in the city is more expensive than working in the countryside. However, the thicker labour market of the city increases the possible matches for workers and employers, which tightens matches.

4.2.1 Basic setting

The model only considers searching workers and job openings. We assume that both workers and employers seek a 'lifetime' deal; hence nobody considers taking a job or filling a vacancy for just a few years.[3] Once a job or a worker is chosen, there is no turning back. Quitting or firing is ruled out. An agent's choice of location is indexed $l \in [0, 1]$, with $0 = \textit{countryside}$ and $1 = \textit{city}$. The countryside is a scattered

[1] The models of, for instance, Helsley & Strange (1990) and Duranton & Puga (2004) relate city formation to matching advantages. Our focus lies on scale effects in the match quality between heterogeneous workers and heterogeneous jobs. The main advantage of the dataset is the detailed information about the heterogeneity of workers, jobs, and matches, but information about location is limited. Therefore, we choose to set up a framework that focuses on scale effects for the match and does not explain city formation. Following Gautier et al. (2010) and Petrongolo & Pissarides (2006), we consider location characteristics as given.

[2] Note that the economic structure does not differ between the city and the countryside. The spatial division is not based on urban versus rural industries.

[3] If workers do consider future job opportunities or, for example, job opportunities after shocks, this would strengthen the advantage of the city as a location (Helsley & Strange, 1990; Strange et al., 2006).

location and its population density remains low, even if many work seekers and employers choose to be located there. City life is more expensive; $\Delta c = c_1 - c_0$ defines the additional costs in the city. These additional costs reflect higher housing prices (which are exogenous in this model) or the commuting price of travel from a cheap location to the city for work.[4]

The quality of workers is defined by their skill level a. We assume that a worker's skills are given and do not vary across locations. When a worker performs a job in the city, the skills are the same as when the job is performed in the countryside. We relax this assumption in the sensitivity analyses. Employers hold vacancies with complexity α. In addition, job complexity is static. Both workers and employers try to optimise their utility: workers search for the most complex and best paid jobs they can obtain. Workers maximise the nominal wage:

$$w(a, l) = \alpha - c_l. \tag{4.1}$$

More complex jobs pay more. All workers earn the same wage for a certain job, regardless of their skills. The variable c_l reflects the location costs of location l. Employers maximise their revenue and seek the most skilled worker willing to accept the job. The revenue depends on the skills of the worker, a, and the costs of the city:

$$r(\alpha, l) = a - c_l. \tag{4.2}$$

The revenue of the job increases with worker skills. For the employer, a more skilled worker is more valuable than a less skilled worker who needs additional job training. The amount of training costs required for the job and, in turn, the employer's revenue, decreases with worker skills (Helsley & Strange, 1990). For simplicity, we further assume that workers and employers face the same location costs.

4.2.2 Search segments

We now define the segments in which workers and employers search for possible matches. A worker with skills a who is willing to settle for a job with complexity α^* is also willing to settle for all jobs with $\alpha > \alpha^*$ as wages increase with complexity. Workers and jobs are classified into segments z, for example, labelled by educational categories. Each worker searches for a job within his or her segment

[4] The model ignores location choices based on social or living preferences. We admit that amenities can play a significant role in location choice and address this factor in the robustness section.

and each employer seeks a worker within the job's segment. Segments are exogenously given. The segments operate as 'labels' for workers and jobs. A worker with a university degree never accepts a job for a high school graduate and employers with a vacancy for a university graduate never invite a high school graduate to a job interview.[5] The labour market can be decomposed into a number of consecutive, non-overlapping segments. The first segment contains the workers with the highest skill levels and the jobs with the highest complexity and wages. Workers and employers never match outside their segments of the market.

Workers maximise their expected nominal wage, given their segment, while choosing a job and do not consider possible promotions or job changes:

$$w(a, l) = maxE_l[w(\alpha_z, l) - w(\alpha_z^-)] - c_l, \tag{4.3}$$

where α_z^- is the least complex job of segment z. A worker's wage is always positive, since accepting the worst-paying job is always more beneficial than remaining a job seeker: $w(\alpha^-) > 0$.

Similar, employers maximise job revenue, given the segment z of the job. Employers consider a one-time match for a lifetime. Once hired, a worker cannot be fired:

$$r(\alpha, l) = maxE_l[r(a_z, l) - r(a_z^-)] - c_l. \tag{4.4}$$

where $r(a_z^-)$ is the revenue the least skilled worker of the segment produces. For an employer, letting a worker perform the job is always more beneficial than leaving the job vacant: $r(a^-) > 0$. Note that in contrast with the standard model of Pissarides (2000) the value of being unemployed and the value of vacancy is zero. Let $S_{z,l}$ be the mass of job seekers and $V_{z,l}$ the mass of vacancies in segment z, location l. All job seekers and vacancies are 'new'; the number of seekers and vacancies is related only to the size of the market and not to market clearing. Larger markets have more seekers and vacancies: $S_{z,1} > S_{z,0}$ and $V_{z,1} > V_{z,0}$. The utility, in terms of wage or revenue, of a segment match is always positive for workers and for employers. Thus, the number of matches m of job seekers to vacancies is

$$m_{z,l} = min(S_{z,l}, V_{z,l}). \tag{4.5}$$

Given the number of vacancies and job seekers, the maximum number of matches

[5] This is a strong assumption to keep things simple. In reality, workers may find jobs outside their education segment.

in the local market is created.[6] If the number of job seekers in the market (z, l) exceeds the number of vacancies, all vacancies are filled and vice versa.

4.2.3 Match requirements

A match between a worker and an employer requires mutual agreement. This mutual agreement requires two conditions:

$$C1 : E_l w(\alpha_z, l) - c_l \geq w(\alpha_z^-, 0). \tag{4.6}$$

Condition C1 is the condition under which a worker in segment z is willing to accept a job with complexity α_z. The expected net wage of the job should equal or exceed the expected income of the least complex job of the segment in the countryside.

$$C2 : E_l r(a_z, l) - c_l \geq r(a_z^-, 0). \tag{4.7}$$

Condition C2 states that an employer holding a vacancy in segment z should be willing to let a worker with skills a_z perform the job. The revenue generated by the worker should equal or exceed the revenue of employing the least skilled worker in the countryside.

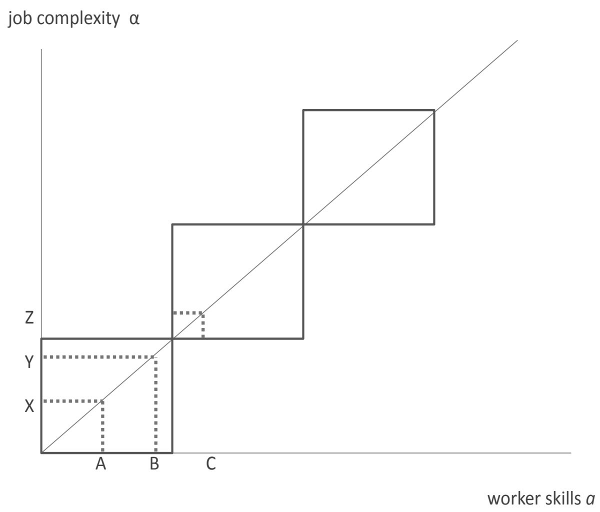
All job seekers face the same problem: the wage of the job with the lowest complexity they accept equals that of the least complex job of their segment performed in the countryside. The upper bound is formed by the job with the highest complexity and wage they are able to obtain. Hence, workers would accept a job in a higher segment. Condition 2 states, however, that an employer would not hire a worker from a lower segment. The range of job possibilities for a worker with skills a in segment z is bounded. The lower bound of the set of jobs for which a worker is willing to settle is bound by the lower bound α_z^- of matches and the upper bound α_z^+. The upper bound sets the job with the highest complexity the worker is able to obtain. The lower bound is the complexity for which condition C1 is just violated, while the upper bound is the highest complexity for which condition C2 holds. The range of possibilities reflects the jobs for which the worker is willing to settle (C1) and able to obtain (C2). The worker searches in the job set $\alpha \in [\alpha_z^-, \alpha_z^+]$. Employ-

[6] The standard model of Pissarides (2000) assumes a Cobb-Douglas function. Our matching function results from the assumption that both the value of being unemployed and the value of having a vacancy are zero. This assumption relates to our empirical analysis which contains a cross-section of matches which are already made. It therefore ignores unemployed workers and vacancies.

ers in the same segment z face a similar problem and search for workers in the set $a \in [a_z^-, a_z^+]$.

Figure 4.1 displays the labour market set-up. Workers are ranked by their skills on the horizontal axis and jobs are ranked by their complexity on the vertical axis. The diagonal represents optimal matches between worker skills and job complexity. The squares represent the market segments. For instance, the first square consists of all low-educated workers and all jobs for low-educated workers. All low-educated workers search for a job within the set of low-educated jobs, which are labelled as low educated and attract only low-educated workers. The label of low educated does not tell the whole story, however. Within the group of low-educated workers, skills vary. Although they are both low-educated, worker B has more skills than worker A, for example. Similarly, the complexity of jobs within the low-educated group varies and job Y is more complex than job X.

Figure 4.1. Matching



4.2.4 Match quality

The number of matches within a location and segment affects the quality of the match between worker skills and job complexity. Within a market with many matches, both agents have more match choices than in a low-density market. Since both parties maximise their utility, the distance between worker skills and job complexity is as small as possible. The density of vacancies and job seekers is higher in the city than in the countryside and both workers and employers are choosier in the city. Therefore, we assume the expected distance between complexity and skills to decrease with the number of matches in the market:

$$E[\alpha_{z,l} - a_{z,l}] = \frac{1}{E(Q_{z,l})} = \frac{1}{m_{z,l}} = \frac{1}{min(S_{z,l}, V_{z,l})}, \tag{4.8}$$

where $Q_{z,l}$ is the quality of the matches in segment z in location l. The intuition is simple: the chance of a worker having the required job skills is smaller in a thin market than in a thick market. When there are only a few workers and vacancies, the match of a worker to a job becomes less efficient.

The spatial variation in match quality results in spatial variation in the expected wage of a worker with skills a in segment z. In a market with better match quality, the gap between worker skills and job complexity is smaller. The thinner the market, the more friction within the matches and the less the expected wage depends on the worker's skills. Following this intuition, we assume

$$E_l[w(\alpha_z, l) - w(\alpha_z^-) - c_l] = (a_z - a_z^-)^{E(Q_{z,l})} + \delta^{1-E(Q_{z,l})}, \tag{4.9}$$

where $w(\alpha_z, l)$ is the expected wage for segment z at location l and $w(\alpha_z^-)$ is the minimum segment wage. The expected wage in a location depends on the worker skills and the match quality and costs of the location. The term $(a_z - a_z^-)^{E(Q_{z,l})}$ defines the part of the expected wage that depends on worker skills, namely, the skill difference between a worker and the least skilled worker in the same segment. The better the match quality in the location, the more the wage depends on the skill difference. The term $\delta^{1-E(Q_{z,l})}$ defines a randomly assigned additional wage caused by the sub-optimality of the matches in the location. The more agents in market z, l, the smaller the distance between skills and complexity and the more the wage difference reflects the skill difference; the importance of $(a_z - a_z^-)$ increases with $E(Q_{z,l})$. The rest of the wage, $\delta^{1-E(Q_{z,l})}$, is a randomly assigned disturbance term caused by a mismatch due to the friction.[7] As explained above, we assume the quality of the match to be better in the city and the impact of the friction to be larger in the countryside: $Q_1 > Q_0 > 0$ for all segments z.

The expected revenue for an employer with a vacancy with complexity α in

[7] By definition, the expected value of δ is 0.

segment z varies across the two locations and is defined similarly:

$$E_l[r(a_{z,l}) - r(a_z^-) - c_l] = (\alpha_z - \alpha_z^-)^{E(Q_{z,l})} + \delta^{1-E(Q_{z,l})}, \tag{4.10}$$

where $\alpha_z - \alpha_z^-$ is the difference in complexity between the job and the least complex job in the same segment.

4.2.5 Location choice

Both workers and employers choose their location before the matching moment. Workers maximise their nominal wage (equation (4.3)) given the conditions C1 and C2 and expected wages at both locations (equation (4.9)). A worker with skills a in segment z maximises

$$E_l w(a_z) = (a_z - a_z^-)^{E(Q_{z,l})} + \delta^{1-E(Q_{z,l})} - c_l. \tag{4.11}$$

The ratio between the expected nominal wage in the city and in the countryside is therefore:

$$\frac{E_1 w(a_z)}{E_0 w(a_z)} = \frac{(a_z - a_z^-)^{E(Q_{1,z})} - \Delta c}{(a_z - a_z^-)^{E(Q_{0,z})}}. \tag{4.12}$$

There is a trade-off between the better matching in the city, ($Q_{1,z} > Q_{0,z}$), and the additional costs Δc of working there. Because the expected value of the disturbance term δ is zero, this term does not affect the trade-off. The relative nominal wage in the city increases with worker skills a_z. However, location costs are higher in the city (Δc) and decrease the nominal city wage. Workers who are relatively skilled, given their segment, benefit more from the better matching in the city than less skilled workers in their segment do. Hence, the less skilled a worker is, the lower the additional costs Δc need to be to locate in the countryside. At a given Δc, there exists a value a_z^* for which all workers with $a_z < a_z^*$ locate in the countryside and all workers with skills $a_z > a_z^*$ locate in the city.

Employers maximise their revenue at a location (equation (4.4)) given the conditions C1 and C2 and expected revenues at both locations (equation (4.10)). An employer with a vacancy with complexity α in segment z maximises

$$E_l r(\alpha_{z,l}) = (\alpha_z - \alpha_z^-)^{E(Q_{z,l})} + \delta^{1-E(Q_{z,l})} - c_l. \tag{4.13}$$

The ratio between the expected nominal revenue in the city and in the countryside is

$$\frac{E_1 r(\alpha_z)}{E_0 r(\alpha_z)} = \frac{(\alpha_z - \alpha_z^-)^{E(Q_{1,z})} - \Delta c}{(\alpha_z - \alpha_z^-)^{E(Q_{0,z})}}. \tag{4.14}$$

Employers face a similar trade-off between the better matching in the city ($Q_{z,1} > Q_{z,0}$) and the additional costs Δc of working in the city. Similar to the worker's problem, the less complex a job, the lower the additional location costs Δc need to be for the employer to locate in the countryside. At a given Δc, there exists a value α_z^* for which all vacancies with $\alpha_z < \alpha_z^*$ locate in the countryside and all vacancies with complexity $\alpha_z > \alpha_z^*$ locate in the city.

As in Gautier et al. (2010), there is an *elite city ordering*[8]: the better workers and more difficult jobs locate in the city to benefit from the better matching because their opportunity costs are higher. Note that in this model the elite ordering occurs within segments (e.g. education groups) and not between for example low and high educated workers.

Again, Figure 4.1 illustrates this mechanism. Workers A and B both search for a job in the lowest segment, segment 1. Since worker B has more skills than worker A ($a_1^B > a_1^A$), in an optimal match worker B's wage is higher than the one of worker A. The distance between worker B's optimal wage and the minimum wage in the segment is higher than that of worker A. Worker B therefore has more wage to lose in a thin market than worker A and a larger incentive to locate in the city than worker A. Worker C has more skills than worker B but operates in another segment. The difference between the optimal and minimum segment wage of worker C is lower than for worker B. Although worker C has more skills, worker C's wage depends less on the tightness of the match than the wage of worker B. Worker B has the strongest incentive of workers A, B, and C to locate in the city with a high $Q_{z,I}$.

4.2.6 Empirical predictions

The model suggests that the higher density in cities results in more productive and tighter matches. Workers with greater skills and employers with more complex vacancies have higher opportunity costs, because they simply have more to lose. In summary, the model results in three hypotheses:

1. Matches between worker skills and job complexity are better in the city than

[8] Note that the difference between locations reflects differences in the density of agents.

in the countryside. Rewriting equation (4.8), we obtain

$$E(Q_{z,l}) = min(S_{z,l}, V_{z,l}) \quad and \quad \frac{\partial E(Q_{z,l})}{\partial L_l} > 0, \tag{4.15}$$

where L reflects the binary location choice: $L_1 = 1$ and $L_0 = 0$.

2. Skilled workers are found more often in the city than in the countryside. Equation (4.12) implies a positive relation between skills and the wage difference between the city and the countryside:

$$\frac{\partial E\Delta w}{\partial a_z} > 0, \tag{4.16}$$

where $E\Delta w = Ew_{1,z} - Ew_{0,z}$ reflects the expected wage difference between the locations as defined in equation (4.12). The larger wage difference between locations results in a positive relation between worker skill level and city location since the cost differences are compensated more:

$$\frac{\partial E(a_z)}{\partial L_l} > 0. \tag{4.17}$$

3. More complex jobs are found more often in the city than in the countryside. Equation (4.14) implies a positive relation between skills and revenue difference between the city and the countryside:

$$\frac{\partial E\Delta r}{\partial \alpha_z} > 0, \tag{4.18}$$

where $E\Delta r = Er_{1,z} - Er_{0,z}$ reflects the expected wage difference between the locations as defined in equation (4.14). Similar to the case of skilled workers, this results in a positive relation between job complexity and location:

$$\frac{\partial E(\alpha_z)}{\partial L_l} > 0. \tag{4.19}$$

4.3 Empirical strategy

4.3.1 Data

We employ the LISS panel to empirically test the theoretical framework. The LISS panel is the core element of a project titled 'Measurement and Experimentation in the Social Sciences' from the Dutch research institute CentERdata. The panel consists of 5,000 households, with a total of 8,000 individuals. This household sample is a true representation, obtained from the Dutch population register. The survey involves no self-selection.

All panel members complete the questionnaires online and update their information monthly. Households without Internet access receive a computer with Internet access. About half of the yearly interview time is reserved for the longitudinal study. The other half is divided among additional questionnaires from researchers. This chapter uses data from one of these additional questionnaires: a survey about job tasks carried out in May 2012. The questionnaire aims to gain insight into the importance of job tasks, the location where workers learned these tasks, and how efficient workers are in performing these tasks. A total of 3,883 household members were asked to fill out the questionnaire, with a response rate of 71.6 percent (2,780 household members).

We match additional personal and career information from several studies of the LISS panel; the background study, the work and schooling study, and the personality study to this dataset. We drop all skilled agricultural, fishery, and forestry workers, since the locations of these occupations depend on natural resources. Only 29 individuals in the sample hold a job in this occupational group. Missing information about matching is replaced with the answers to the same question from the work and schooling study. A total of 13 respondents provided no matching information and 136 respondents provided no skill information. The ratio of city to countryside work location does not vary across missing and non-missing observations.

4.3.2 Variables

Worker skills. An education diploma displays the vast amount of skills a worker holds and defines the worker's segment. Skills tend to be occupation specific; therefore we also distinguish between broad occupational groups. The theoretical model suggests that worker skills vary among students within the same graduation class. Honours, such as student of the year, underline this assumption. Skill variation within an education segment is estimated by the worker's personality. The idea is that more ambitious workers tend to invest more in their own skills. Both cognitive and social skills seem to be important for job performance, as indicated by Heckman et al. (2006), Borghans et al. (2006), and Bacolod et al. (2009).

We measure cognitive investments by the inclusion of five survey statements about a worker's cognitive orientation (see Table 4.1). Scaling varies across statements, which we rescale into three categories: zero if the worker (strongly) disagrees, one if the worker is neutral, and two if the worker (strongly) agrees. The cognitive skills index is standardised with a mean of zero and a standard deviation of one. In the same spirit, we define the social capacity of workers, given their education. Workers with more socially oriented personalities will develop more suitable skills for the performance of social tasks. Table 4.1 presents the five social characteristics of our index. The index for social skills is standardised with a mean of zero and a standard deviation of one.[9]

Job complexity. The dataset does not contain employer information about vacancies or job characteristics. All job information is gathered from the worker. The indicated importance of several job tasks defines the job's complexity. Thus, we assume that workers in more complex jobs indicate higher task importance.[10] We distinguish between tasks that are crucial or very important for a job (core tasks) and tasks that are moderately or barely important for a job (subtasks). A job's complexity increases with the amount of core tasks. Again, we distinguish between cognitive and social job requirements. Table 4.1 defines eight cognitive and eight social job tasks. These tasks form a cognitive task index and a social task index, both standardised with a mean of zero and a standard deviation of one.

Matching. Match quality defines the gap between worker skills and job complexity: $\alpha_z - a_z$, as defined in Section 4.2. The smaller the gap between these two, the better the match. First, we include the question 'How do your knowledge and skills suit the work you do?'. The answer choices range from zero (do not suit my work at all) to 100 (suit my work perfectly). Both the survey about job tasks and the work and schooling study include this question. When the answers differ between the two questions, we use the mean of the two.[11] The second quality measure considers the gap between the importance of cognitive (social) job tasks and the worker's invested skills in such cognitive (social) tasks. The smaller the gap between the (standardised) importance and (standardised) skills, the better the worker is suited for the job. For comparability, we standardise the matching in-

[9] Another possible measure of skill variation could be the indicated effectiveness of workers in performing job tasks. This measure relates more to job tasks than to the degree of investment in developing skills. However, the survey asks about effectiveness directly after questions about the importance of job tasks. We are concerned about measurement error, since we assume workers will be reluctant to indicate they are ineffective in the performance of an important job task. This fear is underlined by weak correlations with other variables. Therefore, we exclude the information about effectiveness from our analyses.

[10] Self-reporting job tasks may induce measurement errors, as discussed in Section 4.3.4.

[11] For 9 percent of the sample, the difference between the two answers is more than one standard deviation.

dexes.

Location. The questionnaire includes two questions about location, one about the location of residence and the other about the work location. Both questions indicate the urban character of the location by its density. Five categories are distinguished by the amount of dwellings per square kilometre: extremely urban (more than 2,500 dwellings), very urban (1,500–2,500 dwellings), moderately urban (1,000–1,500 dwellings), slightly urban (500–1,000 dwellings), and not urban (fewer than 500 dwellings). In line with the theoretical model, we distinguish a city labour market and a countryside labour market. We generate a city dummy indicating whether the place of work consists of more than 1,500 dwellings per square kilometre.

Additional variables. Besides urban character, job complexity, worker skills, and the matching between these two, we include personal characteristics and wage information. Table 4.2 gives an overview of the dependent variables, while Table C.1 in Appendix C presents all the variables, measurements, and summary statistics.

4.3.3 Descriptive statistics

Tables 4.3 presents simple descriptive statistics for our dataset. On average, workers rate the suitability of their skills for their job as 69.86 on a scale of zero to 100, with a standard deviation of 20.09. This indicated suitability is higher among high-skilled workers and workers in cities than across low- and middle-skilled workers and workers located in the countryside. The matching of cognitive skills to cognitive job tasks shows a similar pattern. The cognitive skills of high-skilled workers match their cognitive job tasks better than the skills of low- and middle-skilled workers do. The cognitive job matches are better in the city than in the countryside. Table C.2 in Appendix C shows significant spatial variation. The tables also show a less clear pattern of the matching of social job tasks across skill groups and working locations. Table 4.4 presents the summary statistics by broad occupational groups. Professionals enjoy, on average, the best assignment of skills to jobs, while workers in elementary jobs indicate the worst matches. For cognitive skills the match is especially strong among managers and weak among operators and within elementary occupations.

A worker's cognitive skill level is estimated by the number of cognitive statements with which the worker agrees or strongly agrees. On average, a worker agrees with 0.23 statements out of 5. High-skilled workers agree with more state-

Table 4.1. Skill and task variables

Cognitive skills: cognitively oriented personality statements
1. I like to have the responsibility of handling a situation that requires a lot of thinking. 2. I prefer complex problems to simple problems. 3. I enjoy tasks that involve coming up with good solutions for new problems. 4. I prefer my life to be filled with puzzles that I must solve. 5. The notion of thinking abstractly is appealing to me.
Social skills: socially oriented personality statements
1. I'm interested in other people. 2. I make people feel at ease. 3. I have social recognition. 4. I start conversations. 5. I feel comfortable around other people.
Cognitive tasks
1. Knowledge of use or operation of tools/equipment machinery. 2. Solving problems. 3. Analysing problems. 4. Planning the work of others. 5. Reading long documents. 6. Writing short documents with correct spelling and grammar. 7. Writing long documents with correct spelling and grammar. 8. Simple calculations. 9. Calculations with math and/or statistics.
Social tasks
1. Dealing with people. 2. Working together or in a team. 3. Listening to other people. 4. Teaching people. 5. Making speeches/presentations. 6. Selling a product or service. 7. Persuading or influencing others. 8. Counselling, advising, or caring for customers or clients.

Table 4.2. Job content variables

Matching variables	
Matching all skills	How do your knowledge and skills suit the work you do?
Cognitive skills & cognitive tasks	Absolute difference between cognitive skills and cognitive task importance (see below).
Social skills & cognitive tasks	Absolute difference between social skills and social task importance (see below).
Skill variables	
Investment cognitive skills	Score on 5 statements about cognitively oriented personality – see Table 4.1.
Investment social skills	Score on 5 statements about socially oriented personality – see Table 4.1.
Tasks	
Importance cognitive tasks	Number of cognitive tasks (out of 8) that are core tasks in the worker's job.
Importance social tasks	Number of social tasks (out of 8) that are core tasks in the worker's job.

Note: the definitions and measurement of the variables are displayed in Table C.1 in Appendix C.

Table 4.3. Match quality, skills, and job tasks by education group and location

	Observations	Crude		Low skilled		Medium skilled		High skilled		Countryside		City	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Matching variables													
Matching all skills	2,373	0.70	(0.20)	-0.08	(0.06)	0.15	(0.04)	0.35	(0.03)	0.13	(0.03)	0.25	(0.03)
Cognitive skills & cognitive tasks	1,596	0.93	(0.75)	-0.12	(0.05)	-0.09	(0.04)	0.14	(0.04)	-0.06	(0.03)	0.07	(0.04)
Social skills & cognitive tasks	1,596	1.00	(0.75)	0.06	(0.06)	-0.05	(0.04)	0.03	(0.04)	-0.01	(0.03)	0.01	(0.04)
Skill variables													
Investment in cognitive skills	2,149	0.23	(0.27)	-0.28	(0.04)	-0.14	(0.04)	0.34	(0.05)	-0.08	(0.03)	0.15	(0.04)
Investment in social skills	2,149	0.18	(0.23)	-0.07	(0.05)	-0.04	(0.04)	0.10	(0.04)	-0.08	(0.03)	0.10	(0.04)
Task variables													
Importance of cognitive tasks	1,724	0.33	(0.26)	-0.42	(0.05)	-0.11	(0.04)	0.30	(0.04)	-0.11	(0.03)	0.11	(0.04)
Importance of social tasks	1,724	0.48	(0.27)	-0.32	(0.06)	-0.12	(0.04)	0.27	(0.04)	-0.10	(0.04)	0.11	(0.04)

Note: the definitions and measurement of the variables are displayed in Table C.1 in Appendix C.

ments than low- and middle-skilled workers. Respondents who work in the city have more cognitively oriented personalities than respondents who work in the countryside. The same pattern is apparent for social skills. Managers have the most cognitive and social skills according to themselves, while operators attribute the least skills to themselves.

The last two rows of the table show the summary statistics of core cognitive and core social job tasks. Out of the five possible core cognitive job tasks, workers, on average, indicate that their job consists of 0.33 cognitive tasks. For core social job tasks, this average is 0.48. Similar to cognitive and social skills, high-skilled workers and workers in cities perform more cognitive and social job tasks than low- and middle-skilled workers and workers in the countryside. Again, managers state that their job contains the most cognitive and social job tasks, while elementary workers and operators indicate that their jobs contain the fewest.

As in Teulings (1995), personal characteristics correlate with job characteristics. More cognitive oriented persons perform more cognitive tasks (correlation 0.30 (0.00)) while more social oriented persons perform more social tasks (correlation 0.20 (0.00)). These correlations are stronger in the city than in the countryside. Table C.2 shows no significant correlation between the indexes for match quality. The matches of skills to tasks are better for workers with more skills and more complex jobs. The positive and significant correlation with gross monthly earnings suggests that better matches lead to higher wages.

4.3.4 Empirical model

The theoretical model results in empirical predictions about the distribution of match quality, worker skills, and job complexity across location L. We define a simple empirical strategy following equations (4.15) to (4.19) to test these predictions:

$$y_{i,l} = \alpha_0 + \alpha_1 L_l + \alpha_2 E_i + \alpha_3 Z_i, + \epsilon_{i,l}, \tag{4.20}$$

where $y_{i,l}$ is the dependent variable for worker i in location l and reflects either match quality $Q_{i,l}$, skill level $a_{i,l}$, or job complexity $\alpha_{i,l}$. The term L_l is a dummy variable indicating whether the worker works in the city or the countryside. This dummy captures the impact of the local mass of vacancies and job seekers ($min(S_{z,l}, V_{z,l})$). The worker's segment, z, is defined by the worker's educational background (E_i) and demographic characteristics (Z_i). The theoretical model as-

Table 4.4. Matching skills and tasks across occupations

	1. Managers		2. Professionals		3. Technicians		4. Clerks		5. Service and sales		7. Craft and trade		8. Operators		9. Elementary	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Matching variables																
Matching all skills	0.22	(0.91)	0.25	(0.76)	0.14	(0.84)	-0.22	(1.07)	-0.29	(1.16)	0.03	(0.97)	-0.15	(1.11)	-0.71	(1.33)
Cognitive skills & cognitive tasks	0.18	(1.16)	0.05	(1.04)	0.05	(0.98)	-0.15	(0.82)	-0.07	(0.99)	-0.12	(0.98)	-0.24	(0.87)	0.12	(1.04)
Social skills & cognitive tasks	0.09	(0.99)	0.00	(0.99)	0.01	(1.01)	-0.16	(0.93)	0.05	(0.97)	0.19	(1.25)	-0.09	(0.85)	0.08	(1.06)
Skill variables																
Investment in cognitive skills	0.46	(1.16)	0.21	(1.09)	-0.10	(0.95)	-0.15	(0.84)	-0.23	(0.90)	-0.13	(0.87)	-0.37	(0.63)	-0.25	(0.81)
Investment in social skills	0.19	(1.10)	0.06	(1.02)	0.05	(1.02)	-0.15	(0.93)	-0.01	(0.99)	-0.07	(1.00)	-0.25	(0.69)	-0.08	(0.96)
Job tasks																
Importance of cognitive tasks	0.54	(1.07)	0.18	(0.97)	0.10	(1.00)	-0.22	(0.89)	-0.34	(0.93)	-0.13	(0.85)	-0.53	(0.93)	-0.64	(0.83)
Importance of social tasks	0.48	(0.94)	0.23	(0.91)	0.08	(0.92)	-0.33	(0.92)	-0.02	(1.02)	-0.67	(1.03)	-0.56	(1.05)	-0.55	(1.03)

Note: we use one-digit ISCO occupations. The occupation group of skilled agricultural, fishery, and forestry workers (ISCO 6) is omitted since the location of these occupations depends on natural resources. The definitions and measurement of the variables are displayed in Table C.1 in Appendix C.

sumes workers search for a job within their educational class. However, within educational classes, the task packages and skills required vary strongly between different fields. To control for this heterogeneity, we also estimate a model with standard errors clustered at one-digit occupations and a model with occupational fixed effects.

Several measurement issues can affect the estimation of this empirical model. First, our dataset is a self-reporting survey, which can lead to measurement error. Measurement error affects our results when the error varies between countryside and cities. We have no reason to expect such a spatial variation. Second, the spatial sorting of workers may be driven by certain consumption preferences (Glaeser & Gottlieb, 2006). If this is the case, our estimates of matching and sorting reflect a sorting pattern of workers for consumption instead of for job opportunities. Third, the strong regional differences in the Netherlands, especially that between the Randstad area and the other regions, can result in biased results for the sample of all regions. Fourth, higher skill levels in cities may reflect a more efficient learning mechanism in cities instead of the sorting of more skilled workers into cities. In line with this reasoning, our results could show additional learning effects of city locations. Lastly, the unequal spatial distribution of industrial and service sectors may drive the results, since these sectors have different location advantages and different production structures. Section 4.5 discusses these issues in detail and provides sensitivity analyses.

4.4 Results

4.4.1 Match quality in cities

The theoretical framework argues that the quality of the match of worker skills to job tasks increases with the density of the local market. Table 4.5 presents the results of estimating the empirical model with three measures of match quality. In the first column, the suitability of a worker's skills for a job is explained by location and demographic characteristics. The coefficient of the city dummy is positive and significant: suitability is better in thick labour markets than in thin ones in the Netherlands. Furthermore, the quality of the match increases with age: young workers indicate that their skills suit their job worse than older workers. During their careers, workers self-select into jobs that match their skills better as their knowledge of both their own skills and required job tasks increase with experience. On-the-job training and learning by doing likely improve the match as well. Matches are better for men than for women and better for native workers than for non-native workers. The quality of the match increases with education level, which suggests that education is effective in terms of skill development.

Next, we cluster the standard errors at two-digit occupations (column (2)) and include fixed occupational effects (column (3)) to control for differences across education fields. Both worker skills and the task packages of jobs vary heavily between occupational groups. For instance, managers and clerks perform different tasks and therefore need different skills to perform their tasks. The coefficient of being located in a city remains significant and positive when we control for occupational differences. The match of worker skills to job tasks in cities is, on average, 14 percent of a standard deviation better than that of workers in the same occupational group in the countryside. In absolute terms, this finding is a difference of 2.8 points on a scale of zero to 100. Column (4) shows that relatively skilled workers, given their segment, experience better matches than less skilled workers in their segment.

The second measure of assignment quality considers cognitive skills and cognitive job tasks. Workers who focus on a smaller subset of job tasks and are more specialised develop more specific skills (Becker & Murphy, 1992). Furthermore, highly cognitive workers sort into specialised jobs (Bacolod et al., 2009). The higher the specialisation level of a worker and a job, the more difficulties arise with finding a decent match between the two. Among different skill types, cognitive skills seem to be an important measure for the relevance of the match. The match of cognitive skills to cognitive tasks is also significantly better in cities than in the countryside (column (5)). When we control for broad occupational groups, the city coefficient loses some significance but remains significant and positive (columns (6) and (7)).[12] Workers with abundant cognitive skills face a better match to cognitive job tasks. A worker's social skills do not affect the cognitive match.

Economic activity in cities benefits from proximity, learning, and knowledge spillovers. Considering these advantages, skills that ease or improve communication and interactions with others are especially valued in cities (Bacolod et al., 2009). Furthermore, more social, non-cognitive skills determine labour market outcomes as well (Heckman et al., 2006). Columns (9) to (12) of Table 4.5 show the estimates for the determinants of the assignment of social worker skills to social job tasks. The coefficient for working in a city is positive but insignificant. There is no significant spatial variation in match quality of social skills to social job tasks. Workers with strong social skills have better matches than workers with few social skills.

[12] We obtain more observations for the matching quality of all skills than for the ones of cognitive and

Table 4.5. Matching is better in cities

	Match quality											
	all skills				cognitive skills				social skills			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
City	0.146***	0.146***	0.136***	0.104***	0.095*	0.095*	0.093*	0.060	0.014	0.014	0.015	-0.027
	[0.040]	[0.032]	[0.029]	[0.035]	[0.050]	[0.048]	[0.048]	[0.043]	[0.051]	[0.052]	[0.051]	[0.054]
Age	0.010***	0.010***	0.007**	0.007**	-0.001	-0.001	-0.001	-0.003	-0.001	-0.001	-0.000	0.000
	[0.002]	[0.003]	[0.003]	[0.003]	[0.002]	[0.002]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.002]
Female	-0.075*	-0.075	-0.045	-0.046	-0.114**	-0.114**	-0.088	-0.006	-0.034	-0.034	0.006	-0.062
	[0.040]	[0.054]	[0.055]	[0.057]	[0.050]	[0.046]	[0.053]	[0.053]	[0.051]	[0.046]	[0.048]	[0.047]
Native	0.171**	0.171***	0.150**	0.174***	-0.099	-0.099	-0.090	-0.091	-0.169*	-0.169	-0.166	-0.115
	[0.079]	[0.061]	[0.060]	[0.060]	[0.100]	[0.117]	[0.119]	[0.109]	[0.098]	[0.107]	[0.108]	[0.092]
Medium skilled	0.151***	0.151***	0.075	0.088	0.012	0.012	-0.019	-0.052	-0.080	-0.080	-0.052	-0.062
	[0.059]	[0.055]	[0.053]	[0.056]	[0.063]	[0.061]	[0.063]	[0.058]	[0.068]	[0.059]	[0.064]	[0.059]
High skilled	0.421***	0.421***	0.211***	0.204**	0.226***	0.226***	0.151**	0.020	-0.002	-0.002	-0.013	-0.032
	[0.054]	[0.057]	[0.071]	[0.074]	[0.066]	[0.066]	[0.067]	[0.068]	[0.070]	[0.065]	[0.072]	[0.082]
Cognitive skills				0.079***				0.279***				-0.022
				[0.018]				[0.038]				[0.036]
Social skills				0.046**				-0.033				0.336***
				[0.019]				[0.025]				[0.052]
Constant	-0.761***	-0.761***	-0.537**	-0.534**	0.160	0.160	0.176	0.211	0.268	0.268	0.182	0.242
	[0.144]	[0.213]	[0.221]	[0.231]	[0.171]	[0.176]	[0.187]	[0.187]	[0.184]	[0.205]	[0.219]	[0.215]
Clustered standard errors		YES	YES	YES		YES	YES	YES		YES	YES	YES
Fixed effects			YES	YES			YES	YES			YES	YES
Observations	2,373	2,373	2,373	2,149	1,596	1,596	1,596	1,596	1,596	1,596	1,596	1,596
Adjusted R-squared	0.053	0.053	0.017	0.027	0.016	0.016	0.007	0.074	0.001	0.001	-0.001	0.101

Note: the definitions and measurement of the variables are displayed in Table C.1 in Appendix C. Robust or clustered standard errors are in parentheses. Fixed effects refer to those at the two-digit occupational level (ISCO codes). *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

4.4.2 Worker skills and job tasks in cities

The model suggests self-selection of more skilled workers and complex jobs into cities. Table 4.6 presents the results of an estimation of the spatial distribution of worker skills and job tasks. The cognitive interest of workers in cities is, on average, greater than that of workers in the countryside (column (1)). Column (2) clusters the standard errors by broad occupational groups and column (3) includes fixed effects at the occupational level. The coefficient for working in a dense urban area remains positive and significant. The spatial differences are substantial. City workers have 14 percent of a standard deviation more cognitive skills than workers in the countryside. Given their job, older workers, males, and high-skilled workers have more cognitive skills than younger workers, females, and low-skilled workers.

Columns (4) to (6) present the same estimates for workers' social skills. Workers in cities have more social skills than workers in the countryside. The spatial variation of 10 percent of a standard deviation is somewhat smaller than that for cognitive skills. Females have more social skills, while males have more cognitive skills.

Not only better workers but also better jobs are expected to sort into cities. Here, we consider the importance of several cognitive and social tasks (defined in Section 4.3.2) in job complexity. Jobs in cities demand more cognitive core job tasks than jobs in the countryside (column (7)). Bacolod et al. (2009) find no such spatial differences. Since they only measure the average task package of occupations, this finding suggests spatial variation within the content of jobs. Columns (8) and (9) indeed show significant spatial variation in task packages within broad occupations. Occupations in cities contain 9 percent of a standard deviation more cognitive job tasks than comparable occupations in the countryside. Column (10) shows that workers in cities perform more social job tasks than workers in the countryside. This spatial variation is, however, fully explained by the spatial distribution of jobs (column (11)).

social skills. The results for the matching quality of all skills are similar in both samples of observations.

Table 4.6. Spatial distribution skills and tasks

	Cognitive skills			Social skills			Cognitive tasks			Social tasks		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
City	0.159***	0.159***	0.137***	0.116***	0.116***	0.101***	0.127***	0.127**	0.090*	0.122**	0.122*	0.062
	[0.042]	[0.044]	[0.046]	[0.044]	[0.039]	[0.036]	[0.047]	[0.052]	[0.050]	[0.048]	[0.060]	[0.056]
Age	0.007***	0.007***	0.006***	-0.001	-0.001	-0.002	-0.001	-0.001	-0.004**	0.003	0.003*	0.000
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Female	-0.296***	-0.296***	-0.242***	0.174***	0.174***	0.177***	-0.306***	-0.306***	-0.246***	0.024	0.024	-0.098**
	[0.041]	[0.045]	[0.042]	[0.043]	[0.050]	[0.053]	[0.046]	[0.051]	[0.047]	[0.047]	[0.085]	[0.045]
Native	-0.007	-0.007	-0.020	-0.161**	-0.161*	-0.160*	0.017	0.017	-0.010	0.108	0.108	0.059
	[0.067]	[0.075]	[0.075]	[0.080]	[0.090]	[0.088]	[0.086]	[0.072]	[0.069]	[0.084]	[0.072]	[0.063]
Medium skilled	0.199***	0.199***	0.161***	-0.003	-0.003	-0.016	0.283***	0.283***	0.124**	0.205***	0.205**	0.070
	[0.046]	[0.053]	[0.049]	[0.055]	[0.060]	[0.055]	[0.062]	[0.057]	[0.049]	[0.069]	[0.075]	[0.057]
High skilled	0.625***	0.625***	0.508***	0.147***	0.147**	0.076	0.639***	0.639***	0.362***	0.548***	0.548***	0.232***
	[0.051]	[0.073]	[0.069]	[0.056]	[0.056]	[0.053]	[0.062]	[0.077]	[0.063]	[0.067]	[0.101]	[0.082]
Constant	-0.229*	-0.229**	-0.195*	-0.196	-0.196	-0.106	0.051	0.051	0.337**	-0.637***	-0.637***	-0.071
	[0.128]	[0.101]	[0.107]	[0.144]	[0.136]	[0.127]	[0.156]	[0.182]	[0.161]	[0.167]	[0.193]	[0.118]
Clustered standard errors		YES	YES		YES	YES		YES	YES		YES	YES
Fixed effects			YES			YES			YES			YES
Observations	2,149	2,149	2,149	2,149	2,149	2,149	1,724	1,724	1,724	1,724	1,724	1,724
Adjusted R-squared	0.106	0.106	0.054	0.016	0.016	0.010	0.093	0.093	0.035	0.053	0.053	0.09

Note: cognitive (social) skills refer to the number of cognitive (social) statements the respondent agrees with. Cognitive (social) tasks refer to the number of cognitive (social) tasks the worker performs. The definitions and measurement of the variables are displayed in Table C.1 in Appendix C. Robust or clustered standard errors are in parentheses. Fixed effects refer to those at the two-digit occupational level (ISCO codes). *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

4.5 Further analyses

Previous estimates may be affected by several estimation issues, as discussed in Section 4.3.4. To test the sensitivity of the results to these issues, this section presents several additional analyses. First, in Section 4.5.1 we test the impact of measurement error caused by the self-reporting of the main variables. Next, in Section 4.5.2 we test whether our results reflect the sorting of workers for job opportunities or for consumption preferences. Regional differences in local labour markets are analysed in Section 4.5.3. Fourth, Section 4.5.4 discusses the role of quicker human capital accumulation in cities. Section 4.5.5 presents separate analyses for industrial and service occupations. Lastly, Section 4.5.6 presents a first indicator for the relevance of match quality in urban wage premia. Here, we only present the results for one measure of match quality, namely, the suitability of all skills. The quality of the cognitive match shows similar patterns, with less significant spatial variation. Social match quality never shows significant spatial variation.

4.5.1 Subjective measurement

The dataset consists of self-reported personalities, self-reported job tasks, and self-reported quality of job matches. Autor & Handel (forthcoming) discuss several issues with these kind of surveys. Bias caused by the respondents' subjective answers is our main concern. First, bias can result from the abstract definitions of the variables, resulting in different interpretations among respondents. Second, respondents likely vary in how they distribute scores. For instance, some respondents will label a score as important, whereas others will label the same score as very important. This measurement error affects our results when workers in cities have different biases in their answers than workers in the countryside.

The survey includes questions about task importance and the effectiveness of several tasks for commonly known example jobs. All respondents should have an image of the task package and required skills of these well-known jobs, such as secretary or teacher. The questions measure the respondent's answering bias. The idea is that when a respondent interprets a certain task differently or provides higher scores than others, he or she will do so for the example job as well. We use the relative answers of respondents to proxy for answering bias in match, skill, and task questions. The relative answers about the effectiveness of workers in certain tasks in the example jobs proxy for measurement error in skills, while the relative answers about the importance of tasks proxy for errors in relevance of job tasks.

Lastly, the error in a match is proxied for by the difference between relative importance and relative effectiveness. Appendix C.2 displays the details of the measurement.

The average value of all three proxies is significantly higher in cities than in the countryside. Workers who work in cities attribute relatively more importance, more effectiveness, and better match quality between importance and effectiveness to the job tasks of example jobs. This spatial variation could be driven by different worker attitudes in cities and the spatial sorting of workers. The value of all three proxies also varies significantly across education groups. We test the impact of this spatial variation in measurement error in two steps. First, we test whether the spatial variation remains significant when we control for other characteristics, such as education and gender. Second, we include the proxy in our baseline empirical model to see whether the results change when we control for measurement error.

The first three columns of Table 4.7 show the spatial variation in these three proxies, controlling for the usual factors. Only the spatial variation of attributing effectiveness to a task in an example job remains significant when we control for other characteristics. This finding suggests that measurement error could affect our measure of the sorting of skilled workers in cities, but probably not that of match quality. Columns (4) to (8) present our previous estimates, including the proxy. The proxy for the measurement error has an insignificant coefficient in the matching estimation and a significant coefficient in the skill and task estimations. Respondents who value the importance and effectiveness of job tasks in the example job more have higher skill levels and jobs with more demanding tasks. The proxy for the error of matching is defined by the difference between importance and effectiveness; the insignificant coefficient suggests that the bias in the two cancels out. None of the previous results is affected by the inclusion of the proxy. City workers have a significantly positive bias to their answers compared to countryside workers. This bias does not, however, affect our results.

4.5.2 Consumption preferences

Urban areas facilitate interactions not only between workers and employers, but also between the workers themselves. Many people like to live in urban areas for social interaction and a larger variety of consumption amenities, from schools to theatres (Glaeser & Gottlieb, 2006; Glaeser et al., 2001). Urban consumption variety is deemed a luxury good (Lee, 2010). Thus, richer people tend to value urban consumption variety more than poor people. The relation between skills and wages

Table 4.7. Subjective measurement

	Measurement error proxy			Matching	Skills		Tasks	
	Imp-eff	Effectiveness	Importance		Cognitive	Social	Cognitive	Social
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
City	-0.003	0.082	0.079	0.105***	0.113**	0.125***	0.083	0.053
	[0.036]	[0.066]	[0.050]	[0.035]	[0.052]	[0.036]	[0.052]	[0.056]
Age	0.001	-0.001	0.000	0.007**	0.006**	-0.002	-0.004**	0.000
	[0.002]	[0.003]	[0.002]	[0.003]	[0.003]	[0.002]	[0.002]	[0.002]
Female	0.097**	0.011	0.175**	-0.047	-0.263***	0.185***	-0.277***	-0.130***
	[0.041]	[0.044]	[0.065]	[0.057]	[0.064]	[0.062]	[0.049]	[0.043]
Native	0.037	0.026	0.084	0.174***	-0.035	-0.155*	-0.023	0.062
	[0.068]	[0.058]	[0.112]	[0.060]	[0.081]	[0.090]	[0.070]	[0.061]
Medium skilled	0.015	0.042	0.071	0.088	0.122**	0.039	0.122**	0.067
	[0.045]	[0.060]	[0.060]	[0.056]	[0.050]	[0.065]	[0.047]	[0.058]
High skilled	-0.027	-0.104	-0.142*	0.204**	0.493***	0.080	0.372***	0.244***
	[0.060]	[0.070]	[0.072]	[0.074]	[0.083]	[0.075]	[0.065]	[0.084]
Cognitive skills	0.015	0.097***	0.123***	0.079***				
	[0.028]	[0.023]	[0.037]	[0.018]				
Social skills	-0.003	0.053*	0.042	0.046**				
	[0.019]	[0.028]	[0.032]	[0.019]				
Measurement error proxy				0.011	0.105***	0.078***	0.130***	0.133***
				[0.019]	[0.023]	[0.027]	[0.022]	[0.019]
Constant	-0.229*	-0.006	-0.357*	-0.532**	-0.133	-0.155	0.379**	-0.024
	[0.129]	[0.173]	[0.182]	[0.229]	[0.114]	[0.134]	[0.156]	[0.121]
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES	YES
Fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,149	1,556	1,556	2,149	1,556	1,556	1,683	1,683
Adjusted R-squared	-0.001	0.013	0.024	0.027	0.061	0.016	0.054	0.029

Note: the measurement error proxy is defined for the difference between the importance and effectiveness of tasks (label 'Imp-eff'), the effectiveness of task performance (label 'Effectiveness') and the importance of tasks (label 'Importance'). Appendix C.2 displays the details of the measurement. Cognitive (social) skills refer to the number of cognitive (social) statements the respondent agrees with. The dependent variable 'matching' measures the matching quality of all skills. Cognitive (social) tasks refer to the number of cognitive (social) tasks the worker performs. The definitions and measurement of the variables are displayed in Table C.1 in Appendix C. The Appendix displays the measurement of the proxy as well. Robust or clustered standard errors are in parentheses. Fixed effects refer to those at the two-digit occupational level (ISCO codes). *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

suggests that more skilled people value urban consumption variety more and are more likely to locate in an urban area.

Our estimates of the spatial distribution of skills could reflect the sorting of more skilled workers into cities for consumption preferences instead of for job opportunities. The Netherlands is an interesting case to test whether jobs follow people or people follow jobs. A substantial part of the Dutch labour force (more than 50 percent; see Statistics Netherlands) does not work in the same municipality as they live. Distances are short in the Netherlands and commuting to work is very common. On average, a Dutchman travels 17 kilometres to work. Because many people choose to commute to work in the Netherlands, we can test whether the location of residence or the location of work reflects the matching and sorting patterns we find. In the sample, 27 percent of the individuals do not work and live in a location with the same density; 57 percent of these workers live in the countryside and work in the city, while the other 43 percent live in the city and work in the countryside.

Table 4.8 presents the results of an estimation including a city dummy for the worker's location of residence instead of the location of work. The location of residence does not explain variation within the match of all worker skills to job tasks. Worker skills do vary with the density of the location of residence (see columns (2) and (3)). Workers who live in the city have more cognitive and social skills than workers who live in the countryside. The importance of cognitive and social job tasks does not vary with the density of the location of residence (columns (4) and (5)). Lastly, column (6) explains the quality of the match for a sample of commuters. The coefficient for the city of residence is negative and significant. Workers who commute from a large city of residence to the countryside for work have a worse match than workers who commute the other way.

The complexity of jobs and match quality only increase with the density of the worker's work location and not with the density of the location of residence. This underlines our hypothesis that the density of workers and jobs in cities results in better matching between the two. Our findings suggest that more skilled workers sort into cities of residence for consumption preferences or other reasons, such as the partner's location of work, while their location of work depends on job opportunities.

4.5.3 Regional differences in the Netherlands

Both the theoretical and empirical models neglect a city's hinterland. Cities are assumed to be isolated in space. In the case of the Netherlands, the hinterland across

Table 4.8. City of residence

	Matching	Skills		Tasks		Commuters
		Cognitive	Social	Cognitive	Social	Matching
	(1)	(2)	(3)	(4)	(5)	(6)
City of residence	0.024	0.139***	0.185***	-0.002	0.041	-0.130*
	[0.036]	[0.047]	[0.047]	[0.047]	[0.053]	[0.065]
Age	0.007**	0.006***	-0.002	-0.004**	0.000	0.001
	[0.003]	[0.002]	[0.002]	[0.002]	[0.002]	[0.004]
Female	-0.049	-0.245***	0.175***	-0.250***	-0.099**	-0.133
	[0.057]	[0.043]	[0.053]	[0.048]	[0.045]	[0.086]
Native	0.167***	-0.006	-0.133	-0.020	0.061	0.104
	[0.059]	[0.076]	[0.089]	[0.069]	[0.059]	[0.126]
Medium skilled	0.093	0.174***	-0.002	0.124**	0.073	-0.051
	[0.055]	[0.046]	[0.054]	[0.049]	[0.058]	[0.119]
High skilled	0.215***	0.524***	0.087	0.371***	0.239***	0.049
	[0.074]	[0.067]	[0.054]	[0.064]	[0.082]	[0.142]
Cognitive skills	0.081***					
	[0.018]					
Social skills	0.046**					
	[0.019]					
Constant	-0.494**	-0.215*	-0.176	0.386**	-0.067	0.187
	[0.224]	[0.110]	[0.133]	[0.157]	[0.108]	[0.278]
Clustered standard errors	YES	YES	YES	YES	YES	YES
Fixed effects	YES	YES	YES	YES	YES	YES
Observations	2,149	2,149	2,149	1,724	1,724	615
Adjusted R-squared	0.025	0.055	0.016	0.033	0.008	0.005

Note: the dependent variable 'matching' measures the matching quality of all skills. Cognitive (social) tasks refer to the number of cognitive (social) tasks the worker performs. Commuters work in a city with a different density than the city they live in. The definitions and measurement of the variables are displayed in Table C.1 in Appendix C. Clustered standard errors are in parentheses. Fixed effects refer to those at the two-digit occupational level (ISCO codes). *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

regions differs substantially. In the Randstad provinces, the distance between large cities is much smaller than in other provinces. Several studies therefore refer to the Randstad provinces as one city (Lambooy, 1998). If the Randstad operates as a single labour market, workers and employers search for matches within the Randstad. This suggests that the importance of a city's density should be more important outside the Randstad than within it.

Table 4.9 shows separate estimations for workers located in and outside the Randstad. The matching of skills to job tasks is better in cities than in the countryside in both regions. If we control for self-selection into occupational groups, this spatial variation remains significant only outside the Randstad. Workers who work in cities have more cognitive skills than workers in the countryside in both regions. In the Randstad, city workers also have more social skills. The complexity of jobs does not vary across cities or the countryside in the Randstad, whereas it does outside the Randstad.

The results in Table 4.9 suggest that the more integrated labour market in the Randstad diminishes the spatial variation in match quality. Since the Randstad is often seen as one labour market, we assume that the Randstad operates more efficiently in the matching of workers to jobs because a relatively large market is created. The variation in the scope of the labour market likely affects the optimal spatial unit of observation in the Netherlands. Analyses for an alternative spatial unit may bias the results. This so-called modifiable area unit problem (MAUP) seems to bias our results for the Randstad area (see Briant et al. (2008) for a discussion on the MAUP).

4.5.4 Human capital accumulation

Many studies suggest that cities stimulate knowledge spillovers and learning (Jaffe et al., 1993; Rosenthal & Strange, 2008; Glaeser & Ressenger, 2010). The quicker and deeper human capital accumulation of workers in cities may be the driving force behind the higher productivity rates in these cities. Glaeser & Maré (2001), for instance, show that workers start earning an urban wage premium three to five years after their move to the city.

Considering our estimates, the quicker and better human capital accumulation in cities could result in a quicker development of workers' skills with respect to their tasks in cities. If workers in cities learn more and faster than workers in the countryside, their skill development towards job tasks will be better and faster as well. The results for the spatial distribution of worker skills could reflect a learning

Table 4.9. Regional differences in the Netherlands

	Matching		Skills				Tasks			
			Cognitive		Social		Cognitive		Social	
Region	Randstad	Other region	Randstad	Other region	Randstad	Other region	Randstad	Other region	Randstad	Other region
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
City	0.104	0.142**	0.157**	0.121**	0.174***	0.012	0.017	0.133**	0.026	0.071
	[0.093]	[0.056]	[0.076]	[0.058]	[0.060]	[0.055]	[0.099]	[0.055]	[0.100]	[0.068]
Age	0.009**	0.006	0.005	0.006**	-0.003	-0.001	-0.003	-0.005*	0.003	-0.001
	[0.004]	[0.004]	[0.003]	[0.002]	[0.003]	[0.002]	[0.003]	[0.003]	[0.004]	[0.003]
Female	-0.000	-0.065	-0.310***	-0.210***	0.091	0.229***	-0.246***	-0.246***	-0.111	-0.098
	[0.073]	[0.073]	[0.061]	[0.046]	[0.077]	[0.056]	[0.068]	[0.074]	[0.086]	[0.072]
Native	0.117	0.197***	-0.007	-0.007	-0.067	-0.200**	0.021	-0.090	0.071	-0.004
	[0.080]	[0.071]	[0.093]	[0.077]	[0.131]	[0.092]	[0.089]	[0.109]	[0.065]	[0.107]
Medium skilled	0.281**	-0.015	0.143*	0.153**	-0.121	0.008	0.209**	0.075	-0.014	0.097
	[0.114]	[0.069]	[0.081]	[0.062]	[0.108]	[0.062]	[0.078]	[0.067]	[0.108]	[0.060]
High skilled	0.262*	0.185***	0.515***	0.497***	-0.112	0.168**	0.328***	0.399***	0.025	0.327***
	[0.138]	[0.066]	[0.082]	[0.104]	[0.101]	[0.069]	[0.106]	[0.090]	[0.146]	[0.089]
Cognitive skills	0.038	0.113***								
	[0.031]	[0.031]								
Social skills	0.104***	0.007								
	[0.030]	[0.027]								
Constant	-0.762***	-0.444	-0.065	-0.279*	0.125	-0.259	0.312*	0.418	0.037	-0.028
	[0.253]	[0.295]	[0.191]	[0.152]	[0.248]	[0.186]	[0.161]	[0.253]	[0.223]	[0.252]
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	822	1327	822	1327	822	1327	665	1,059	665	1,059
Adjusted R-squared	0.036	0.032	0.066	0.045	0.003	0.014	0.018	0.047	-0.004	0.015

Note: the dependent variable 'matching' measures the matching quality of all skills. Cognitive (social) tasks refer to the number of cognitive (social) tasks the worker performs. The definitions and measurement of the variables are displayed in Table C.1 in Appendix C. The Randstad sample contains all workers who work in the following provinces: Noord-Holland, Zuid-Holland, and Utrecht. Clustered standard errors are in parentheses. Fixed effects refer to those at the two-digit occupational level (ISCO codes). *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

mechanism if these additional skills reflect newly learned skills instead of initial skills a worker had before the job match.

Table 4.10 shows estimates that test this hypothesis. The density of the work location does not explain the development of the job match between 2010 and 2012. City workers also do not learn more cognitive and social tasks at work (columns (2) and (3)) than workers in the countryside. Moreover, workers in dense cities learned their cognitive skills more often at school than workers in the countryside.

Table 4.10. Learning in cities

	Matching 2010–2012	Tasks learned at work Cognitive	Social
	(1)	(2)	(3)
City	-1.022	-0.038**	-0.011
	[1.158]	[0.018]	[0.020]
Age	-0.276***	0.006***	0.007***
	[0.051]	[0.001]	[0.001]
Female	1.228	0.018	0.007
	[1.219]	[0.025]	[0.015]
Native	-0.686	0.023	0.018
	[2.677]	[0.031]	[0.024]
Medium skilled	-7.475***	-0.058**	-0.034
	[2.124]	[0.022]	[0.023]
High skilled	-9.673***	-0.123***	-0.110***
	[1.795]	[0.017]	[0.024]
Cognitive skills	-0.940	-0.024**	-0.009
	[0.585]	[0.012]	[0.007]
Social skills	0.575	0.013	0.003
	[0.929]	[0.008]	[0.006]
Constant	-54.573***	0.454***	0.504***
	[3.684]	[0.083]	[0.062]
Clustered standard errors	YES	YES	YES
Fixed effects	YES	YES	YES
Observations	1,567	1,501	1,496
Adjusted R-squared	0.023	0.077	0.088

Note: the dependent variable 'matching' measures the matching quality of all skills. Tasks learned at work is a dummy variable indicating whether cognitive (social) tasks are learned at work or not. The definitions and measurement of the variables are displayed in Table C.1 in Appendix C. Clustered standard errors are in parentheses. Fixed effects refer to those at the two-digit occupational level (ISCO codes). *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

4.5.5 Industrial and service jobs

Location advantages vary across several stages and tasks of the production process. For instance, a metalworker performing routine tasks in a factory and an innovator working for the same industry but in the research and development department face different advantages of dense areas. Different locations are beneficial for different stages of product life cycle (Harrison et al., 1996), firm life cycle (Duranton & Puga, 2001), and industry life cycle (Desmet & Rossi-Hansberg, 2009). This results in an unequal distribution of these production stages over space. Our theoretical model, however, suggests that the amount of vacancies and job seekers is larger in cities for all workers and employers. If, for instance, manufacturing jobs are overrepresented in the countryside, this would result in scale benefits in the countryside for these jobs instead of in the city. More land-intense and less knowledge-intense product processes result in less agglomeration economies for goods production than for idea production (Glaeser & Ponzetto, 2010). Therefore, we distinguish between industrial and service occupations. Industrial occupations focus on producing goods, while service occupations focus on either producing ideas or providing services. Indeed, 47 percent of the service occupations are performed in the city, while only 31 percent of industrial occupations are.

Table 4.11 presents separate estimations for both occupation types. The coefficient for city work location is positive and insignificant for industrial occupations and positive and very significant for service occupations. Both the weaker spatial distribution of industrial occupations and the smaller number of observations can explain the insignificant coefficient for industrial occupations. Columns (3) to (6) show the distribution of worker skills for both occupation types. Again, only the spatial distribution of the service sector is significant. We do not find a significant spatial distribution for the importance of cognitive and social tasks for either type of occupation.

4.5.6 Explaining regional wage differences

Our results show that the matching of worker skills to job tasks is of better quality in the cities than in the countryside. Here, we test whether this better match quality determines part of the urban wage premium in the Netherlands. Clearly, a full assessment of the role of matching in urban wage premia is beyond the scope of this chapter and not feasible with our dataset. This section presents a simple back-of-the-envelope estimation and suggests that more productive labour matches in

Table 4.11. Industrial and service occupations

	Matching		Skills				Tasks			
			cognitive		social		cognitive		social	
	Industrial	Service	Industrial	Service	Industrial	Service	Industrial	Service	Industrial	Service
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
City	0.166	0.094**	-0.033	0.164***	0.009	0.122***	0.198	0.070	0.313	0.016
	[0.124]	[0.035]	[0.102]	[0.049]	[0.185]	[0.028]	[0.216]	[0.047]	[0.186]	[0.055]
Age	0.000	0.007**	0.013***	0.005**	-0.007*	-0.001	0.001	-0.005**	0.001	-0.000
	[0.004]	[0.004]	[0.004]	[0.002]	[0.003]	[0.002]	[0.003]	[0.002]	[0.006]	[0.002]
Female	-0.401*	-0.003	-0.169	-0.249***	0.358*	0.158**	-0.443**	-0.232***	-0.069	-0.109**
	[0.218]	[0.054]	[0.173]	[0.041]	[0.173]	[0.058]	[0.175]	[0.051]	[0.200]	[0.044]
Native	0.106	0.197***	-0.298	0.030	-0.083	-0.182*	0.077	-0.033	0.240	0.017
	[0.185]	[0.062]	[0.176]	[0.083]	[0.154]	[0.102]	[0.179]	[0.072]	[0.178]	[0.063]
Medium skilled	0.041	0.097	0.350***	0.117**	0.011	-0.029	0.249**	0.089	0.264***	0.002
	[0.129]	[0.058]	[0.108]	[0.050]	[0.096]	[0.063]	[0.103]	[0.055]	[0.071]	[0.075]
High skilled	0.284	0.204**	0.563***	0.484***	-0.103	0.090	0.450**	0.338***	0.520***	0.158
	[0.180]	[0.079]	[0.160]	[0.074]	[0.108]	[0.059]	[0.204]	[0.070]	[0.124]	[0.100]
Cognitive skills	0.132***	0.071***								
	[0.031]	[0.019]								
Social skills	-0.003	0.055**								
	[0.048]	[0.020]								
Constant	0.193	-0.660**	-0.433	-0.149	-0.127	-0.113	0.046	0.438**	-0.865***	0.158
	[0.389]	[0.255]	[0.379]	[0.094]	[0.390]	[0.131]	[0.338]	[0.181]	[0.263]	[0.114]
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	363	1,786	363	1,786	363	1,786	287	1,437	287	1,437
Adjusted R-squared	0.031	0.028	0.049	0.058	0.016	0.010	0.033	0.035	0.039	0.005

Note: the dependent variable 'matching' measures the matching quality of all skills. Cognitive (social) tasks refer to the number of cognitive (social) tasks the worker performs. The definitions and measurement of the variables are displayed in Table C.1 in Appendix C. Industrial occupations have ISCO codes 13, 31, 61, 62, 69, and 70–92. Clustered standard errors are in parentheses. Fixed effects refer to those at the two-digit occupational level (ISCO codes). *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

the cities result in higher wages.

Table 4.12 presents the results of a simple wage model. Workers in Dutch cities earn, ceteris paribus, 13 percent of a standard deviation more than workers in the countryside (column (1)). Column (2) shows that a one standard deviation better job match increases wages by 23 percent of a standard deviation. When we include both variables, both coefficients decrease slightly (column (3)).

Column (4) includes a worker's cognitive and social attitudes as additional skill information. These hardly affect the wage returns of the match quality. The coefficient of the urban wage premium does decrease slightly. Cognitive skills are valued positively, while social skills have no wage returns. Borghans et al. (2008) argue that the supply and demand of skills determine their wage returns. These authors' results resemble ours and suggest that social skills are overrepresented relative to cognitive skills.

Additionally, we follow the task approach literature and proxy for worker skills with job tasks (see Acemoglu & Autor (2011) for a review of this literature). This approach assumes that job tasks reflect work activities that produce output. The ongoing self-selection of workers into job tasks implies an interplay between workers skills and job tasks (Autor & Handel, forthcoming). Columns (5) to (10) include information about a worker's job tasks and the job's broad occupational group. The performance of both cognitive and social job tasks is valued positively. A substantial part of the urban wage premium is explained by different job tasks: wage returns decrease by 8 to 12 percent of a standard deviation. In addition, the coefficient of match quality decreases substantially, from 0.22 to 0.10. Column (6) includes both skills and tasks and shows that the latter are especially valued.

Lastly, columns (7) to (10) show fixed effects regressions explaining variation within broad occupational groups. The urban wage premia and the returns to match quality decrease when we include occupational fixed effects. Hence, the spatial distribution of occupations explains a substantial part of the spatial wage differences in the Netherlands. Columns (9) and (10) show substantial explanatory power of job tasks. As we control for additional cognitive and social job tasks, the spatial wage variation in the Netherlands becomes insignificant. This finding suggests that the spatial wage variation reflects different activities and not increasing returns to scale. Only the economic activity of workers explains spatial wage differences. The wage return of match quality remains significant but decreases to 9 percent of a standard deviation.

Table 4.12. Wage returns

	Gross monthly earnings									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
City	0.134***		0.120***	0.129***	0.080**	0.090**	0.092***	0.099***	0.055	0.064*
	[0.041]		[0.040]	[0.042]	[0.039]	[0.040]	[0.032]	[0.030]	[0.034]	[0.036]
Matching		0.226***	0.224***	0.221***	0.098***	0.088***	0.188***	0.195***	0.082**	0.079**
		[0.030]	[0.030]	[0.033]	[0.027]	[0.028]	[0.031]	[0.036]	[0.030]	[0.032]
Age	0.111***	0.096***	0.095***	0.094***	0.070***	0.071***	0.073***	0.072***	0.061***	0.061***
	[0.015]	[0.014]	[0.014]	[0.014]	[0.014]	[0.015]	[0.013]	[0.014]	[0.016]	[0.016]
Age squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Female	-0.845***	-0.836***	-0.835***	-0.829***	-0.720***	-0.714***	-0.736***	-0.730***	-0.651***	-0.637***
	[0.041]	[0.040]	[0.040]	[0.043]	[0.040]	[0.042]	[0.067]	[0.078]	[0.058]	[0.066]
Native	0.075	-0.017	0.002	0.000	0.033	0.014	-0.029	-0.024	-0.002	-0.021
	[0.073]	[0.067]	[0.068]	[0.071]	[0.062]	[0.065]	[0.084]	[0.080]	[0.068]	[0.072]
Medium skilled	0.208***	0.187***	0.177***	0.138**	0.103*	0.096*	0.059	0.030	-0.002	-0.002
	[0.059]	[0.058]	[0.058]	[0.059]	[0.057]	[0.057]	[0.059]	[0.065]	[0.059]	[0.066]
High skilled	0.843***	0.786***	0.761***	0.724***	0.578***	0.586***	0.531***	0.500***	0.396***	0.397***
	[0.055]	[0.055]	[0.055]	[0.057]	[0.059]	[0.059]	[0.064]	[0.064]	[0.060]	[0.065]
Cognitive skills				0.045**		0.002		0.027*		-0.004
				[0.022]		[0.022]		[0.015]		[0.018]
Social skills				-0.017		-0.024		-0.024		-0.034
				[0.023]		[0.022]		[0.024]		[0.022]
Cognitive tasks					0.144***	0.150***			0.099***	0.103***
					[0.024]	[0.025]			[0.027]	[0.024]
Social tasks					0.068***	0.065***			0.068**	0.070*
					[0.022]	[0.023]			[0.031]	[0.034]
Constant	-2.052***	-1.575***	-1.621***	-1.586***	-1.101***	-1.115***	-1.029***	-0.996***	-0.773**	-0.765*
	[0.362]	[0.321]	[0.325]	[0.341]	[0.331]	[0.344]	[0.320]	[0.361]	[0.372]	[0.400]
Clustered standard errors							YES	YES	YES	YES
Fixed effects							YES	YES	YES	YES
Observations	1,490	1,490	1,490	1,354	1,157	1,071	1,490	1,354	1,157	1,071
Adjusted R-squared	0.393	0.427	0.430	0.435	0.443	0.450	0.308	0.307	0.325	0.325

Note: the definitions and measurement of the variables are displayed in Table C.1 in Appendix C. We indicate gross monthly wages as missing when a person earns nothing, less than nothing, or more than 10,000 euros a month. Robust or clustered standard errors are in parentheses. Fixed effects refer to those at the two-digit occupational level (ISCO codes). *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

4.6 Conclusion

This chapter estimates the spatial variation in the match quality of worker skills to job tasks in the Netherlands. We argue that the assignment of heterogeneous workers to heterogeneous jobs is better in a larger market. Tighter matches attract relatively skilled workers and relatively complex jobs to these large markets to optimise returns to their skills and complexity. Within the debate about the sorting of skilled workers, we show that workers indeed sort into cities for better matching opportunities. This pattern is apparent in the spatial distribution of occupations as well. The better matching of worker skills to job tasks results in higher individual wages for workers with the same occupation but with a better match.

This chapter contributes to the literature about agglomeration economies by measuring labour market pooling directly, see the work of Rosenthal & Strange (2004) and Glaeser & Gottlieb (2009) for reviews. Earlier work of, among others, Helsley & Strange (1990), Kim (1990), and Wheeler (2001) frames the idea of labour market pooling and the impact of scale on match quality. Extending the work of Petrongolo & Pissarides (2006), we measure the quality of the match between skills and tasks. The finding that the match quality of skills to jobs is better in cities indicates the labour market advantages of economic concentration.

CHAPTER 5

RETURNS TO COMMUNICATION

IN SPECIALISED AND DIVERSIFIED US CITIES[*]

5.1 Introduction

A key factor in today's urban wealth is the means by which cities reduce costs of communication. Rapid progress in transport, information and communication technologies lowered the costs of production at distance. Still, in 2009 metropolitan areas were responsible for 85 percent of US employment, income and production. The significance of personal communication for innovation is a fundamental aspect of the current economic success of cities. The economic structure of cities varies; diversified cities focusing on producing ideas and specialised cities focusing on producing products successfully coexist in the US. Is communication equally important and valued within both city types?

Variation in the advantages of clustering of economic activity resulted in the existence of different economic city structures. Typically two types of cities coexist in the US: cities with a specialised industrial structure and cities with a diversified industrial structure (Duranton & Puga, 2000). Within specialised cities firms benefit from cost sharing, labour matching and learning from similar firms. The production costs are relatively low in these cities and the focus lies on producing products. A diversified environment with a wide variety of firms and ideas is beneficial for innovation and producing ideas. The knowledge spillovers are more extensive in diversified cities but the production costs are higher. Especially for young firms and products the flows of ideas within diversified cities are key to success, while more mature firms flourish in specialised cities (Duranton & Puga, 2001; Desmet & Rossi-Hansberg, 2009). These variations in trade-offs between knowledge spillovers and

[*] This chapter is based on Kok (forthcoming)

production costs suggest that communication is less important within specialised cities. However, this suggestion does not reconcile with the assigned role of knowledge spillovers to the success of specialised clusters such as Silicon Valley.

In this chapter we focus on the role of communication within the coexistence of diversified and specialised cities. We measure the individual returns to communication job tasks in a cross-section of both city types in the US. Workers, who communicate more in and outside the organisation, earn higher wages. The main contribution to regional science and policy is our finding that the importance of communication decreases with the specialisation level of cities.

First, a simple framework is set out to guide our empirical analyses. The framework captures an economy with perfect competition, free firm entry, full mobility of labour and spatial wage differences. The differences in wages across local labour markets are compensated with differences in productivity, labour ability and other local characteristics. In equilibrium both firms and workers are indifferent towards location. The productivity of a firm increases with the specialisation level of the city when the firm operates in the dominant industry of the city, hence the industry in which the city specialises. The productivity benefits of local communication decrease with the specialisation level of the city.

Second, we estimate the returns to communication job tasks for workers in the largest 168 US cities in 2009. Individual data from the Current Population Survey (CPS) is combined with the job characteristics from the Occupational Information Network (ONET) Skill Survey. The performance of communication job tasks is defined by the work context and work activities information from the ONET Skill Survey. We start by estimating simple wage regressions in which we test the correlation between communication job tasks and individual wage, conditional on several individual and city characteristics. We find a positive relation between the number of communication job tasks a worker performs and his wage. Furthermore, our estimates show that this relation is present in both specialised and diversified cities but diminishes with the specialisation level of the city. The correlation between wage and communication is significantly stronger in diversified cities than in specialised cities.

Third, we control for differences in unobserved ability and perform IV-estimates. The occupational communication job tasks are instrumented with a language-skill proxy. Workers with weaker language-skills are assumed to be less likely to perform communication job tasks. The language-skill proxy measures the share of workers in an occupation who did not grow up in an English-speaking household.

Several tests prove that the language-skill proxy does not measure the wage impact of cultural differences. Following Ciccone & Hall (1996) historical population (1930) is used as an instrument for current city size or the extent to which the industrial structure is either specialised or diversified. The IV-estimates correspond to the OLS-estimates. A one standard deviation increase in the importance of communication, increases wages by 18 percent of a standard deviation. However, in cities with a specialised sectoral structure, these returns are about 16 percent of a standard deviation. The returns are somewhat higher in large cities: about 21 percent of a standard deviation. The returns to communication do not vary with the diversity level of the city. The variation in returns to communication over city types explains part of the lower wages in specialised cities and part of the higher wages in larger cities.

Lastly, we carry out several robustness checks and analyse alternative specifications. First, we test the sensitivity of the measure of communication and measure the returns to the relative importance of communication, non-routine interactive tasks as in Autor et al. (2003) and people skills as in Bacolod et al. (2009). Next, we perform an additional test on the effect of unobserved ability and allow the returns to communication to vary across skill level (Glaeser & Maré, 2001). The results are robust to all these specifications. Moreover, the results hold for both industrial sectors and service sectors.

Our work is based on a small theoretical literature explaining the coexistence of diversified and specialised cities. Duranton & Puga (2001) and Desmet & Rossi-Hansberg (2009) set up a dynamic general-equilibrium model that explains the coexistence of the two city types within the life-cycle of respectively firms and industries. Glaeser & Ponzetto (2010), Gaspar & Glaeser (1998) and Ioannides et al. (2008) model two rival spatial effects of technological progress. All these papers underline their theory with empirical analyses. Furthermore, Harrison et al. (1996), Kelley & Helper (1999) and Feldman & Audretsch (1999) document the contributions of sectoral diversity towards new production processes and new products.

A very broad and extensive literature indicates the (non random) coexistence of diversified and specialised cities (Duranton & Puga, 2000; Ellison & Glaeser, 1999) and the relative advantages at the city level (see Glaeser & Gottlieb (2009) for an overview). The importance of communication in the current wealth of cities relates to empirical contributions of (among others) Jaffe et al. (1993), Rauch (1993), Charlot & Duranton (2004), Bacolod et al. (2009) and Florida et al. (2012). Our work adds to these contributions by focussing on the variation in returns to communication between different city types. Therefore, we focus on the suggested microfoundations of the coexistence of these two city types as in Duranton & Puga (2001).

The rest of the chapter is structured as follows. The next section discusses a simple framework underlying our ideas and Section 5.3 sets out the estimation strategy of this framework. Section 5.4 describes the construction of the database and some descriptive statistics. Section 5.5 presents the OLS-estimates and Section 5.6 the IV-estimates. In Section 5.7 several other specifications are tested for robustness. Section 5.8 concludes.

5.2 Spatial wage differences and communication

Before we present the estimates of the returns to communication we set out a framework which captures the underlying mechanism. Our framework explains the existence of spatial wage differences and the role of communication. It relies on the assumption that in equilibrium wage differences can exist while workers and firms should be indifferent to location. Local markets (l) are characterised by (both observed and unobserved) ability, productivity level, price level, and industrial structure (specialisation level).

5.2.1 General setting

We consider an economy with perfect competition, free firm entry and full mobility of labour. Firms either focus on mass-products or on new and developing products. Firm's output is a function of productivity (A), number of workers (L) and city characteristics (C): $Y = f(A, L, C)$. These factors are mutually dependent. The productivity of a firm, for example, depends on its workers and its location and varies between mass-production and relatively newly initiated production (Duranton & Puga, 2001; Desmet & Rossi-Hansberg, 2009). The free entry assumption implies that firms obtain zero profits. As often noted in the literature, large spatial wage differences exist (Glaeser & Maré, 2001). The spatial wage differences are compensated by spatial variation in the input factors productivity, labour and city characteristics. In equilibrium workers and firms are indifferent regarding location l. The spatial variation in A and C explains why not all workers move to the high wage cities and not all firms move away from these cities.

5.2.2 Spatial distribution of firms

Following the theoretical work of Duranton and Puga, firms locate in a less specialised (or diversified) city during the learning stage in which they develop their ideal production process. In these 'nursery' cities firms learn from the ideas and knowledge of a broad variety of firms. Human capital externalities are crucial for the productivity and innovation of new products since the cross-fertilisation of ideas and knowledge stimulates the generation of new ideas (Lucas, 1988; Duranton & Puga, 2001; Desmet & Rossi-Hansberg, 2009). When firms find their optimal production process and move to mass-production they relocate to more specialised cities. Specialised cities house a co-agglomeration of similar firms which enables firms to share, match and learn from their direct competitors.

5.2.3 Productivity

The ability of the local work force varies over space (Combes et al., 2008). All firms in location l benefit from a productive labour force (ϕ_l). The determinants of local productivity vary with the local specialisation level (ρ_l). Firms who focus on mass-production and locate in specialised cities benefit from sharing facilities, matching labour and knowledge spillovers from similar firms. If the firm operates in the dominant local industry, productivity rises with the specialisation level (M^{ρ_l}).[1] A mature firm in the textile industry benefits from the co-location of textile industry and a high local specialisation level in this industry.

As indicated, both firms in specialised and diversified cities benefit from learning and communication with other firms. The cross-fertilisation of ideas is more likely to happen when people meet face-to-face. Not only is face-to-face contact a very efficient communication technology, it also helps solving incentive problems and more importantly facilitates learning and human capital externalities (Storper & Venables, 2004).[2] The amount of local knowledge spillovers and communication depends on the allocation of labour between core work activities and communication tasks. Core work activities are the job tasks of the worker's occupation. Communication tasks contain the communication with other workers (inside or outside the firm) about work activities. θ is the fraction of labour spent on communication tasks. The firm allocates labour optimally between work activities and

[1] M is the productivity effect of operating in the local dominant industry. This effect increases with the specialisation level of the city.

[2] This explains why human capital spillovers and learning are bound by distance (Jaffe et al., 1993; Jacobs, 1969).

communication tasks given local characteristics. However, learning and communication are more crucial for firms in less specialised cities who still optimise their production process by learning from others (Duranton & Puga, 2001; Desmet & Rossi-Hansberg, 2009).

To sum up, the productivity of a firm (A) depends on whether the firm operates in the local dominant industry (M), the specialisation level of the local industry (ρ_l), the amount of local communication (θL) and the ability or productivity of the local work force (ϕ_l). Firms which operate in the local industry experience a productivity which increases with the local specialisation level. The productivity benefits of local communication, on the other hand, decrease with the specialisation level of the city. Therefore,

$$A = M^{\rho_l} E^{1-\rho_l} \phi_l, \tag{5.1}$$

where: $0 < \rho_l < 1$, $E = d\theta L$ and d is a scalar indicating the level of productivity of the labour force.

Labour input to produce output only includes the fraction of labour spent on work activities ($(1-\theta)L$). Output is produced with labour spent on work activities (which decreases with the fraction spent on communication) at a productivity level that increases with the fraction spent on communication:

$$Y = A(1-\theta)L. \tag{5.2}$$

5.2.4 Optimal allocation of labour

Output is only produced with work activities while wages and rents are paid for both communication tasks and work activities (L). Local wages (W_l) and rents (R_l) are given. Congestion costs cause the local rents to rise with the size of the local market. Profits are defined as follows:

$$\pi = A(1-\theta)L - W_l L - R_l. \tag{5.3}$$

There is a trade-off between increasing productivity by spending labour on communication and increasing production output by spending labour on work activities. This trade-off varies with the local level of specialisation ρ_l. Firms maximise profits π, given the local dominant industry, specialisation level and rents, and optimally allocate labour between communication tasks and work activities. They optimise the following equation:

$$\pi = M^{\rho_l}(d\theta L)^{1-\rho_l}\phi_l(1-\theta)L - W_l L - R_l. \tag{5.4}$$

Optimising equation (5.4) leads to the following optimal allocation of labour between core activities $(1-\theta)$ and communication about core activities (θ), given the local specialisation level ρ_l:

$$(1-\theta) = \frac{\theta}{1-\rho_l}. \tag{5.5}$$

Substituting the optimal allocation of labour into equation (5.4) it follows that:

$$\pi = b\phi_l M^{\rho_l}(\theta L)^{2-\rho_l} - W_l L - R_l, \tag{5.6}$$

where $b = \frac{d^{1-\rho_l}}{1-\rho_l}$.

5.2.5 Individual wages

Firm entry is free which implies zero profits. This leads to the following total labour costs:

$$W_l L = b\phi_l M^{\rho_l}(\theta L)^{(2-\rho_l)} - R_l. \tag{5.7}$$

We assume that individual wages correspond to individual ability. Setting L to 1, individual wage of worker i is then:

$$W_i = b\phi_i M^{\rho_l}(\theta_i)^{(2-\rho_l)} - R_l. \tag{5.8}$$

The individual wage is determined by a constant, the worker's ability (ϕ_i), the level of local specialisation (ρ_l), whether the worker works in the dominant industry (M), the fraction of labour which the worker spends on communication (θ_i), and the average local rent costs (R_l). If the worker works in the dominant local industry, his wage rises with the local industrial specialisation of the relevant industry. However, the wage benefits of communication decrease with the local level of specialisation:

$$\frac{\partial W_i}{\partial \theta_i} > 0. \tag{5.9}$$

$$\frac{\partial W_i^2}{\partial \theta_i \rho_l} < 0. \tag{5.10}$$

5.3 Empirical strategy

5.3.1 Reduced form

We bring equation (5.8) to the data and estimate the reduced form for worker i in city l.

$$ln \quad w_{i,l} = \alpha_1 + \alpha_2\hat{\phi}_i + \alpha_3\hat{M}_i + \beta_1\hat{\theta}_i + \beta_2\hat{\rho}_l + \beta_3\hat{R}_l + \gamma_1(\hat{\theta}_i \cdot \hat{\rho}_l) + \gamma_2(\hat{M}_i \cdot \hat{\rho}_l) + \epsilon_{i,l}, \tag{5.11}$$

where $w_{i,l}$ is the hourly wage earnings of individual i, in city (Metropolitan Statistical Area) l. Individual ability is estimated by $\hat{\phi}_i$: a set of standard, demographical controls (age, age squared, gender, race and marital status), a set of occupational dummies and a set of education dummies of the highest grade completed. $\hat{M}_i$ represents the productivity effect of mass-production and indicates whether individual i works in the dominant industry in city l or not. Indicator $\hat{\theta}_i$ denotes the estimate of the performance of communication tasks by worker i.[3] The local level of specialisation is estimated with the Regional Specialisation Index (RSI). The RSI calculates the maximum over-representation (subject to national share) of an industry in the city. $\hat{\rho}_l = \max_l(\log E_{l,j} - logE_j)$ in which $E_{l,j}$ represents the employment share of industry j in city l and E_j the employment share of industry j in national employment. We allow the returns to communication to vary with the local specialisation level ($\gamma_1(\hat{\theta}_i * \hat{\rho}_l)$). The returns to working in the local dominant industry vary with the local level of specialisation as well ($\gamma_2(\hat{M}_i \cdot \hat{\rho}_l)$). Lastly, ($\hat{R}_l$) indicates the average rent in city l.

5.3.2 Measurement

The estimation of this empirical model requires a number of assumptions. First, the indicator for communication tasks ($\hat{\theta}_i$) and its interaction with local industrial specialisation ($\hat{\theta}_i \cdot \hat{\rho}_l$) are measured at aggregated levels and do not vary by worker. The dependent variable ($w_{i,l}$) is however measured at the worker level. This leads

[3] As explained in the next section, data limit us to measure communication at the occupational level.

to underestimation of the standard errors as indicated by Moulton (1990). To avoid this issue, we cluster standard errors at the occupational level.

Second, endogeneity issues may bias our OLS-estimates. The ability of individuals is estimated and not fully observed. The measurement error $\epsilon_{i,l}$ includes ability characteristics (A_i) such as talent and work discipline and some measurement error at the individual and city level ($\mu_{i,l}$): $\epsilon_{i,l} = A_i + \mu_{i,l}$. When A_i correlates with the local specialisation level ρ_l or city rents R_l, we cannot isolate the effect of these indicators on wages and the estimates become biased. To deal with endogeneity, Section 5.6 shows the results when instrumenting communication.

Third, specialisation and diversity are not opposite measures. The RSI_i measures the over-representation of an industry in city l while the local diversity level reflects the mixture of industries within the city. Thus, the regional diversity index (RDI_l) captures all industries in the city while RSI_l only includes information on the dominant industry.[4] We experiment with including both RSI_l and RDI_l.

Fourth, specialised cities tend to be smaller than diversified cities (Duranton & Puga, 2000).[5] Hence, $\hat{\rho}_l$ correlates with city size. The correlation between the size and the specialisation (and diversity) is too strong to include both in the regressions. Therefore, we attempt additional estimates with city size instead of specialisation or diversity and a cross-term of city size with communication.

Lastly, work activities might also involve communication. Especially low skilled service occupations often involve several communication tasks such as waiting tables. We aim however to measure the returns to communication *about* job activities, for example a worker who informs his manager about the results of his analyses. To distinguish between these two forms of communication we include information about communication work activities as well. Communication work activities are defined as the ONET work activity 'performing for or working directly with the public' with the description: 'Performing for people or dealing directly with the public. This includes serving customers in restaurants and stores, and receiving clients or guests'.

[4] RDI_l is defined as $RDI_l = \frac{1}{\sum_j (E_{l,j}/E_j)}$ where $E_{l,j}$ represents employment in industry j in city l and E_j national employment in industry j.

[5] This is the case in our dataset as well, see Section 5.4.2.

5.4 Data

5.4.1 Database construction

We use individual wage data for 2009 provided by the Current Population Survey (CPS). For each individual it contains information on personal characteristics (education level, age, marital status etc), occupation, industry, wage and location. Occupations are converted to a time-consistent scheme of 326 occupations as in Autor & Dorn (forthcoming). Our sample consists of working individuals living in Metropolitan Statistical Areas (MSAs) in 2009, aged between 16 and 65, working outside the agricultural sector. We exclude all self-employed workers. This results in a sample of 83,078 individuals.

Wages are measured by hour. Following Lemieux (2006), outliers are removed by trimming very small (hourly wage below $ 1) and very large values (hourly wage above $101) of wages. Hourly wages above $101 are top coded within the CPS and are therefore replaced with the 1.4 top coded value. For missing wage values we apply a no-imputation approach. The no-imputation method excludes the wages of missing cases but counts them when calculating occupational sizes (Mouw & Kallenberg, 2010).

Communication job tasks and work activities are collected from the Occupational Information Network (ONET) Skill Survey. The ONET data characterises the workers abilities, interest, knowledge, skills, work activities, work context and work values, by occupation. Three types of work activities and three work context items are included as communication job tasks. They measure the importance of:

- Establishing and maintaining interpersonal relationships (label ’relations’)
- Communicating with persons outside organisation (label ’external communication’)
- Communicating with supervisors, peers, or subordinates (label ’internal communication’)
- Face-to-face discussions (label ’face-to-face’)
- Work with work group or team (label ’teamwork’)
- Contact with others (label ’contact’)

Table 5.1 lists the ONET definition of these communication job tasks. We standardise the scores of these variables (a mean of zero and a standard deviation of one) to equalise scaling. The communication job task scores of the occupations are matched to the occupations in the CPS database. A Communication-Index is estimated by using a principal component analysis:

$$Y = \sum_{i=1}^{i=6} (\beta_i Communication_i + \epsilon_i). \tag{5.12}$$

Y is the constructed index based on the input of the six communication tasks represented by i. The estimates are presented in Appendix A.4, together with the correlations between the communication tasks. The principal component loadings (β_i) could be viewed as weights and are rather equal for all communication tasks in the first component. The first component explains about 0.60 percent of the total variation in the six tasks. The first component explains a substantial larger variation than the other components. Therefore, the first component is defined as the Communication-Index. It should be noted that the measurement error in the component is not taken into account in the analyses below.

Employment figures are gathered from the Local Area Unemployment Statistics from the Bureau of Labor Statistics Additional. The employment figures include information about the total city employment and the employment by industry (which is used for the construction of the local specialisation level). Lastly, additional city data, such as average rents, are collected from the Census Decennial Database.

Appendix A describes the data sources, the used classifications and Appendix A.4 includes a list of all the used variables, measurements and source.

5.4.2 Descriptive statistics

Before we proceed to present a set of estimates, we first discuss the descriptive statistics for our dataset. Table 5.2 provides an overview of the characteristics of our entire sample of 83,078 individuals.[6] The average worker earns 22 US dollars per hour, is 40 years old and works in a city with almost 1.3 million employees. One out of two workers is female. Individuals who perform more communication job tasks earn higher wages, live more often in diversified cities, are more often high skilled and female.

The last column of Table 5.1 shows the correlations between the performance of the six communication job tasks and individual wages. All the correlations are positive and significant. The establishment of relations, communication outside

[6] Note that the measurement of the variables results in possible values which conflict with the assumptions of the theoretical model. This does not affect the interpretation of the results.

Table 5.1. Communication job tasks

	Definition by ONET	Correlation with wage
Relations	Developing constructive and cooperative working relationships with others, and maintaining them over time.	0.34***
External communication	Communicating with people outside the organisation, representing the organisation to customers, the public, government, and other external sources. This information can be exchanged in person, in writing, or by telephone or e-mail.	0.39***
Internal communication	Providing information to supervisors, co-workers, and subordinates by telephone, in written form, e-mail, or in person.	0.35***
Face-to-face	How often do you have to have face-to-face discussions with individuals or teams in this job?	0.27***
Team work	How important is it to work with others in a group or team in this job?	0.14***
Contact	How much does this job require the worker to be in contact with others (face-to-face, by telephone, or otherwise) in order to perform it?	0.05***
Communication-index	Principal-component index of the above six tasks	0.35***

Note: source Current Population Survey 2009, n=81,262. *** significant at the 1% level.

Table 5.2. Summary statistics

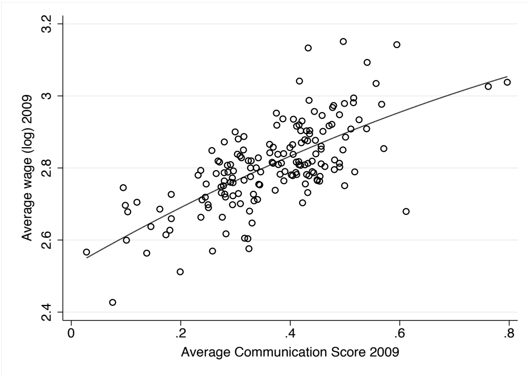
	Mean	Standard deviation	Minimum	Maximum	Correlation with communication
Hourly wage	21.90	16.26	2.49	230.6	0.35***
Log hourly wage	2.88	0.62	0.91	0.54	0.40***
Communication Index	0.43	0.99	-3.11	2.46	
Specialisation -city	0.00	1.00	-1.46	3.74	-0.05***
Diversity -city	0.00	1.00	-2.17	1.69	0.02***
Employment - city	1,311,017	1,136,008	60,580	4,328,589	0.01***
Dominant industry	0.01	0.11	0.00	1.00	-0.03***
High-school drop-out	0.08	0.27	0.00	1.00	-0.31***
High-school	0.26	0.44	0.00	1.00	-0.28***
Some college	0.29	0.45	0.00	1.00	-0.02***
College graduate	0.37	0.48	0.00	1.00	0.44***
Communication job activities	2.55	0.98	1.00	4.83	0.33***
Non-white	0.21	0.41	0.00	1.00	-0.04***
Non-married	0.45	0.50	0.00	1.00	-0.11***
Age	40.00	12.44	16.00	64.00	0.10***
Female	0.52	0.5	0.00	1.00	0.14***

Note: source Current Population Survey 2009, n=81,262. *** significant at the 1% level.

the organisation and communication with workers inside the organisation show the strongest correlations with individual wages. The measure for contact in general only weakly correlates with wages. Cities which house many communication intensive occupations also obtain high average wages (correlation of 0.71, significant at the 1 percent level, see Figure 5.1). The relation between local wages and local communication, as predicted in equation (5.6), does hardly show any outliers. Ann Arbor has the most communication intensive labour market and is the sixth city on the wage ranking. Canton-Massillon has the least communication intensive labour market and only 17 of the 168 cities have a lower average wage than Canton-Massilon.

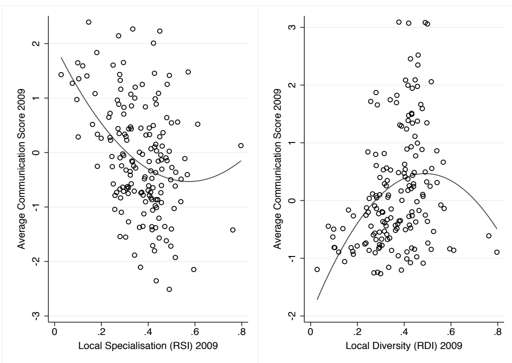
Equation (5.1) suggests that cities with a lower specialisation level benefit more from the performance of communication tasks. Indeed, workers in diversified cities perform on average more communication tasks, while workers in specialised cities perform less communication tasks (see Figure 5.2). Given a certain level of diversity or specialisation, the variation in communication is however large between cities. Appendix A.4 presents a correlation matrix of all variables.

Figure 5.1. Wages and communication in cities



Note: source Current Population Survey 2009. City level data, n=168. The correlation is 0.71 (0.00) and significant at the 1% level. Communication is measured as the average score on the Communication-Index as defined in Section 5.4. Wage is measured as average hourly wage 2009 in logs.

Figure 5.2. Communication in specialised and diversified cities



Note: source Current Population Survey 2009. City level data, n=168. The correlations are respectively -0.40 (0.00) and 0.33 (0.00) and significant at the 1% level. RSI_l and RDI_l are measured as described in Section 5.3. Communication is measured as the average score on the Communication-Index as defined in Section 5.4.

5.5 OLS-estimates

Before we address causality, we present a set of OLS-estimates to show the relationship between wage and communication in a more rigorous way. Column 1 in Table 5.3 presents the estimates of a straightforward wage regression. We find the usual returns to education, see Rauch (1993) and Bacolod et al. (2009). Both industrial specialisation and diversity correlate negatively with individual wages. The positive correlation between local diversity and individual wage (as found in Section 5.4.2) turns negative when we control for demographic and educational factors. Workers who work in the dominant local industry (M in equation (5.11)) earn substantially more than workers who do not work in the dominant industry. This effect increases with the specialisation level of the city. Notable is the positive impact of rents on wages which indicates the cities' role as centre for consumption (Glaeser et al., 2001). All the covariates, such as age and gender, obtain the expected sign and size.

Next, we test whether the correlations between wages and communication vary with the city's industrial specialisation and diversity level. Column (3) includes a cross-term between communication and the local specialisation level (all variables are standardised to ease comparison). The coefficient of the cross-term is negative and significant: the correlation between wage and performed communication tasks is weaker in specialised cities. The linear impact of communication remains positive and significant, while the size of the coefficient of local specialisation decreases. Column (4) performs the same regression but includes a cross-term between communication and sector diversity instead of sector specialisation. The coefficient of the cross-term is positive and significant. Both in specialised and in diversified cities workers in communication intensive jobs earn more, but this relation is stronger in diversified cities and weaker in specialised cities.

Lastly, we allow the relation between wages and communication to vary across city size. Diversified cities are on average larger than specialised cities. Column (5) presents the baseline results including city size instead of industrial structure and column (6) presents the results including the cross-term as well. The correlations between wage and performed communication tasks are stronger in larger cities. Workers in larger cities who perform communication tasks earn more than workers in small cities with the same task package. The positive coefficient of the cross-term between city size and communication outweighs the negative linear coefficient for communication.

Similar to the theory of Section 5.2, individual wages increase with the worker's ability, the worker's communication when the local industry is not very specialised and the specialisation level when the worker works in the dominant industry. The OLS-estimates suggest that a one standard deviation more communication job tasks increases individual wage with about 16 percent of a standard deviation.[7] In specialised cities this is 13 percent of a standard deviation, in diversified cities 18 percent and 20 percent of a standard deviation in large cities.

[7] The standard deviation of the dependent variable is 0.62. One standard deviation more communication results in an increase of 0.098 in the individual wage which is 15.8 percent of 0.62.

Table 5.3. Returns to communication, specialised and diversified cities (OLS)

	Dependent: individual wage (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Communication		0.098***	0.099***	0.098***		0.098***
		[0.019]	[0.019]	[0.019]		[0.019]
COM*specialisation			-0.019***			
			[0.004]			
COM*diversity				0.010***		
				[0.003]		
COM*size						0.023***
						[0.004]
Specialisation	-0.038***	-0.038***	-0.030***	-0.038***		
	[0.004]	[0.004]	[0.004]	[0.004]		
Diversity	-0.009**	-0.009***	-0.009**	-0.014***		
	[0.004]	[0.004]	[0.003]	[0.004]		
Size					0.043***	0.035***
					[0.004]	[0.003]
Dominant industry	0.105***	0.118***	0.115***	0.119***	0.107***	0.120***
	[0.026]	[0.023]	[0.023]	[0.023]	[0.026]	[0.020]
DOM*specialisation	0.066***	0.065***	0.057***	0.060***	0.047***	0.039**
	[0.018]	[0.018]	[0.017]	[0.017]	[0.017]	[0.017]
Drop-out	-0.204***	-0.185***	-0.183***	-0.185***	-0.204***	-0.181***
	[0.013]	[0.014]	[0.014]	[0.014]	[0.014]	[0.014]
College	0.080***	0.069***	0.069***	0.069***	0.080***	0.070***
	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]
College grad	0.363***	0.348***	0.347***	0.348***	0.373***	0.357***
	[0.021]	[0.020]	[0.020]	[0.020]	[0.021]	[0.020]
Rent	0.046***	0.047***	0.047***	0.047***		
	[0.003]	[0.003]	[0.003]	[0.003]		
Communication job	-0.018	-0.041***	-0.041***	-0.041***	-0.018	-0.041***
	[0.015]	[0.013]	[0.013]	[0.013]	[0.015]	[0.013]
Non-white	-0.085***	-0.082***	-0.082***	-0.082***	-0.062***	-0.061***
	[0.008]	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]
Non-married	-0.056***	-0.055***	-0.056***	-0.055***	-0.055***	-0.055***
	[0.007]	[0.006]	[0.006]	[0.006]	[0.007]	[0.006]
Age	0.049***	0.048***	0.049***	0.048***	0.050***	0.049***
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]
Age squared	-0.496***	-0.497***	-0.485***	-0.487***	-0.485***	-0.489***
	[0.035]	[0.034]	[0.034]	[0.034]	[0.034]	[0.033]
Female	-0.184***	-0.183***	-0.183***	-0.183***	-0.185***	-0.184***
	[0.014]	[0.013]	[0.013]	[0.013]	[0.014]	[0.013]
Occupation dummies	YES***	YES***	YES***	YES***	YES***	YES***
Observations	82,705	82,705	82,705	82,705	81,262	81,262
R-squared	0.438	0.445	0.446	0.445	0.433	0.440

Note: individual data. Communication represents the Communication-Index as defined in Section 5.4. Specialisation refers to the RSI_l, diversity to the RDI_l as defined in Section 5.3. City size is measured in standardised logs. Dominant industry is a dummy variable indicating whether the worker works in the local dominant industry or not. The dominant local industry obtains the highest specialisation level. High-school graduates are the reference group for education. Communication job refers to the importance of communication work activities in the job as defined in Section 5.4. See Appendix A.4 for a detailed description of the variables, measurement and data sources. Regressions also include a constant. Clustered standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

5.6 IV-estimates

The main issue with OLS wage estimates is a possible omitted ability bias. Equation (5.1) distinguishes between an ability and a productivity effect. This distinction is hampered if workers in highly productive cities or jobs are simply 'better' in an unobserved way. Ability characteristics such as talent, work discipline and ambition are unobserved in our analyses. For instance, relatively talented workers might be attracted to certain cities. Diversified cities tend to be larger and house more amenities than the smaller, specialised cities. Talented workers could value these amenities more than less talented workers. Talent of workers is however not measured. In OLS-estimates the higher wages within these cities are assigned to higher local productivity of these cities while they might simply reflect higher (unobserved) ability levels of their workers. The same feature might bias the impact of communication on wage. It could be the case that communication intensive jobs offer more carrier opportunities in the long run. Workers with a relatively high ambition are more likely to sort into these jobs. In the OLS-estimates, the high wages of these jobs are related to the communication intensity while the impact of worker ambition is unobserved. Combes et al. (2009) refer to this issue as the 'endogenous quality of labour' problem.

5.6.1 Instruments

Communication

We construct a language-skill proxy as an instrument for communication job tasks.[8] We assume workers with weaker language-skills to be less likely to perform communication job tasks. Transferring tacit knowledge is key to communication job tasks and strongly affected by language-skills. Language-skills are proxied by information on the worker's country of birth and the worker's parents. The country of birth indicates the worker's mother tongue. We assume workers who grew up in an English speaking household to obtain better language-skills (in the US) than workers who grew up in a non-English speaking household. The language-skill proxy obtains four values which are described in Table A.5 in Appendix A.4.

Our instrument should be exogenous and not affecting wage via other channels

[8] Charlot & Duranton (2004) instrument communication job tasks with the use of computers and internet at the work floor. The Current Population Survey includes similar information for the year 2000. However, we cannot rule out possible endogeneity of computer use. Workers may sort by ability into communication and computer intensive jobs for the same reasoning. Specification tests underline that computer use at the job is endogenous.

than communication. Clearly, the country of birth is not chosen by the individual and is exogenous. However, we do not observe the actual household language which might be endogenous. We assume such an effect to be negligible. Another possible issue with the proxy is that it might capture the sorting of migrants into certain cities. Figures A.1 and A.2 in Appendix A.4 present the relations between city's specialisation level, diversity level and communication level and the average native share in local occupations. The proxy does not seem to capture such sorting patterns.

Language-skills may affect wages via other channels than communication. For instance, the language-skill proxy captures cultural differences which could affect wage as well. Lewis (2011) finds that this effect is rather small. We test the validity of the instrument in Table 5.4. The first column shows a wage regression including both communication job tasks and the language-skill proxy. After controlling for communication, the language-skill proxy does not affect wage. Columns (2) and (3) show the OLS-estimates for communication and physical job tasks. Physical tasks are defined as 'handling and moving objects' and correlate negatively with wage. The next two columns show the first stage results for IV-estimates instrumenting respectively communication and physical job tasks with the language-skill proxy. The proxy correlates strongly with communication jobs tasks and not with physical job tasks. Columns (6) and (7) present the IV-estimates. The IV-estimates for communication (column (6)) correspond with the OLS-estimates. The IV-estimates for physical tasks are insignificant. The significant wage effect of physical tasks diminishes in the IV-estimates. These results indicate that our language-skill proxy does not measure a cultural wage effect.

Specialisation and diversity

Ciccone & Hall (1996) introduced the standard way to tackle the endogeneity problem of city size and productivity. The spatial population distribution in the US is (to some extent) persistent over time. The division of employment across cities is remarkably constant. Thus, the size of a city today can be predicted by the size of the city many decades ago. Today's main drivers of productivity strongly differ from the historical drivers. Thus, historical population of a city strongly correlates with today's city size but does not affect the current wages in the city. Clearly, today's wages cannot affect historical city population. This makes historical population a valid instrument for current city size, at least when the instrument is measured in the far past. For an extensive discussion on the validity and exogeneity of historical population as an instrument we refer to the work of Ciccone & Hall (1996) and Combes et al. (2009).

The sectoral specialisation and diversity of cities is correlated with size (respectively -0.66 and 0.57, significant at the 1 percent level). Therefore, we instrument sectoral specialisation and diversity with population in 1930.

The MSA population in 1930 is composed using Census Historical County Population figures. For each county this database includes decennial information on its population. We include population in 1930 since this is the first year with a decent covering across counties. Next, we sum county population by MSA (1990 definition) to construct MSA population in 1930. The MSA population in 1930 varies between 9,897 and 7,524,736 inhabitants.

5.6.2 Relevance of the instruments

Before we turn to the IV-estimates we test the relevance of our instruments. The correlation between population in 1930 and sectoral specialisation in 2009 is -0.51 and significant at the 1 percent level. For sectoral diversity this correlation is 0.61 (significant at the 1 percent level). Also the instrument of communication is strongly correlated with the communication index (0.58, significant at the 1 percent level).

Columns (1) and (2) of Table 5.5 show the first stage estimates for communication job tasks. In column (1) we include the city's sectoral specialisation and diversity level as explanatory variables while in column (2) we include city's population in 1930. The language-skill proxy seems to be a sound instrument for communication. Natives are relatively more present in communication-intensive occupations. The covariates show the usual sign and coefficients. By definition, the communication intensity of occupations does not vary across cities. This explains the insignificant coefficients of historical and sectoral structure.[9] Communication work activities (e.g. waiting tables) and communication job tasks (communicating about work activities) are positively correlated. The F-statistics show that the instrument for communication is valid.[10] Columns (3) and (4) present the first stage results for sectoral specialisation, with and without instrumenting communication as well. To produce interpretable results, we include the log of historical population. Historical city size is a decent predictor for current sectoral specialisation. The F-statistics indicate that historical population is a valid instrument for current specialisation level. Lastly, columns (5) and (6) show the first stage estimates for the industrial

[9] The importance of communication is measured at the occupation level and independent of location.
[10] F-statistics are generated for the additional instruments only (communication and population in 1930).

Table 5.4. Instrumental variables are valid

	OLS-estimates			IV-estimates: First stage		IV-estimates: Second stage	
	Wage (1)	Communication (2)	Physical (3)	Communication (4)	Physical (5)	Communication (6)	Physical (7)
Communication	0.096***	0.098***				0.109***	
	[0.023]	[0.019]				[0.042]	
Language-skill proxy	0.006			0.434***	0.008		
	[0.021]			[0.097]	[0.072]		
Physical			-0.058*				5.837
			[0.031]				[52.293]
Specialisation	-0.038***	-0.038***	-0.037***	-0.001	0.016**	-0.038***	-0.132
	[0.004]	[0.004]	[0.004]	[0.004]	[0.007]	[0.004]	[0.859]
Diversity	-0.009***	-0.009***	-0.009**	0.003	0.007	-0.009***	-0.051
	[0.004]	[0.004]	[0.004]	[0.003]	[0.006]	[0.004]	[0.385]
Dominant industry	-0.184***	-0.185***	-0.201***	-0.050**	0.056**	-0.183***	-0.517
	[0.014]	[0.014]	[0.013]	[0.023]	[0.022]	[0.015]	[2.891]
DOM*specialisation	0.069***	0.069***	0.079***	0.063***	-0.022	0.068***	0.202
	[0.008]	[0.008]	[0.009]	[0.016]	[0.023]	[0.009]	[1.117]
Drop-out	0.348***	0.348***	0.351***	0.108***	-0.212***	0.346***	1.594
	[0.020]	[0.020]	[0.019]	[0.029]	[0.044]	[0.021]	[11.041]
College	-0.041***	-0.041***	-0.014	0.175***	0.060	-0.043**	-0.373
	[0.013]	[0.013]	[0.015]	[0.051]	[0.070]	[0.017]	[3.293]
College grad	0.047***	0.047***	0.046***	-0.002	-0.007**	0.047***	0.090
	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.003]	[0.389]
Rent	0.119***	0.118***	0.100***	-0.086**	-0.080	0.120***	0.577
	[0.023]	[0.023]	[0.028]	[0.039]	[0.055]	[0.024]	[4.205]
Communication job	0.064***	0.065***	0.068***	-0.022	0.041	0.065***	-0.174
	[0.018]	[0.018]	[0.018]	[0.037]	[0.042]	[0.018]	[2.179]
Non-white	-0.081***	-0.082***	-0.084***	-0.004	0.015	-0.081***	-0.170
	[0.007]	[0.007]	[0.008]	[0.016]	[0.016]	[0.007]	[0.846]
Non-married	-0.055***	-0.055***	-0.056***	-0.014	0.006	-0.055***	-0.090
	[0.006]	[0.006]	[0.006]	[0.009]	[0.011]	[0.006]	[0.301]
Age	0.049***	0.048***	0.049***	0.010**	-0.006	0.048***	0.084
	[0.003]	[0.003]	[0.003]	[0.004]	[0.006]	[0.003]	[0.306]
Age squared	-0.485***	-0.485***	-0.491***	-0.118**	0.074	-0.484***	-0.928
	[0.034]	[0.034]	[0.035]	[0.049]	[0.060]	[0.034]	[3.869]
Female	-0.183***	-0.183***	-0.182***	-0.027	0.043	-0.182***	-0.435
	[0.013]	[0.013]	[0.013]	[0.032]	[0.052]	[0.013]	[2.255]
Occupation dummies	YES***	YES***	YES***	YES***	YES***	YES***	YES***
Observations	82,705	82,705	82,705	82,705	82,705	82,705	82,705
R-squared	0.445	0.445	0.441	0.736	0.626	0.445	
F-test				19.81	0.01		

Note: individual data. F-tests are done for additional instruments communication and population 1930. See Appendix A.4 for a detailed description of the variables, measurement and data sources. Regressions also include a constant. Clustered standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

diversity level of cities. Historical city size predicts current sectoral diversity even more precise than it predicts current sectoral specialisation. In diversified cities, workers perform more communication work activities while communication about these activities is indifferent from the average.

5.6.3 Results

In Table 5.6 the returns to communication are allowed to vary with city specialisation level, diversity level and city size. For each city characteristic (specialisation, diversity and size) we first present the baseline regression in which communication is instrumented with our language-skill proxy and the characteristic with population in 1930. The next column shows the IV-estimates with additional cross-terms between the language-skill proxy and the city characteristics. The IV-estimates provide similar results as the OLS-estimates.

The returns to communication remain positive and significant. An increase of the communication job tasks of one standard deviation raises the individual wage with about 18 percent of a standard deviation. The returns to communication are about 16 percent of a standard deviation in specialised cities (column (2)).

In large cities the returns to communication are somewhat higher (about 21 percent, column (6)). The coefficient of the cross-term between communication and diversity level becomes insignificant (column (4)). Especially in large, not specialised cities workers earn more when they perform more communication tasks.

The variation in returns to communication between different city types partly explains the lower wages in specialised cities. The negative specialisation wage premium decreases from 9 percent of a standard deviation to 8 percent. The urban wage premium decreases from 4 percent of a standard deviation to 2 percent of a standard deviation.

Table 5.5. First stage regressions

	Communication		Specialisation		Diversity	
	(1)	(2)	(3)	(4)	(5)	(6)
Language-skill proxy	0.434***	0.434***		0.010		-0.013
	[0.097]	[0.097]		[0.013]		[0.008]
Population 1930		0.001	-0.372***	-0.372***	0.477***	0.477***
		[0.002]	[0.005]	[0.005]	[0.003]	[0.003]
Communication			-0.003		0.006	
			[0.011]		[0.007]	
Specialisation	-0.001					
	[0.004]					
Diversity	0.003					
	[0.003]					
Dominant industry	-0.086**	-0.088**	-0.214***	-0.212***	-0.580***	-0.582***
	[0.039]	[0.039]	[0.068]	[0.068]	[0.059]	[0.058]
DOM*specialisation	-0.022	-0.023	0.845***	0.845***	0.135*	0.136*
	[0.037]	[0.039]	[0.036]	[0.036]	[0.072]	[0.071]
Drop-out	-0.050**	-0.050**	-0.023	-0.020	-0.023	-0.028*
	[0.023]	[0.023]	[0.015]	[0.014]	[0.015]	[0.015]
College	0.063***	0.063***	0.041***	0.040***	0.021***	0.023***
	[0.016]	[0.016]	[0.009]	[0.009]	[0.007]	[0.007]
College grad	0.108***	0.109***	-0.035***	-0.036***	-0.009	-0.007
	[0.029]	[0.029]	[0.010]	[0.010]	[0.015]	[0.015]
Communication job	0.175***	0.175***	-0.000	-0.003	0.016***	0.020***
	[0.051]	[0.051]	[0.007]	[0.007]	[0.006]	[0.006]
Rent	-0.002	-0.001	-0.277***	-0.277***	0.051***	0.051***
	[0.003]	[0.003]	[0.006]	[0.006]	[0.005]	[0.005]
Non-white	-0.004	-0.004	-0.162***	-0.161***	-0.101***	-0.102***
	[0.016]	[0.016]	[0.010]	[0.010]	[0.010]	[0.010]
Non-married	-0.014	-0.014	-0.040***	-0.040***	-0.014**	-0.014**
	[0.009]	[0.009]	[0.008]	[0.008]	[0.007]	[0.007]
Age	0.010**	0.010**	-0.007***	-0.007***	0.004**	0.004**
	[0.004]	[0.004]	[0.002]	[0.002]	[0.002]	[0.002]
Age squared	-0.118**	-0.118**	0.088***	0.088***	-0.037*	-0.037*
	[0.049]	[0.049]	[0.028]	[0.028]	[0.019]	[0.019]
Female	-0.027	-0.027	0.023***	0.023***	-0.012	-0.012
	[0.032]	[0.032]	[0.008]	[0.008]	[0.007]	[0.007]
Occupation dummies	1.464***	1.465***	-0.139***	-0.154***	-0.026	-0.004
Observations	82,705	82,705	82,705	82,705	82,705	82,705
R-squared	0.736	0.736	0.365	0.365	0.383	0.383

Note: individual data. See Appendix A.4 for a detailed description of the variables, measurement and data sources. Regressions also include a constant. Clustered standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Table 5.6. IV-estimates

	Dependent: individual wage (log)					
City instrument:	Specialisation		Diversity		Size	
	(1)	(2)	(3)	(4)	(5)	(6)
Communication	0.110***	0.108***	0.110***	0.110***	0.108***	0.106**
	[0.042]	[0.041]	[0.042]	[0.042]	[0.042]	[0.042]
COM*specialisation		-0.015***				
		[0.005]				
COM*diversity				0.006		
				[0.005]		
COM*size						0.027***
						[0.007]
Specialisation	-0.055***	-0.047***	-0.033***	-0.033***		
	[0.010]	[0.009]	[0.004]	[0.004]		
Diversity	-0.017***	-0.016***	0.000	-0.002		
	[0.005]	[0.005]	[0.005]	[0.005]		
Size					0.022***	0.012**
					[0.004]	[0.005]
Other controls	YES***	YES***	YES***	YES***	YES***	YES***
Observations	82,705	82,705	82,705	82,705	8,1262	81,262
R-squared	0.444	0.445	0.444	0.445	0.445	0.446

Note: individual data. City characteristics are instrumented by population in 1930. Communication is instrumented by language-skill proxy. Cross-terms are interactions of instruments. Regressions include controls for dominant industry, a cross-term of dominant industry with specialisation, education dummies, communication work activities, age, age squared, gender, marital status, occupational dummies and a constant. See Appendix A.4 for a detailed description of the variables, measurement and data sources. Clustered standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

5.7 Robustness

We test the robustness of our estimates by considering four robustness checks. Here, we only present the IV-estimates including the cross-term between communication job tasks and local specialisation level. The OLS-estimates and IV-estimates including the other cross-terms provide similar results and are available upon request. First, we test the sensitivity of the results towards the measure of communication (Section 5.7.1). Second, Section 5.7.2 discusses an additional test for the impact of unobserved ability. Next, we add cross-terms between communication and individual skill level to our analyses (Section 5.7.3). Lastly, the measure of local specialisation level is put to the test (Section 5.7.4).

5.7.1 Other measures of communication

To address the validity of our results we test three alternative ways to measure communication job tasks. First, we measure communication job tasks as the share of all job tasks. This indicator measures the importance of communication relative to other job tasks instead of the absolute importance of communication. Columns (1) and (2) in Table 5.7 present the IV-estimates. The relative returns to communication are significantly larger than the absolute returns to communication. An increase of one standard deviation in relative communication leads to an increase of 41 percent of a wage standard deviation. Within specialised cities this return is only 32 percent of a standard deviation. The returns to communication do not differ across local diversification levels while the returns in large cities are 52 percent of a standard deviation.

Second, we consider the wage returns to non-routine interactive tasks. Information and communication technology (ICT) acts as a substitute for some tasks and a complement for others (Autor et al., 2003). Computer technology replaces labour in performing routine tasks that can easily be described with programmed rules, such as the repetitive tasks of clerks and cashiers (Bresnahan, 1999). On the other hand, non-routine tasks, such as managing others, legal writing and selling, cannot, as of yet, be described as a set of programmable rules. Non-routine tasks require an adaptive attitude of the worker; these are typically tasks involving communication, interaction and knowledge transfer. The rival effects of computer technology on routine tasks on the one hand and non-routine on the other hand relate to the rival spatial effects of technology as indicated by Glaeser & Ponzetto (2010), Gaspar & Glaeser (1998) and Ioannides et al. (2008). Autor & Dorn (forthcoming) show that cities which initially specialised in routine-intensive occupations obtain employment and wage polarization after 1980. Clearly, non-routine interactive and communication tasks are strongly related (0.72, significant at the 1 percent level). The first stage regression shows a strong correlation between the language-skills proxy and the non-routine interactive tasks of an occupation.[11] Columns (3) and (4) of Table 5.7 present the IV-estimates with the linear and cross-terms of non-routine interactive tasks instead of communication job tasks. The IV-estimates indicate a positive return to the performance of non-routiness interactive tasks of about 25 percent of a standard deviation. This return is - as expected - somewhat lower in specialised cities (about 21 percent of a standard deviation) and somewhat higher in diversified cities (about 30 percent of a standard deviation).

The last measure of communication stems from the work of Borghans et al. (2006) and Bacolod et al. (2009) and measures the interpersonal skill requirements of the job: the importance of 'people skills'. We calculate the importance of 'people skills' by the importance of six ONET skill variables: social perspectives, coordination, persuasion, negotiation, instruction and service orientation. The last three columns of Table 5.7 present the results. Including people skills instead of communication job tasks does not change the results. There are positive wage returns to the performance of people skills in cities, these returns increase with the size of city and decrease with the specialisation level of the city.

5.7.2 Unobserved ability

Sorting of workers by unobserved ability is a commonly acknowledged measurement issue for spatial wage estimations (Combes et al., 2008). Ideally, we would eliminate unobserved worker heterogeneity using a large panel of individuals. The CPS is not a panel but has a time dimension. We aggregate the individual data to the city level (MSA) to obtain a panel of cities. Additionally to our IV-estimates we take the first difference of local variables and remove the unobserved ability bias using the time dimension.

As discussed in Section 5.3.2, unobserved ability (A_i) could cause biased estimates when it correlates with other explanatory variables. We assume that unobserved ability A_i (such as personal talent, ambition and work discipline) is time invariant. Taking the first difference removes the eventual ability bias. To do so, we

[11] The index is defined as in Acemoglu & Autor (2011). The index is standardised with a mean of zero and a standard deviation of one. Appendix A.4 describes the measurement of this index.

Table 5.7. Other measures of communication - IV-estimates

	Dependent: individual wage (log)					
	Relative communication		Non-routine interactive		People skills	
	(1)	(2)	(3)	(4)	(5)	(6)
Communication	0.264**	0.259**	0.162**	0.160**	0.133***	0.131***
	[0.132]	[0.131]	[0.077]	[0.077]	[0.049]	[0.049]
COM*specialisation		-0.047***		-0.023***		-0.019***
		[0.016]		[0.008]		[0.006]
Specialisation	-0.049***	-0.005	-0.058***	-0.057***	-0.052***	-0.042***
	[0.010]	[0.020]	[0.011]	[0.011]	[0.010]	[0.010]
Diversity	-0.013***	-0.012**	-0.018***	-0.017***	-0.015***	-0.014***
	[0.005]	[0.005]	[0.006]	[0.006]	[0.005]	[0.005]
Other controls	YES***	YES***	YES***	YES***	YES***	YES***
Observations	82705	82,705	82,705	82,705	82,705	82,705
R-squared	0.413	0.414	0.412	0.412	0.435	0.436

Note: individual data. Relative communication is the importance of communication relative to all other work activities and work context. Non-routine interactive tasks are measured as in Acemoglu & Autor (2011). Regressions include controls for dominant industry, a cross-term of dominant industry with specialisation, education dummies, communication work activities, age, age squared, gender, marital status, occupational dummies and a constant. See Appendix A.4 for a detailed description of the variables, measurement and data sources. Clustered standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

add a time dimension to equation (5.11):

$$\begin{aligned} ln \quad w_{i,l,y} = \alpha_1 + \alpha_2\hat{\phi}_i + \alpha_3\hat{M}_{i,y} + \beta_1\hat{\theta}_{i,y} + \beta_2\hat{\rho}_{l,y} + \beta_3\hat{R_{l,y}} + \gamma_1(\hat{\theta_{i,y}} \cdot \hat{\rho_{l,y}}) \\ + \gamma_2(\hat{M}_{i,y} \cdot \hat{\rho_{l,y}}) + \epsilon_{i,l,y}. \end{aligned} \tag{5.13}$$

Individual ability ($\hat{\phi}_i$) is constant over year y. The amount of communication tasks the worker performs ($\hat{\theta_{i,y}}$), the specialisation level ($\hat{\rho}_{l,y}$) and the size of the city ($\hat{R_{l,y}}$) may change over time. The measurement error includes the ability of worker i (A_i which is constant over time and place) and some measurement error at the individual, city, time level ($\mu_{i,l,y}$): $\epsilon_{i,l,y} = A_i + \mu_{i,l,y}$

To obtain a panel of cities we aggregate all indicators to the city level l:

$$\begin{aligned} \Delta ln \quad w_l = \alpha_2\Delta\hat{\phi}_l + \alpha_3\hat{M}_l + \beta_1\Delta\hat{\theta}_l + \beta_2\Delta\hat{\rho}_l + \beta_3\Delta\hat{R}_l + \gamma_1(\Delta\hat{\theta}_l \cdot \Delta\hat{\rho}_l) \\ + \gamma_2(\Delta\hat{M}_l \cdot \Delta\hat{\rho}_l) + \Delta\epsilon_l, \end{aligned} \tag{5.14}$$

in which $\Delta\epsilon_l$ does not include unobserved ability. Table 5.8 presents the estimates of this model for the period 2006-2009. The results hold for several time periods. The estimates resemble the IV-estimates. The change in communication tasks at the MSA level between 2006 and 2009 is positively related with the change in MSA wage. The coefficients of the cross-term with sector specialisation is negative and significant, the cross-term with diversity insignificant and the cross-term with size is positive and significant.

5.7.3 Skill level

Especially the spatial clustering of high-skilled workers relates to higher local wages (Glaeser & Maré, 2001; Glaeser & Gottlieb, 2009). Skilled workers cluster in certain cities (e.g. New York, San Francisco) and these cities tend to be the ones with higher wages (Rauch, 1993) and higher growth rates (Glaeser et al., 1995). Table 5.5 showed strong correlations between the sectoral structure of cities and the skill level of their inhabitants. Do high-skilled workers benefit more from performing communication tasks than low-skilled workers? The first two columns of Table 5.9 present the IV-estimates including cross-terms between communication and educational dummies. The cross-terms are insignificant while our variables of interest are hardly affected by the inclusion of these additional explanatory variables.

5.7.4 Industrial structure

Lastly, we test the sensitivity of the results to changes in the measure of the local industrial structure. The bias in the classification of sectors might hamper the estimates of our indicators for the local industrial specialisation and diversity level. Overall, manufacturing sectors are defined at a more detailed level in the classification than service sectors. A diverse local structure of manufacturing sectors therefore obtains a higher RDI than a diverse local structure of service sectors. Indeed, the variation in specialisation and diversity in manufacturing sectors is larger than the variation in service sectors. The last column of Table 5.9 presents IV-estimates in which only manufacturing sectors (column 3) and only service sectors (column 4) are included in the RSI. The returns to communication job tasks vary with the local specialisation level of both manufacturing and service sectors. As expected, the variation in the local manufacturing specialisation obtains a stronger impact than the variation in the local service sector.

Table 5.8. First differences at MSA level

	Dependent: change average MSA wage (2006-2009)				
	(1)	(2)	(3)	(4)	(5)
Communication		0.086***	0.069***		0.062***
		[0.015]	[0.017]		[0.017]
COM*specialisation		-0.030***			
		[0.005]			
COM*diversity			0.012		
			[0.009]		
COM*size					0.018**
					[0.008]
Specialisation	-0.055***	-0.043***	-0.055***		
	[0.002]	[0.002]	[0.001]		
Diversity	-0.010***	-0.011***	-0.016***		
	[0.002]	[0.002]	[0.004]		
Size				-0.005	0.040***
				[0.008]	[0.004]
Dominant industry	0.042	0.099	0.072	0.579**	0.144**
	[0.078]	[0.068]	[0.070]	[0.241]	[0.060]
DOM*specialisation	0.132**	0.111**	0.122**	-0.372**	0.070*
	[0.053]	[0.047]	[0.049]	[0.163]	[0.041]
Drop-out	-0.279***	-0.243***	-0.253***	-0.172**	-0.250***
	[0.027]	[0.024]	[0.024]	[0.083]	[0.022]
College	0.094***	0.095***	0.096***	0.086	0.087***
	[0.019]	[0.016]	[0.017]	[0.059]	[0.015]
College grad	0.378***	0.378***	0.377***	0.429***	0.369***
	[0.023]	[0.020]	[0.021]	[0.072]	[0.019]
Rent	-0.024	-0.030*	-0.027	-0.061	-0.029**
	[0.018]	[0.016]	[0.016]	[0.057]	[0.015]
Communication job	-0.073***	-0.073***	-0.070***	-0.026	-0.076***
	[0.016]	[0.014]	[0.015]	[0.051]	[0.013]
Non-white	0.009***	0.009***	0.009***	0.004**	0.009***
	[0.001]	[0.001]	[0.001]	[0.002]	[0.000]
Non-married	0.000	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Age	-0.192***	-0.202***	-0.202***	-0.276***	-0.202***
	[0.024]	[0.021]	[0.022]	[0.077]	[0.019]
Age squared	0.328***	0.247***	0.256***	0.385***	0.265***
	[0.032]	[0.037]	[0.039]	[0.103]	[0.034]
Female	0.241***	0.177***	0.180***	0.348***	0.200***
	[0.036]	[0.038]	[0.040]	[0.113]	[0.035]
Occupation dummies	YES***	YES***	YES***	YES***	YES***
Observations	165	165	165	165	165
R-squared	0.981	0.986	0.985	0.807	0.988

Note: city data (aggregated individual data). See Appendix A.4 for a detailed description of the variables, measurement and data sources. Clustered standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Table 5.9. Additional variation: skill levels, industry and services - IV-estimates

	Dependent: individual wage (log)			
	Skill cross-terms		Manufacturing	Services
	(1)	(2)	(3)	(4)
Communication	0.114***	0.113***	0.182***	0.125***
	[0.036]	[0.036]	[0.055]	[0.047]
COM*specialisation		-0.016***	-0.069***	-0.015
		[0.005]	[0.021]	[0.009]
COM*drop-out	-0.002	-0.003		
	[0.006]	[0.005]		
COM*college	0.001	0.001		
	[0.007]	[0.007]		
COM*college grad	-0.013	-0.015		
	[0.041]	[0.041]		
Specialisation	-0.055***	-0.048***	-0.037***	-0.047***
	[0.010]	[0.010]	[0.008]	[0.006]
Diversity	-0.017***	-0.016***	-0.002	0.001
	[0.005]	[0.005]	[0.003]	[0.017]
Other controls	YES***	YES***	YES***	YES***
Observations	82,705	82,705	82,705	82,705
R-squared	0.443	0.444	0.446	0.445

Note: individual data. All variables are standardised with a mean of zero and a standard deviation of one. Regressions include controls for dominant industry, a cross-term of dominant industry with specialisation, education dummies, communication work activities, age, age squared, gender, marital status, occupational dummies and a constant. See Appendix A.4 for a detailed description of the variables, measurement and data sources. Clustered standard errors are in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

5.8 Discussion

The debate in the empirical literature and economic regional policy has been largely about stimulating fruitful environments. The success of clusters like Silicon Valley and diversified cities such as New York City stimulated many scientific and policy projects on this subject and incited a massive literature on agglomeration economies. Many papers focus on the question whether specialised or diversified cities are the most fruitful environments. Duranton & Puga (2001) were the first to point out that both types are important in a system of cities. The question remains however how to induce such a beneficial environment and whether the advantages of proximity are similar in both city types.

A major advantage of cities seems to lay in the role of proximity for the communication of tacit knowledge and for learning from each other. Jaffe et al. (1993) show that distance bounds patent citation. Bacolod et al. (2009) and Florida et al. (2012) show that the returns to certain skills, such as social skills, increase with city size. Charlot & Duranton (2004) find positive returns to communication in French cities. This chapter takes a step towards unravelling the advantages of specialised and diversified cities by analysing the role of communication in both city types. We show substantial wage returns to communication in both diversified and specialised US cities. Given their occupation, workers who communicate more are more valued by firms. These returns decrease however with the specialisation level of the urban area. Communication is positively valued in all city environments but plays more of a key role in diversified cities.

In line with the work of Duranton & Puga (2001) and Desmet & Rossi-Hansberg (2009) we relate these findings to differences in the production processes of firms across specialised and not-specialised (diversified) cities. The higher value of communication in diversified cities seems to result from a more crucial role of learning in these cities. Specialised and diversified cities have different comparative advantages. Within their location choice, firms exploit these local comparative advantages. Our results suggest that social and communication skills are more valued in diversified than in specialised cities. In terms of urban policy implications, the results indicate that there is no one-policy-fits-all urban development policy as the comparative advantages vary across city types.

Chapter 6

Summary

and research agenda

In the age of information and communication technology, cities continue to play a key role in production processes. Both workers and firms concentrate in expensive cities. The current role of cities in economic processes has been extensively studied. A major road in these analyses is studying the impact of skill structure on city development. This study takes a new angle and investigates the role of a city's task structure. The breaking-up of production processes into separate parts makes such a task approach insightful. A task approach elucidates what it is that cities facilitate in today's production processes. The previous chapters showed substantial interactions between tasks, occupations and cities. Below, the main findings of this study are summarised following the considered connections and the questions raised in the introduction. This section ends with an agenda for further research.

The economic importance of proximity and especially human interactions explains the role of cities in the age of information and communication technology (Gaspar & Glaeser, 1998). Cities remain important in production processes, but the characteristics of their value altered. Today, cities specialise in certain functions instead of industries (Duranton & Puga, 2005) and focus more on producing ideas instead of goods (Glaeser & Ponzetto, 2010). These changes in comparative advantages and structures of cities demand a new approach towards explaining the rise and fall of cities. Chapter 2 analyses the connection between task structure and city development and focuses on the question:

1. Does the connectivity between tasks explain employment development across cities?

To answer this question the chapter employs a task approach in explaining the rise and fall of the largest 168 US cities in the period 1990-2009. The initial task structure of 41 job tasks of these cities explains a substantial part of the changes in the distribution of employment across them. Moreover, it is shown that the spatial interconnectivity between job tasks is a key comparative advantage of US cities today. The benefits from vicinity to other job tasks and human interactions vary across tasks. Some tasks heavily depend on proximity, while others can be easily performed in isolation. The relevance of interconnectivity of tasks affected the employment development across cities in the last decades. Cities with a large share of occupations that initially performed tasks that demand close performance of other tasks, gained in terms of employment in the last decades. The initial economic structure of these cities fitted the comparative advantage of cities. However, cities with a large share of occupations which used to perform tasks that can be easily performed at distance experienced a loss in employment share. The additional explanatory power of the interconnectivity of the task structure to the one of the skill structure underlines the relevance of a task approach. Successful cities of yesterday contained industries that benefit from co-agglomeration, successful cities of today perform job tasks that benefit from co-agglomeration.

The task approach also contributes to documenting economic activities across space. Employing task data visualises underlying spatial variation within jobs and industries. Chapter 3 studies the connection between cities and the structure of job tasks and the relating question:

2. Do workers in large cities specialise in a smaller subset of tasks and develop more specialised skills than workers located outside these cities?

This chapter is an empirical investigation of this question and employs German task data from the BIBB. Defining jobs as occupation-industry combinations, the task packages of jobs across German cities of different size are measured. The task packages vary substantially between small, medium and large German cities. First, jobs hold a smaller subset of tasks when they are performed in large cities than in small cities. Workers in large cities are more specialised. Second, the demanded skills for a job go up with the size of the local market. Likely, the higher specialisation level of workers in large cities generates more time to focus on their core task(s) and develop more task specific skills (Becker & Murphy, 1992). Chapter 3 shows that comparing jobs (or industries) and their wages across space is a risky approach. The different job contents across space likely impact spatial wage differentials.

The thick labour markets in cities not only affect the content of jobs but also the assignment of worker skills to job contents. Workers tend to be more efficient if their skills match their job tasks well. The existence of a large variety of worker skills and job contents makes the match between the two complex. Education levels do not capture the full range of worker skills, while occupation codes do not capture the full range of job tasks. Chapter 4 analyses the connection between cities and the match of worker skills to job tasks. By doing so it answers the question:

3. Does the thick labour market in large cities result in better matches of heterogeneous workers to heterogeneous job tasks?

Empirically, this chapter applies a survey among Dutch workers who, among others, indicate the quality of the match between their skills and job tasks. The match of heterogeneous worker skills to heterogeneous job tasks is better among workers in dense cities than among workers in the scarcely populated Dutch countryside. The availability of more choice, both from the supply and the demand side, increases the fastidiousness and improves the quality of the match between the two. Chapter 4 takes additional skill investments on top of education into account and distinguishes occupations by their job tasks. More skilled workers and more complex jobs are more often observed in the dense cities. These better workers and better jobs suffer more from a bad job match than less good workers and jobs. This explains their over-representation in dense cities with high quality matches. Referring to the positive wage returns of the quality of labour matches, we conclude that matches are more efficient in dense labour markets in the Netherlands.

Lastly, this study applies a task approach in a contribution to the debate about the advantages of specialised and diversified city environments. Both types of cities seem to be important in a system of cities relating to different stages in production processes. The most salient agglomeration economy of cities is that proximity lowers communication costs and facilitates human interactions to simulate learning. Chapter 5 studies whether advantages of proximity are similar in both city types. It does so by comparing the connection between city structure and the value of communication jobs tasks and answers the question:

4. Are communication tasks equally valuable in specialised and diversified cities?

Empirically, this chapter relates the differences in production processes of firms across both city types to different advantages of communication in these cities in the United States. Within diversified cities, communication occurs between workers from different fields. This complicates communication since workers have different knowledge, use different jargon, etc. The communication in specialised cities reflects communication between workers with similar job tasks. Therefore, communication is likely more specialised but also less complex as workers 'understand' each other better. Chapter 5 compares the wage returns to performing communication job tasks across 168 US cities. Within all cities, the wage returns of communication jobs tasks are positive. The returns to communication are however larger in diversified than in specialised cities. This suggests that comparative advantages vary across city types. Firms exploit these comparative advantages with their location choice.

6.1 Research agenda

The task approach is a relatively new and emerging approach. It faces many challenges, both conceptually and empirically as indicated by, among others, David Autor (2013). Applying the task approach in the context of cities and regions generates additional challenges, especially regarding empirical analyses. This study suggests that a task approach is worthwhile pursuing at the city level and contributes to the understanding of city economies.

Table 6.1 displays the sources for national task data. The task measurement in all these datasets faces substantial limitations.[1] Additionally, only two of the five datasets include regional variation in tasks. This regional information is however categorised and does not include information about actual location. It is therefore impossible to relate important regional factors, such as industrial structure and education level, to these datasets.

Besides these data limitations, several challenges for further research and pressing questions remain. Here, we suggest four ways in which further research could contribute to the understanding of what happens inside and outside cities.

Table 6.1. Sources of task data

	Country	Spatial information	Main other limitations
Dictionary of Occupational Titles (DOT)	United States	None	Infrequent updates Likely status quo bias
Occupational Information Network (ONET)	United States	None	No time series
IAB/BIBB labor force data	Germany	Categorised by population	Questions vary over time Self-reported information
British Skill Survey	United Kingdom	None	Self-reported information
Dutch Skill Survey	Netherlands	Categorised by density	Self-reported information

First, more insight in what spatially connects the performance of tasks is valu-

[1] See Autor (2013) for a discussion about the limitations.

able. The contribution of cities in the current economy seems to lay in the importance of proximity and scale benefits. Chapter 2 shows that the connectivity between tasks explains the development of cities. But what determines this interconnectivity? Why are some tasks performed in close vicinity and others not? Following the influential work of Blinder (2006) an extensive literature focuses on the 'offshorability' classification of tasks and jobs. The key attribute that defines a task as 'offshorable' is that is does not depend on face-to-face worker contact or close proximity between worker and customer. However, it might be possible to offshore a task but not preferable to perform it in isolation. For example, a researcher is able to perform all his tasks alone and far away from co-workers and customers. He will be more productive when he is located in the close vicinity of other workers and captures the benefits from knowledge spillovers. However, it does not matter where this cluster of co-workers and customers is located. Put simply: the offshorability measure does not include the proximity and scale benefits of certain tasks. A clear definition of what bundles tasks together in space would provide more insight in global supply chains and the role of cities in these chains.

Second, taking into account the hinterland of cities would be a valuable extension. Currently, most studies view cities as isolated places while the variation in the hinterland of cities likely affects their economic structure as well. The hinterlands of cities in rich countries, for instance, fulfil a different role for city economies than hinterlands in poor countries. Cities which are located in a densely populated region experience different relations with their hinterland than cities in a scarcely populated region. Networks of cities divide labour in another way than isolated places. More insight in the role of various hinterlands and geographical linkages would add to the knowledge about city economies and their task structures.

A third interesting extension would be to include worker skills in the analyses of the interactions between tasks, jobs and cities. A broad literature relates the success of cities to the preference of skilled workers to live in certain cities (Combes et al., 2008; Lee, 2010). High-skilled workers sort into different task packages compared to less skilled workers. Furthermore, the task packages of jobs may be flexible towards the worker who performs them. Unravelling the interplay between task packages and location preferences of skilled workers is an interesting and important challenge for further research.

Lastly, the assignment of skills to job tasks and the task structure of economies altered the last decades. Little is known about the spatial changes in the task structures. To place the role of cities in past, current and future economic structures, such analyses are however very relevant.

In summary, the relatively new interest in the task approach results in many interesting observations but also in several pressing questions and challenges. The main challenge seems to be overcoming the data limitations which result in a lack of knowledge about spatial patterns of job contents. The substantial spatial variation within German jobs and Dutch jobs (see Chapter 3 and 4) and the substantial changes over time in the United Kingdom (Akcomak et al., 2013) indicate a possible measurement error caused by these limitations. The continuing key role of cities in today's production processes makes knowledge about spatial task patterns and the connections between tasks, skills, jobs and cities relevant for analyses and policy on regional labour markets.

APPENDIX A

US DATA

A.1 Data description

Current Population Survey — May Outgoing Rotation Group

The Current Population Survey (CPS) is a monthly household survey of the US government. It contains information about employment and other labour-market variables. For each person it provides information on occupation, industry, hours worked, earnings, education, and unionisation. The data also contain background variables such as age, sex, race, ethnicity, geographic location. We use the May Merged Outgoing Rotation Group (MORG) files in which more detailed information about earnings and working hours are available. We use the years 1990-2009 because the residence of the respondent is available in terms of Metropolitan Statistical Areas (MSA). We assume that the respondents work in the same MSA as they live. In 1990 67 percent of the respondents lives in a MSA, in 2009 this is almost 75 percent.

ONET

Task information is gathered from the ONET Database (www.onetcenter.org). For each occupation, this database provides information about the importance of 41 work activities. Work activities are defined as 'General types of job behaviours occurring on multiple jobs'. Initial information of the ONET database is based on data from occupation analysts. This information is supplemented and updated by ongoing surveys of each occupation's worker population and occupation experts. The level of importance of the activities is measured by the question: How important is the work activity to the performance of the job? The importance is scaled from 1 (not important at all) to 5 (extremely important). The database consists of a cross-section, which is updated over time. The 3.0 version is used for this study.

Local Area Unemployment Statistics

The employment data for counties is collected from the Local Area Unemployment Statistics of the Bureau of Labor Statistics (BLS). We use county data for employment statistics instead of Metropolitan Statistical Areas (MSAs). The border definitions of MSAs change over time, so growth statistics are biased. Counties are merged into MSAs following the 1990 definition of the Census. Details on the construction of the city classifications are given below.

Census 1990 and 2000

The share of high-skilled people and the mean rent by MSA is gathered from the Census data. For each MSA it contains information on the number of people that have obtained at least a Bachelor's degree in 1990 and 2000. We do not have information on the share of high-skilled people or rents by city for other years.

A.2 Classifications

Cities

In 2009, MSAs were responsible for more than 85 percent of the employment, income, and production of products and services in the United States (Bureau of Economic Analysis). MSAs are defined by the nature of their economic activity. The scope of regional economic activity increases over time, which is replicated in the borders of the MSA classification. To analyse the development of economic structure within cities, we need a consistent city classification. To do so, we use MSA definitions by combining counties following the 1990 definition of the Census. Since county borders do not change over time, our MSA classification represents cities, which do not change in geographical size over time. Due to a change in classification of MSAs in 2005 we loose a small fraction of our sample. The definition of the Census is optimised for this break in classification. Our city classification consists of 168 MSAs, which are stable over time.

Industries

The Census industry classification changes within the period 1990-2009. We use a three-digit consistent classification provided by David Dorn and used in Autor & Dorn (forthcoming). The classification includes 142 industries at the three-digit level. For the two-digit level we distinguish 11 industries. We exclude the agriculture industry. To verify the CPS distribution of industries over MSAs we compare it with the County Business Pattern data. Using data from the County Business Patterns instead of the CPS does not change the results.

Occupations

The Census classification for occupations changes over time as well. We make use of the occupation classification in Autor & Dorn (forthcoming). This classification includes 326 occupations, which are consistently defined over time.

A.3 Data appendix chapter 2

Tasks are defined as 'General types of job behaviours occurring on multiple jobs'. The ONET database provides the importance of 41 work activities for occupations following the Standard Occupation Classification (SOC 2000). The SOC occupational classification scheme of the ONET database is matched to the Census 2000 occupational classification scheme. This scheme is collapsed to the 326 consistent occupations. Table A.1 shows the 41 tasks, their task group and descriptive statistics.

Table A.1. Summary statistics of tasks

Task name (ONET definition)	Task group (ONET definition)	Employment share	Mean	Standard deviation	Task connectivity
Getting Information	Information input	0.03	4.08	0.42	0.30
Monitor Processes, Materials, or Surroundings	Information input	0.03	3.42	0.48	-1.39
Identifying Objects Actions, and Events	Information input	0.03	3.66	0.43	-0.97
Inspecting Equipment, Structures, or Material	Information input	0.02	3.17	0.81	-1.65
Estimating the Quantifiable Characteristics of Products, Events, or Information	Information input	0.02	2.92	0.49	-0.84
Judging the Qualities of Things, Services, or People	Mental processes	0.03	3.16	0.43	-0.50
Processing Information	Mental processes	0.03	3.32	0.62	0.84
Evaluating Information to Determine Compliance with Standards	Mental processes	0.03	3.32	0.56	0.19
Analysing Data or Information	Mental processes	0.02	3.09	0.71	0.97
Making Decisions and Solving Problems	Mental processes	0.03	3.70	0.53	0.86
Thinking Creatively	Mental processes	0.03	3.11	0.64	0.92
Updating and Using Relevant Knowledge	Mental processes	0.03	3.48	0.57	0.86
Developing Objectives and Strategies	Mental processes	0.02	2.77	0.57	0.98
Scheduling Work and Activities	Mental processes	0.02	2.93	0.54	0.89
Organising, Planning, and Prioritising Work	Mental processes	0.03	3.46	0.50	0.86
Performing General Physical Activities	Work output	0.02	2.98	0.86	-1.69
Handling and Moving Objects	Work output	0.02	3.01	0.87	-1.71
Controlling Machines and Processes	Work output	0.02	2.81	0.95	-1.68
Operating Vehicles, Mechanised Devices, or Equipment	Work output	0.02	2.45	0.98	-1.68
Interacting With Computers	Work output	0.03	3.12	1.02	0.96
Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment	Work output	0.02	1.97	0.71	-1.00
Repairing and Maintaining Mechanical Equipment	Work output	0.02	2.22	0.91	-1.70
Repairing and Maintaining Electronic Equipment	Work output	0.02	2.01	0.67	-1.48
Documenting/Recording Information	Work output	0.03	3.32	0.71	0.77
Interpreting the Meaning of Information for Others	Interacting with others	0.02	3.04	0.61	1.01
Communicating with Supervisors, Peers, or Subordinates	Interacting with others	0.03	3.82	0.42	0.29
Communicating with Persons Outside Organisation	Interacting with others	0.03	3.10	0.80	0.82
Establishing and Maintaining Interpersonal Relationships	Interacting with others	0.03	3.53	0.51	0.84
Assisting and Caring for Others	Interacting with others	0.02	2.73	0.69	-0.80
Selling or Influencing Others	Interacting with others	0.02	2.41	0.68	0.31
Resolving Conflicts and Negotiating with Others	Interacting with others	0.02	2.83	0.58	0.70
Performing for or Working Directly with the Public	Interacting with others	0.03	2.86	0.96	0.00
Coordinating the Work and Activities of Others	Interacting with others	0.02	2.93	0.50	0.60
Developing and Building Teams	Interacting with others	0.02	2.81	0.50	0.41
Training and Teaching Others	Interacting with others	0.02	3.03	0.50	-0.81
Guiding, Directing, and Motivating Subordinates	Interacting with others	0.02	2.65	0.54	0.44
Coaching and Developing Others	Interacting with others	0.02	2.76	0.51	-0.21
Provide Consultation and Advice to Others	Interacting with others	0.02	2.60	0.58	1.02
Performing Administrative Activities	Interacting with others	0.02	2.65	0.64	0.87
Staffing Organisational Units	Interacting with others	0.02	1.93	0.53	0.77
Monitoring and Controlling Resources	Interacting with others	0.02	2.41	0.55	0.62

Note: summary statistics based on the task values across 326 occupations following the classification as defined in A.2. ONET Groups refers to the work activities groups as defined by ONET. Employment share is the average employment share in city employment as defined in Section 2.3. Task connectivity is defined in equation (2.6).

Table A.2. Variables

Variable	Definition	Year of measurement	Measurement	Source
Employment growth	MSAs employment growth 1990 - 2009	1990-2009	Standardised change in logs	Local Area Unemployment Statistics
Employment	MSAs employment	1990	Standardised log	Local Area Unemployment Statistics
Connectivity	MSAs average task connectivity, see equation (2.5)	1990	Standardised mean	CPS matched to ONET
Industrial specialisation	MSAs maximum over-representation of an industry, see equation (2.7)	1990	Standardised mean	CPS
Labour suitability	MSAs quality of the local labour pool relative to the industrial structure, see equation (2.9)	1990	Standardised mean	CPS
Social skills	MSAs share of social skills (ONET definition) in employment	1990	Standardised share	CPS matched to ONET
Routine tasks	MSAs ratio of routine task importance versus non-routine importance. Defined as in Acemoglu & Autor (2011)	1990	Standardised ratio	CPS matched to DOT
Computer use	MSAs average importance of computer use as defined in Section 2.4.3	1990	Standardised score	CPS matched to ONET
High-skilled	MSAs share of workers with at least a bachelor degree	1990	Standardised share	Census Decennial Database
Rent	MSAs mean rent	1990	Standardised rent	Census Decennial Database
January temperature	Average State January temperature	1990	Standardised temperature	Census
July temperature	Average State July temperature	1990	Standardised temperature	Census
Regional dummies	MSAs location dummy, defined as in the Census Regional Division		Dummy variables	Census Regional Division
Skill - connectivity	MSAs average connectivity of ONET Skills, see equation (2.6)	1990	Standardised mean	CPS matched to ONET
Co-agglomeration	MSAs average co-agglomeration of task employment, see equation (2.10)	1990	Standardised mean	CPS matched to ONET
HHI	MSAs score on the inverse Hirschman-Hefindahl index, see equation (2.11)	1990	Standardised score	CPS matched to ONET

Table A.3. Correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Employment growth 1990-2009	1.00															
(2) Employment	-0.09 (0.23)	1.00														
(3) Connectivity	0.02 (0.81)	0.88 (0.00)	1.00													
(4) Industrial specialisation	-0.11 (0.17)	-0.58 (0.00)	-0.59 (0.00)	1.00												
(5) Labour suitability	0.03 (0.69)	0.82 (0.00)	0.89 (0.00)	-0.54 (0.00)	1.00											
(6) Social skills	0.20 (0.01)	0.04 (0.62)	0.17 (0.03)	-0.15 (0.05)	0.11 (0.15)	1.00										
(7) Routine tasks	-0.16 (0.04)	0.36 (0.00)	0.31 (0.00)	-0.25 (0.00)	0.40 (0.00)	-0.30 (0.00)	1.00									
(8) Computer use	0.07 (0.36)	0.43 (0.00)	0.61 (0.00)	-0.40 (0.00)	0.38 (0.00)	0.11 (0.17)	-0.08 (0.27)	1.00								
(9) High skilled	0.16 (0.04)	0.35 (0.00)	0.48 (0.00)	-0.34 (0.00)	0.28 (0.00)	0.06 (0.45)	-0.10 (0.20)	0.69 (0.00)	1.00							
(10) Rent	-0.06 (0.45)	0.42 (0.00)	0.38 (0.00)	-0.38 (0.00)	0.26 (0.00)	0.01 (0.90)	-0.05 (0.55)	0.31 (0.00)	0.48 (0.00)	1.00						
(11) January temperature	0.34 (0.00)	-0.04 (0.59)	-0.08 (0.31)	-0.13 (0.09)	-0.02 (0.77)	0.16 (0.04)	-0.09 (0.22)	-0.13 (0.09)	-0.04 (0.62)	0.23 (0.00)	1.00					
(12) July temperature	0.21 (0.01)	-0.05 (0.56)	-0.09 (0.24)	-0.02 (0.85)	-0.03 (0.68)	0.12 (0.12)	-0.07 (0.40)	-0.15 (0.05)	-0.13 (0.08)	-0.15 (0.06)	0.72 (0.00)	1.00				
(13) North-east	-0.30 (0.00)	0.15 (0.06)	0.14 (0.07)	-0.07 (0.38)	0.10 (0.20)	-0.21 (0.01)	0.23 (0.00)	0.07 (0.40)	-0.02 (0.79)	0.12 (0.12)	-0.33 (0.00)	-0.32 (0.00)	1.00			
(14) Midwest	-0.36 (0.00)	-0.01 (0.92)	0.00 (0.96)	0.10 (0.20)	-0.04 (0.64)	-0.04 (0.58)	-0.05 (0.52)	0.03 (0.72)	-0.07 (0.40)	-0.24 (0.00)	-0.60 (0.00)	-0.31 (0.00)	-0.21 (0.01)	1.00		
(15) South	0.20 (0.01)	-0.10 (0.20)	-0.11 (0.16)	0.00 (0.99)	-0.04 (0.60)	0.11 (0.16)	0.02 (0.75)	-0.13 (0.09)	-0.11 (0.16)	-0.28 (0.00)	0.57 (0.00)	0.78 (0.00)	-0.30 (0.00)	-0.48 (0.00)	1.00	
(16) West	0.37 (0.00)	0.01 (0.89)	0.01 (0.85)	-0.05 (0.52)	0.01 (0.92)	0.08 (0.31)	-0.15 (0.05)	0.08 (0.33)	0.21 (0.01)	0.49 (0.00)	0.21 (0.01)	-0.35 (0.00)	-0.20 (0.01)	-0.31 (0.00)	-0.45 (0.00)	1.00

Note: P-values are in parentheses. n = 168 cities. All variables are measured in 1990. Definitions and sources of the variables are displayed in Table A.2.

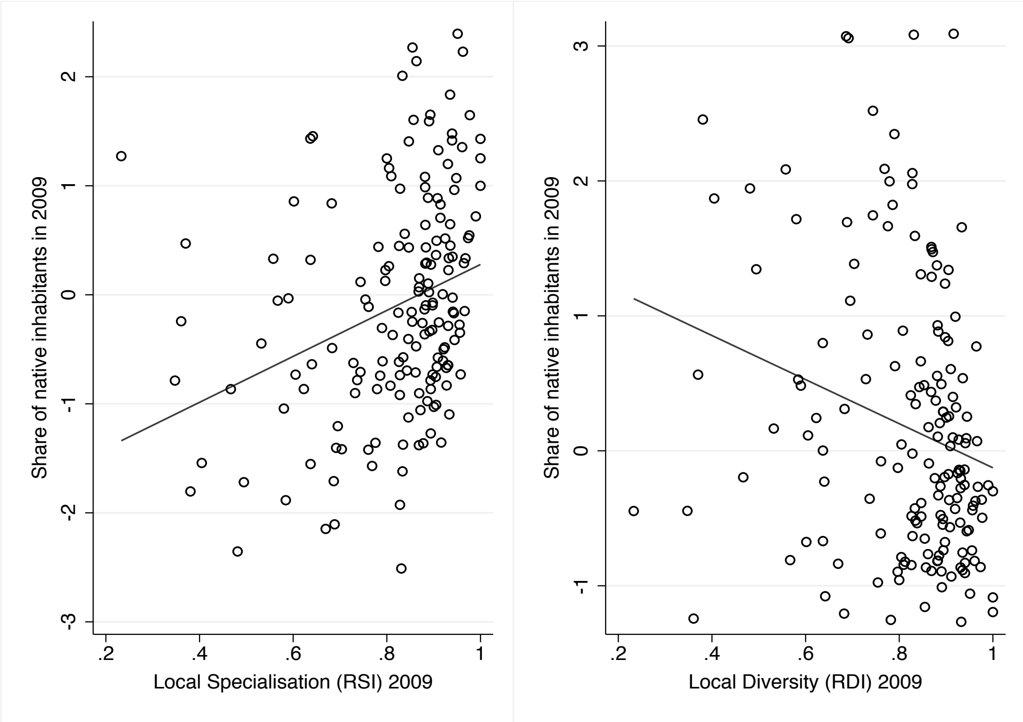
Table A.4. Regressions - task group combinations

	Information input	Work output	Mental processes
Work output	-0.049		
	[0.043]		
Mental process	0.065	0.047	
	[0.054]	[0.046]	
Interacting with others	0.032	0.035	-0.083
	[0.038]	[0.039]	[0.054]

Note: Regressions include initial employment share (1990), employment in both task groups separately and the control variables as well. Only the interaction term between employment in two task groups is presented. For instance, the cell Information input - Work output shows the coefficient of the interaction term of employment in information input and employment in work output tasks (both in 1990 while the regression furthermore includes the initial employment, the employment shares in information input and work output in 1990 and the control variables). Robust standard errors are in parentheses. The coefficients are insignificant.

A.4 Data appendix chapter 5

Figure A.1. Native inhabitants in specialised and diversified cities



Note: source Current Population Survey 2009. City level data, n=168. The correlations are respectively 0.30 (0.00) and -0.23 (0.00) and significant at the 1% level. RSI_l and RDI_l are measured as described in Section 5.3. Native inhabitants are defined as workers born in the US and are measured as share of employment.

Table A.5. Variables

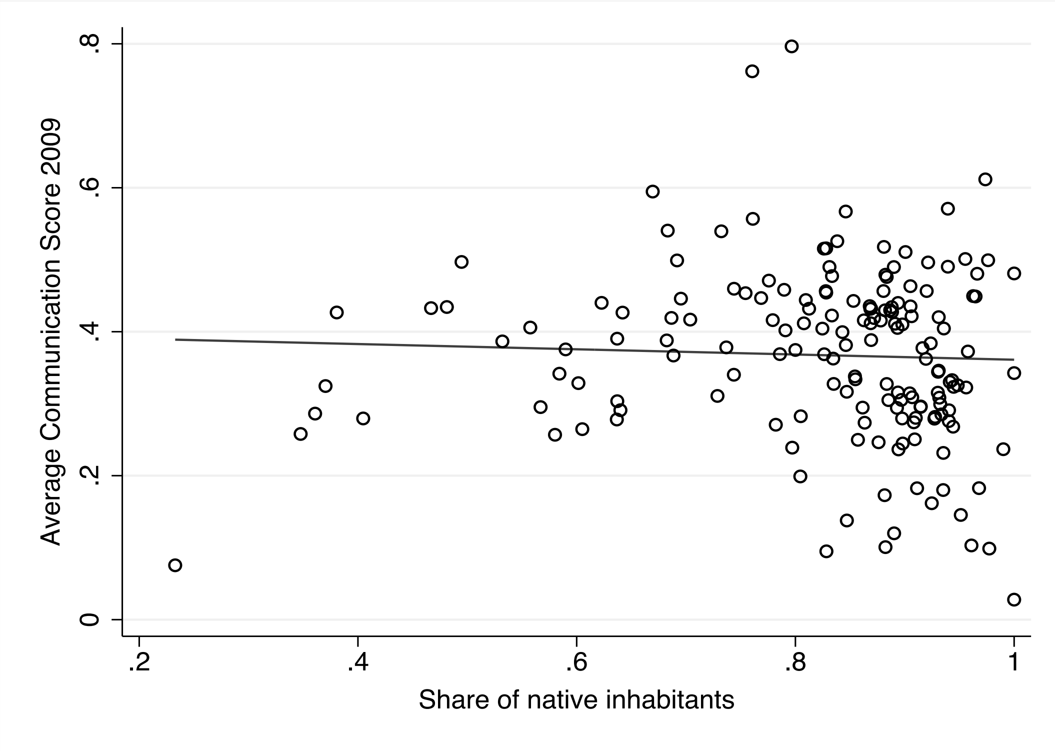
Variables	Description	Measurement	Source
Wage	Hourly wage	Individual level, logs Note: top coded as described in Section 5.4	Current Population Survey 2009
Communication	Principal component index by occupation Constructed by the standardised scores of the six communication tasks as described in Section 5.4	Occupational level, standardised scores	ONET Skill Survey 2000
Specialisation	Regional Specialisation Index by city $\hat{\rho}_l$ $\hat{\rho}_l = \max_l(\log E_{l,j} - log E_j)$ in which $E_{l,j}$ represents employment share of industry j in city l and E_j the employment share of industry j in national employment.	City level, standardised scores	Current Population Survey 2009
Diversity	Regional Diversity Index by city RDI_l $RDI_l = \frac{1}{\sum_j E_{l,j}/E_j}$	City level, standardised scores	Current Population Survey 2009
Dominant industry	Dummy variable indicating whether the individual works in the dominant local industry or not The dominant industry is the industry with the highest specialisation level in the city.	Individual level, dummy variable	Current Population Survey 2009
Control variables			
Drop-out	Dummy variables indicating whether the individual dropped-out of high-school	Individual level, dummy variable	Current Population Survey 2009
High-school	Dummy variable indicating whether the highest completed education of the individual was high-school	Individual level, dummy variable	Current Population Survey 2009
Some College (College)	Dummy variable indicating whether the highest completed education of the individual was some college	Individual level, dummy variable	Current Population Survey 2009
College (College grad)	Dummy variable indicating whether the highest completed education of the individual was college	Individual level, dummy variable	Current Population Survey 2009
Communication job activities	Standardised score on the ONET variable performing for or working directly with the public'.	Occupational level, standardised scores	ONET Skill Survey 2000
Non-white	Race measurement, when the individual originates from a non-white race the dummy equals unity.	Individual level, dummy variable	Current Population Survey 2009
Non-married	When the individual is not married, the dummy equals unity	Individual level, dummy variable	Current Population Survey 2009
Age and age squared	Age and age squared of the individual	Individual level	Current Population Survey 2009
Female	When the individual is a female, the dummy equals unity		Current Population Survey 2009
Occupation dummies	Dummy variables for each two-digit occupation group	Occupational level, dummy variables	Current Population Survey 2009
Additional / robustness variables			
Size	Employment by MSA	City level, standardised logs	Local Unemployment Figures 2009
Language-skill proxy	Average score on the following category: Who originates from a non-English speaking country? Category 1: the worker him/herself Category 2: both parents of the worker Category 3: one of the parents of the worker Category 4: nobody	Occupational level, standardised shares	Current Population Survey 2009
Population 1930	County population in 1930, summed by MSA	City level, standardised logs	Census Historical Population Figures
Relative communication	Share of communication job tasks within the total score of job tasks by occupation		ONET Skill Survey 2000
Non-routine interactive	Occupational score on the non-routine interactive job tasks as defined in Acemoglu & Autor (2011)	Occupational level, score	ONET Skill Survey 2000
Rent	Standardised average rent by MSA in 2000	City level, standardised averages	Census 2000

Table A.6. Correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)d	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
(1) Wage	1.00																				
(2) Communication	0.40	1.00																			
	(0.00)																				
(3) Specialisation	-0.11	-0.04	1.00																		
	(0.00)	(0.00)																			
(4) Diversity	0.03	0.01	-0.45	1.00																	
	(0.00)	(0.00)	(0.00)																		
(5) Dominant industry	0.04	-0.03	0.00	-0.09	1.00																
	(0.00)	(0.00)	(0.53)	(0.00)																	
(6) Drop-out	-0.26	-0.30	-0.00	-0.00	0.01	1.00															
	(0.00)	(0.00)	(0.68)	(0.52)	(0.00)																
(7)High-school	-0.22	-0.27	0.04	-0.01	-0.01	-0.17	1.00														
	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)															
(8) College	-0.12	-0.01	0.06	-0.01	-0.02	-0.19	-0.38	1.00													
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)														
(9) College grad	0.46	0.42	-0.09	0.02	0.01	-0.22	-0.46	-0.49	1.00												
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)													
(10) Communication job	-0.09	0.32	-0.01	0.01	-0.04	-0.08	-0.08	0.05	0.08	1.00											
	(0.00)	(0.00)	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)												
(11) Non-white	-0.07	-0.04	-0.13	-0.01	0.00	-0.02	0.02	0.01	-0.02	0.02	1.00										
	(0.00)	(0.00)	(0.00)	(0.00)	(0.41)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)											
(12) Non-married	-0.24	-0.11	-0.02	-0.01	-0.01	0.05	0.03	0.06	-0.12	0.07	0.08	1.00									
	(0.00)	(0.00)	(0.00)	(0.05)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)										
(13) Age	0.29	0.10	0.00	0.00	0.02	-0.09	0.01	-0.05	0.09	-0.10	-0.02	-0.34	1.00								
	(0.00)	(0.00)	(0.17)	(0.20)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)									
(14) Age squared	0.26	0.09	0.01	0.00	0.02	-0.08	0.02	-0.04	0.07	-0.08	-0.03	-0.30	0.99	1.00							
	(0.00)	(0.00)	(0.01)	(0.36)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)								
(15) Female	-0.16	0.13	0.00	-0.00	-0.03	-0.05	-0.03	0.05	0.01	0.24	0.05	0.06	0.02	0.02	1.00						
	(0.00)	(0.00)	(0.37)	(0.78)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)							
(16) Size	0.10	0.03	-0.66	0.57	0.01	0.03	-0.04	-0.05	0.07	-0.00	0.07	0.02	-0.01	-0.01	-0.01	1.00					
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.41)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)						
(17) Communication proxy	0.27	0.67	-0.01	0.00	-0.03	-0.31	-0.16	0.06	0.26	0.27	-0.06	-0.07	0.08	0.07	0.11	-0.00	1.00				
	(0.00)	(0.00)	(0.04)	(0.23)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.74)					
(18) Population 1930	0.06	0.02	-0.51	0.61	-0.05	0.01	-0.02	-0.04	0.05	-0.01	0.03	0.00	-0.00	-0.00	-0.00	0.72	0.00	1.00			
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.33)	(0.82)	(0.47)	(0.64)	(0.00)	(0.69)				
(19) Relative communication	0.11	0.52	-0.03	0.01	-0.01	-0.13	-0.07	-0.01	0.15	0.34	-0.06	-0.02	0.03	0.03	0.04	0.03	0.45	0.01	1.00		
	(0.00)	(0.00)	(0.00)	(0.06)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)			
(20) Non-routine Index	0.35	0.76	-0.02	0.00	-0.01	-0.19	-0.22	-0.07	0.38	0.13	-0.05	-0.12	0.11	0.09	0.04	0.02	0.42	0.01	0.29	1.00	
	(0.00)	(0.00)	(0.00)	(0.16)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
(21) Rent	0.12	0.02	-0.36	0.13	0.00	0.00	-0.06	-0.03	0.08	-0.00	0.15	0.02	-0.00	-0.00	-0.01	0.27	-0.01	0.14	0.02	0.01	1.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.24)	(0.75)	(0.00)	(0.00)	(0.00)	(0.84)	(0.00)	(0.00)	(0.86)	(0.30)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	

Note: n=62,705. P-values are in parentheses. See Table A.5 for a detailed description of the variables, measurement and data sources.

Figure A.2. Communication and native inhabitants



Note: source Current Population Survey 2009. City level data, n=168. The correlation is -0.08 (0.34) and not significant. Communication is measured as the average score on the Communication-Index as defined in Section 5.4. Native inhabitants are defined as workers born in the US and are measured as share of employment.

Table A.7. Correlations among communication tasks

Variables	(1)	(2)	(3)	(4)	(5)	(6)
(1) Relations	1.000					
(2) External communication	0.800	1.000				
(3) Internal communication	0.658	0.603	1.000			
(4) Face-to-face	0.479	0.447	0.500	1.000		
(5) Teamwork	0.420	0.332	0.512	0.544	1.000	
(6) Contact	0.579	0.522	0.308	0.472	0.535	1.000

Table A.8. PCA results for communication tasks

	Communication-Index loadings for first principal component
Relations	0.456
External communication	0.429
Internal communication	0.416
Face-to-face	0.386
Teamwork	0.371
Contact	0.386
Explained variance	0.599

APPENDIX B

GERMAN DATA

B.1 Data description

The empirical analyses in Chapter 3 employ the the survey of the working population in Germany carried out by the German Federal Institute for Vocational Training (BIBB) and the Federal Institute for Occupational Safety and Health (BAuA).[1] Since 1979, the BIBB survey questions the German labour force about qualifications, career history and detailed job characteristics etc. Chapter 3 employs the most recent wave of the BIBB; the 2006 wave. This wave consists of a representative sample of about 20,000 Germans. For more information about the survey and the dataset we refer to the work of Rohrbach-Schmidt (n.d.).

[1] Hereafter we refer to this dataset as the BIBB dataset.

Table B.1. Included tasks

Job tasks	Cognitive skills	Task requirements
Manufacturing of goods	Natural scientific	Deadline pressure
Measuring, testing, quality control	Manual / craft	Work is stipulated in the minutest details
Operating, controlling machines	Pedagogic	One and the same work cycle is repeated
Repairing	Legal	Confronted with new problems that remain to be understood
Purchasing, selling	Project management	Process optimisation / trying out new things
Transporting, storing, shipping	Medical or custodial	Are you failed or disturbed
Promoting, marketing, public relations	Lay-outing, designing, visualising	Required output is stipulated in the minutest details
Organising, making plans, working out operations	Mathematical, statistical	Doing things you haven't learned before
Research, development	German language	Simultaneously keep an eye on diverse processes or tasks
Teaching, training	Computer application software	Do very small mistakes lead to big financial losses
Gathering information, investigating, documenting	Technical	Reaching the limits of your capacities
Consulting, advising	Foreign language	Need to work very quick
Entertaining, accommodating, preparing food		
Taking care, healing		
Protecting, guarding, observing, controlling traffic		
Working with computers		
Job Characteristics	Specialised Skills	
Having to react to and solving unforeseeable problems	Finance	
Notifying / communicating difficult issue in an intelligible to all way	Book-keeping	
Convincing others, compromising	Fiscal	
Making tough choice on your own responsibility	Accounting	
Recognising and closing own knowledge gaps	Credit system	
Speech-making, giving talks	Controlling	
Having contact to customers, clients, patients	Sales	
Dealing with a range of duties and responsibilities	Business administration	
Being responsible for the well-being of other		

Table B.2. List of included variables

	Measurement	Mean	S.D.
Specialisation	Number of tasks that are performed ’sometimes or ’rarely’ by the worker	15.61	57.89
Required cognitive skills	Number of performed cognitive core tasks - as defined in Section 3.3.2	1.66	1.09
Small city	Dummy variables: city of residence houses less than 20,000 inhabitants	0.41	0.49
Medium city	City of residence houses between 20,000 and 100,000 inhabitants	0.26	0.44
Large city	City of residence houses more than 100,000 inhabitants	0.33	0.47
Unskilled	No degree	0.03	0.18
Low skilled	Obtained high-school degree	0.04	0.19
Medium skilled	Obtained operational college degree	0.58	0.49
High skilled	Obtained college or university degree	0.35	0.48
Age	Age of the individual	42.45	9.54
Gender	Dummy variable with value 0 for males and 1 for females	0.49	0.50
Native speaker	Dummy variable indicating whether German is the workers mother tongue	0.94	0.23
Job	Three-digit occupation and two-digit industry combination This results in 1739 unique jobs.		
Job average	The mean of the dependent variable for the for the occupation-industry combination.		

Table B.3. Correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Specialisation	1.00											
(2) Cognitive tasks	-0.35	1.00										
	(0.00)											
(3) Small city	0.02	-0.06	1.00									
	(0.04)	(0.00)										
(4) Medium city	-0.01	-0.01	-0.50	1.00								
	(0.33)	(0.44)	(0.00)									
(5) Large city	-0.01	0.07	-0.58	-0.41	1.00							
	(0.21)	(0.00)	(0.00)	(0.00)								
(6) Unskilled	-0.12	-0.00	0.02	0.00	-0.03	1.00						
	(0.00)	(0.66)	(0.01)	(0.59)	(0.00)							
(7) Low skilled	-0.04	0.01	-0.04	-0.01	0.05	-0.04	1.00					
	(0.00)	(0.14)	(0.00)	(0.15)	(0.00)	(0.00)						
(8) Medium skilled	-0.04	-0.15	0.09	0.02	-0.11	-0.21	-0.23	1.00				
	(0.00)	(0.00)	(0.00)	(0.06)	(0.00)	(0.00)	(0.00)					
(9) High skilled	0.10	0.16	-0.08	-0.01	0.10	-0.14	-0.15	-0.86	1.00			
	(0.00)	(0.00)	(0.00)	(0.11)	(0.00)	(0.00)	(0.00)	(0.00)				
(10) Age	-0.05	-0.04	0.05	0.00	-0.05	0.05	-0.05	-0.07	0.08	1.00		
	(0.00)	(0.00)	(0.00)	(0.85)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)			
(11) Female	-0.10	0.13	-0.01	-0.00	0.01	0.03	0.02	0.08	-0.10	-0.01	1.00	
	(0.00)	(0.00)	(0.39)	(0.83)	(0.27)	(0.00)	(0.04)	(0.00)	(0.00)	(0.32)		
(12) German	0.03	0.01	0.06	-0.02	-0.05	-0.08	-0.07	0.03	0.02	0.08	0.00	1.00
	(0.00)	(0.23)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.69)	

Note: n = 15,670. Table B.2 displays the definitions of the variables. P-values are in parentheses.

B.2 Replication estimates of Duranton & Jayet (2011) for Germany

Following Duranton and Jayet, Section 3.3.3 analyses whether scarce occupations are more often performed in large cities. Ideally we estimate the employment share for each sector j, city size l and occupation o combination $E^j_{o,l}$:

$$E^j_{l,o} = a_0 + a_1(1/E^j_o) + a_2 N_l + a_3 N^j_l + a_4(1/E^j_o) * N_l + \epsilon^j_{l,o} \tag{B.1}$$

in which E^j_o represents the average employment share of occupation o in sector j. N_l is a city size dummy and N^j_l is a dummy for each city category and sector combination. However, there are too many zeros in the data to estimate this regression. Therefore, we use fixed effects and dummies for each sector and occupation combination (a^j_o), for each sector and city size combination (b^j_l) and for each city size and scarcity level combination ($d_{m(l),r(j,o)}$). Scarcity is defined as the scarcity of occupation o within sector j, we measure this in terms of quartiles.

$$E^j_{l,o} = a^j_o + b^j_l + d_{m(l),r(j,o)} + \epsilon^j_{l,o} \tag{B.2}$$

To make this estimation computationally tractable, we focus on the probability of an individual to end up in each of these cells. We assume this probability follows a logit form:

$$\pi^j_{l,o} = \frac{exp(Y^j_{l,o})}{\sum_{i=1,...L} \sum_{l=1,...O} exp(Y^j_{i,l})} \tag{B.3}$$

with: $Y^j_{l,o} = \alpha^j_l + \beta^j_o + \xi_{m(l),r(j,o)}$
For more detailed information we refer to the work of Duranton & Jayet (2011).

APPENDIX C

DUTCH DATA

C.1 Data description

Chapter 4 employs the Longitudinal Internet Studies for the Social Sciences (LISS) panel of 3,000 Dutch individuals. This panel is the core element of a project titled 'Measurement and Experimentation in the Social Sciences' from the Dutch research institute CentERdata. The chapter combines information from the background study, the work and schooling study, the personality study and an additional questionnaire about job tasks (carried out in May 2012). We drop all skilled agricultural, fishery, and forestry workers, since the locations of these occupations depend on natural resources.

The website of the LISS panel (http://www.lissdata.nl/) provides detailed information about the panel and questionnaires and provides access to the data.

Table C.1. List of variables

	Dependent variables		Mean	SD
Matching all skills	How do your knowledge and skills suit your work?	Standardised	0.00	1.00
Matching cognitive skills	Inverse of the difference between standardised cognitive job tasks and standardised cognitive skills.	Standardised	0.00	1.00
Matching social skills	Inverse of the difference between standardised social job tasks. and standardised social skills. See Table 4.2 for the social skills.	Standardised	0.00	1.00
Cognitive skills	Number of cognitive statements with which the worker agrees or strongly agrees.	Standardised	0.00	1.00
Social skills	Number of cognitive statements with which the worker agrees or strongly agrees, See Table 4.2 for cognitive and social statements.	Standardised	0.00	1.00
Cognitive tasks	Number of core cognitive tasks the worker performs.	Standardised	0.00	1.00
Social tasks	Number of core social tasks the workers performs, See Table 4.2 for cognitive and social job tasks.	Standardised	0.00	1.00
Gross monthly wage	Personal gross monthly income in euros.	Standardised	0.00	1.00
	Explanatory variables			
City	Dummy variable indicating whether the worker works in a city with at least 1,500 dwellings per square kilometre.	Dummy	0.45	0.50
Age	Age of the worker.	In years	44.69	12.17
Female	Dummy variable indicating whether the worker is male or female.	Dummy	1.53	0.50
Native	Dummy variable indicating whether the worker is native Dutch or not.	Dummy	0.91	0.29
Skill dummies	Indicates the worker's highest diploma obtained.			
Low skilled	Worker did not obtain a diploma.	Dummy	0.20	0.40
Medium skilled	Intermediate vocational education diploma. or high school diploma at the pre-university level.	Dummy	0.37	0.48
High skilled	Higher vocational education or university diploma.	Dummy	0.38	0.49
Occupation	One-digit ISCO occupations	Categorised 1-10		

Note: all standardised variables have a mean of zero and a standard deviation of one. We indicate gross monthly wages as missing if a person earns nothing, less than nothing, or more than 10,000 euros a month.

Table C.2. Correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(1) Age	1.00														
(2) Female	-0.10	1.00													
	(0.00)														
(3) Native	0.02	0.02	1.00												
	(0.23)	(0.37)													
(4) Low skilled	0.15	0.01	0.00	1.00											
	(0.00)	(0.77)	(0.90)												
(5) Medium skilled	-0.13	0.04	-0.03	-0.38	1.00										
	(0.00)	(0.04)	(0.09)	(0.00)											
(6) High skilled	-0.01	-0.04	0.05	-0.39	-0.60	1.00									
	(0.47)	(0.04)	(0.02)	(0.00)	(0.00)										
(7) City	0.02	-0.02	-0.06	-0.12	-0.05	0.14	1.00								
	(0.35)	(0.25)	(0.00)	(0.00)	(0.01)	(0.00)									
(8) Match quality	0.15	-0.05	0.06	-0.07	-0.10	0.15	0.09	1.00							
	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)								
(9) Match quality, cognitive	-0.01	-0.06	-0.03	-0.06	-0.07	0.11	0.06	0.03	1.00						
	(0.68)	(0.01)	(0.26)	(0.02)	(0.01)	(0.00)	(0.01)	(0.32)							
(10) Match quality, social	-0.01	-0.02	-0.05	0.03	-0.04	0.02	0.01	0.01	0.10	1.00					
	(0.82)	(0.44)	(0.05)	(0.28)	(0.15)	(0.39)	(0.61)	(0.70)	(0.00)						
(11) Cognitive skills	0.08	-0.17	0.00	-0.16	-0.11	0.26	0.12	0.15	0.29	0.08	1.00				
	(0.00)	(0.00)	(0.88)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)					
(12) Social skills	-0.02	0.08	-0.05	-0.02	-0.05	0.07	0.07	0.09	0.06	0.32	0.28	1.00			
	(0.38)	(0.00)	(0.03)	(0.41)	(0.03)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)				
(13) Cognitive tasks	-0.02	-0.16	0.01	-0.21	-0.07	0.24	0.10	0.22	0.26	0.07	0.30	0.15	1.00		
	(0.40)	(0.00)	(0.82)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)			
(14) Social tasks	0.01	0.01	0.04	-0.15	-0.08	0.21	0.09	0.22	0.13	0.10	0.23	0.20	0.52	1.00	
	(0.60)	(0.83)	(0.10)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
(15) Gross monthly earnings	0.19	-0.45	0.01	-0.22	-0.20	0.36	0.13	0.32	0.08	-0.02	0.25	-0.01	0.36	0.26	1.00
	(0.00)	(0.00)	(0.68)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.57)	(0.00)	(0.79)	(0.00)	(0.00)	

Note: the p-values are in parentheses. The definitions and measurement of the variables are displayed in Table C.1.

C.2 Proxy measurement error

Respondents were asked to indicate the importance of a certain job task for an example job followed by the effectiveness at performing that task in that occupation. For two separate tasks, the respondent was questioned about the importance and effectiveness of two example jobs. Table C.3 shows the task–occupation combinations.

The proxy for the measurement error is a respondent's indicated importance for a task–occupation combination relative to its average indicated importance.

Table C.3. Task–occupation combinations for example jobs

Task	Example jobs	
Dealing with people	Secretary	Car mechanic
Persuading/influencing others	Nurse	Teacher
Physical strength	Grocer	Policeman
Dexterity	Plumber	Salesperson
Solving problems	Ticket collector	Journalist
Simple mathematics	Cashier	Real estate agent

Bibliography

Acemoglu, D. & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *In Handbook for Labour Economics Volume 4, Part B (Amsterdam: Elsevier)*, 1043-1171.

Akcomak, I., Borghans, L. & Ter Weel, B. (2011). Measuring and interpreting trends in the division of labour in the Netherlands. *De Economist*, *159 (4)*, 435-482.

Akcomak, I., Kok, S. & Rojas-Romagosa, H. (2013). Technology, offshoring and the task-conent of occupations: Evidence from the United Kingdom. *CPB Discussion Paper*, *No. 233*.

Autor, D. (2013). The 'task approach' to labor markets: An overview. *NBER Working Paper*, *No. 18711*.

Autor, D. & Dorn, D. (2009). This job is 'getting old': Measuring changes in job opportunities using occupational age structure. *American Economic Review Papers and Proceedings*, *99 (2)*, 45-51.

Autor, D. & Dorn, D. (forthcoming). The growth of low-skill service jobs and the polarization of the U.S. labor market. *American Economic Review*.

Autor, D. & Handel, M. (forthcoming). Putting tasks to the test: Human capital, job tasks and wages. *Journal of Labor Economics*.

Autor, D., Katz, L. & Kearney, M. (2006). The polarization of the U.S. labor market. *American Economic Review*, *96 (2)*, 189-194.

Autor, D., Katz, L. & Krueger, A. (1998). Computing inequality: Have computers changed the labor market? *Quarterly Journal of Economics*, *113 (4)*, 1169-1214.

Autor, D., Levy, F. & Murnane, R. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, *118 (4)*, 1279-1334.

Bacolod, M., Blum, B. & Strange, W. (2009). Skills in the city. *Journal of Urban Economics*, *65 (2)*, 136-153.

Bacolod, M., Blum, B. & Strange, W. (2010). Elements of skill: Traits, intelligences, education, and agglomeration. *Journal of Regional Science*, *50 (1)*, 245-280.

Baldwin, R. (2010). Integration of the North American economy and new paradigm globalisation. *CEPR Discussion Paper*, *No. 7523*.

Baldwin, R. & Evenett, S. (2012). The UK in a global world. How can the UK focus on steps in global value chains that really add value? In D. Greenaway (Ed.), (p. 71-128). Centre for Economic Policy Research.

Baldwin, R. & Nicoud, F. (2010). Trade-in-goods and trade-in-tasks: An integrating framework. *NBER Working Paper*, *No. 15882*.

Baumgardner, J. (1988a). The division of labor, local markets, and worker organization. *Journal of Political Economy*, *96*, 509-527.

Baumgardner, J. (1988b). Physicians' services and the division of labor across local markets. *Journal of Political Economy*, *96*, 948-982.

Becker, G. & Murphy, K. (1992). The division of labor, coordination costs and knowledge. *Quarterly Journal of Economics*, *107 (4)*, 1137-1160.

Berry, C. & Glaeser, E. (2005). The divergence of human capital levels across cities. *Papers in Regional Science*, *84 (3)*, 407-444.

Besley, T., Persson, T. & Sturm, D. (2010). Political competition, policy and growth: Theory and evidence from the US. *Review of Economic Studies*, *77 (4)*, 1329-1352.

Blinder, A. (2006). Offshoring: The next industrial revolution. *Foreign Affairs*, *85*, 113-128.

Bloom, N., Garicano, L., Sadun, R. & Van Reenen, J. (2009). The distinct effects of information technology and communication technology on firm organization. *NBER working paper*, *No. 14975*.

Blum, B. & Goldfarb, A. (2006). Does the internet defy the law of gravity? *Journal of International Economics*, *70 (2)*, 384-405.

Borghans, L. & Ter Weel, B. (2004). What happens when agent T gets a computer? The labor market impact of cost efficient computer adoption. *Journal of Economic Behavior & Organization*, *54 (2)*, 137-151.

Borghans, L. & Ter Weel, B. (2006). The division of labour, worker organisation, and technological change. *Economic Journal*, *116 (509)*, F45-F72.

Borghans, L., Ter Weel, B. & Weinberg, B. (2006). People people: Social capital and the labor-market outcomes of underrepresented groups. *NBER Working Paper*, *No. 11985*.

Borghans, L., Ter Weel, B. & Weinberg, B. (2008). Interpersonal styles and labor market outcomes. *Journal of Human Resources*, *43 (4)*, 815-858.

Bresnahan, T. (1999). Computerisation and wage dispersion: An analytical reinterpretation. *Economic Journal*, *109 (456)*, 390-415.

Bresnahan, T., Brynjolfsson, E. & Hitt, L. (2002). Information technology, workplace organization and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics*, *117 (1)*, 339-376.

Briant, A., Combes, P. & Lafourcade, M. (2008). Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *CEPR Discussion Paper*, *No. 6928*.

Burdett, K. & Coles, M. (1997). Marriage and class. *Quarterly Journal of Economics*, *112*, 141-168.

Burrows, P. (1995). The global chip payoff. *Business Week*, *Augustus 7*.

Cairncross, F. (1997). *The death of distance. How the communications revolution will change our lives*. Harvard Business School Press, Boston, Massachusetts.

Caselli, F. & Coleman, W. (2001). Cross-country technology diffusion: The case of computers. *American Economic Review*, *91 (2)*, 328-335.

Charlot, S. & Duranton, G. (2004). Communication externalities in cities. *Journal of Urban Economics*, *56 (3)*, 581-613.

Ciccone, A. & Hall, R. (1996). Productivity and the density of economic activity. *American Economic Review*, *86 (1)*, 54-70.

Combes, P., Duranton, G. & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, *63 (2)*, 723-742.

Combes, P., Duranton, G. & Gobillon, L. (2009). The economics of agglomeration. In E. Glaeser (Ed.), (p. 15-65). Chicago and London: University of Chicago Press.

Combes, P., Duranton, G., Gobillon, L. & Roux, S. (2012). Sorting and the local wage and skill distributions in France. *Regional Science and Urban Economics*, *42 (6)*, 913-930.

Costa, D. & Kahn, M. (2001). Power couples. *Quarterly Journal of Economics*, *116*, 1287-1315.

Criscualo, C. & Garicano, L. (2010). Offshoring and wage inequality: Using occupational licensing as a shifter of offshoring costs. *American Economic Review*, *100 (2)*, 439-443.

Desmet, K. & Rossi-Hansberg, E. (2009). Spatial growth and industrial age. *Journal of Economic Theory*, *144 (6)*, 2477-2502.

Duranton, G. & Jayet, H. (2011). Is the division of labour limited by the extent of the market? Evidence from French cities. *Journal of Urban Economics*, *69 (1)*, 56-71.

Duranton, G. & Puga, D. (2000). Diversity and specialisation in cities: Why, where and when does it matter? *Urban Studies*, *37 (3)*, 533-555.

Duranton, G. & Puga, D. (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review*, *91 (5)*, 1454-1477.

Duranton, G. & Puga, D. (2004). Micro-foundations of urban agglomeration economies. *In Handbook of Regional and Urban Economics, Volume 4 (Amsterdam: Elsevier)*, 2063-2117.

Duranton, G. & Puga, D. (2005). From sectoral to functional urban specialisation. *Journal of Urban Economics*, *57 (2)*, 343-370.

Dustmann, C., Ludsteck, J. & Schnberg, U. (2009). Revisting the German wage structure. *Quarterly Journal of Economics*, *124 (2)*, 809-842.

Eeckhout, J., Pinheiro, R. & Schmidheiny, K. (2010). Spatial sorting: Why New York, Los Angeles and Detroit attract the greatest minds as well as the unskilled. *CEPR Discussion Paper*, *No. 8151*.

Ellison, G. & Glaeser, E. (1997). Geographic concentration in US manufacturing industries: A dartboard approach. *Journal of Political Economy*, *105 (5)*, 889-927.

Ellison, G. & Glaeser, E. (1999). The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review*, *89 (2)*, 311-316.

Ellison, G., Glaeser, E. & Kerr, W. (2010). What causes industry agglomeration? Evidence of coagglomeration patterns. *American Economic Review*, *100 (3)*, 1195-1213.

Elvery, J. (2010). City size and skill intensity. *Regional Science and Urban Economics*, *40*, 367-379.

Feenstra, R. (2010). *Offshoring in the global economy. Microeconomic structure and macroeconomic implications*. MIT Press, Cambridge Mass.

Feldman, M. & Audretsch, D. (1999). Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review*, *42 (2)*, 409-429.

Firpo, S., Fortin, N. & Lemieux, T. (2009). Occupational tasks and changes in the wage structure. *IZA Discussion Paper, University of British Columbia*, *No. 5542*.

Florida, R., Mellander, C., Stolarick, K. & Ross, A. (2012). Cities, skills and wages. *Journal of Economic Geography*, *12 (2)*, 355-377.

Friedman, T. (2005). *The world is flat. A brief history of the twenty-first century*. Farrar, Straus & Giroux, New York.

Garicano, L. & Hubbard, T. (2009). Specialization, firms and markets: The division of labour between and within law firms. *Journal of Law, Economics and Organisation*, *25*, 339-371.

Garicano, L. & Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy. *Quarterly Journal of Economics*, *121 (4)*, 1383-1435.

Gaspar, J. & Glaeser, E. (1998). Information technology and the future of cities. *Journal of Urban Economics*, *43 (1)*, 136-156.

Gathmann, C. & Schnberg, U. (2010). How general is human capital? A task-based approach. *Journal of Labor Economics*, *28 (1)*, 1-49.

Gautier, P., Svarer, M. & Teulings, C. (2010). Marriage and the city: Search frictions and sorting of singles. *Journal of Urban Economics*, *67*, 206-218.

Glaeser, E. & Gottlieb, J. (2006). Urban resurgence and the consumer city. *Urban Studies*, *43 (8)*, 1275-1299.

Glaeser, E. & Gottlieb, J. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature*, *47 (4)*, 983-1028.

Glaeser, E., Kallal, H., Scheinkman, J. & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, *100 (6)*, 1126-1152.

Glaeser, E. & Kerr, W. (2009). Local industrial conditions and entrepreneurship: How much of the spatial distribution can we explain? *Journal of Economics and Management Strategy*, *18 (3)*, 623-663.

Glaeser, E., Kolko, J. & Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, *1*, 27-50.

Glaeser, E. & Maré, D. (2001). Cities and skills. *Journal of Labor Economics*, *19 (2)*, 316-342.

Glaeser, E. & Ponzetto, A. (2010). Agglomeration economics. In E. Glaeser (Ed.), (p. 303 - 337). The University of Chicago Press.

Glaeser, E., Ponzetto, A. & Tobio, K. (2012). Cities, skills and regional change. *Regional Studies*, *iFirst article*.

Glaeser, E. & Ressenger, M. (2010). The complementarity between cities and skills. *Journal of Regional Science*, *50 (1)*, 221-244.

Glaeser, E., Schienkman, J. & Shleifer, A. (1995). Economic growth in a cross-ection of cities. *Journal of Monetary Economics*, *36 (1)*, 117-143.

Glaeser, E. & Shapiro, J. (2003). Urban growth in the 1990s: Is city living back? *Journal of Regional Science*, *43 (1)*, 139-165.

Goos, M. & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *Review of Economics and Statistics*, *89 (1)*, 118-133.

Goos, M., Manning, A. & Salomons, A. (2009). Job polarization in Europe. *American Economic Review: Papers and Proceedings*, *99 (2)*, 58-63.

Grossman, G. & Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring. *American Economic Review*, *98 (5)*, 1978-1997.

Harrison, B., Kelley, M. & Gant, J. (1996). Specialization versus diversity in local economies: The implications for innovative private-sector behavior. *Cityscape: A Journal of Political Development and Research*, *2 (2)*, 61-93.

Heckman, J., Stixrud, J. & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24 (3)*, 411-482.

Helsley, R. & Strange, W. (1990). Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics, 20 (2)*, 189-212.

Henderson, J., Kuncoro, A. & Turner, M. (1995). Industrial development in cities. *Journal of Political Economy, 103 (5)*, 1067-1090.

Hill, D. (2011). Population growth and the rise of mega-cities: Can technology help? *Singularityhub 'http://singularityhub.com/', June 23.*

Ioannides, Y., Overman, H., E., R.-H. & Schmidheiny, K. (2008). The effect of information and communication technologies on urban structure. *Economic Policy, 23 (54)*, 201-242.

Jacobs, J. (1969). *The economy of cities.* Random house, New York.

Jaffe, A., Trajtenberg, M. & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidence by patent citations. *Quarterly Journal of Economics, 108 (3)*, 577-598.

Jensen, B. & Kletzer, L. (2005). Tradable services: Understanding the scope and impact of services offshoring. *Institute for International Economics Working Paper, No. 05-9.*

Katz, L. (2000). Understanding the digital economy: Data, tools and research. In E. Brynjolfsson & B. Kahin (Eds.), (p. 217-244). MIT Press, Cambridge Mass.

Kelley, M. & Helper, S. (1999). Firms size and capabilities, regional agglomeration, and the adoption of new technology. *Economics of Innovation and New Technology, 8 (2)*, 79-103.

Kim, S. (1990). Labor heterogeneity, wage bargaining, and agglomeration economies. *Journal of Urban Economics, 28 (2)*, 160-177.

Kim, S. (1991). Heterogeneity of labor market and city size in an open spatial economy. *Regional Science and Urban Economics, 21 (1)*, 109-126.

Kok, S. (forthcoming). Applied modelling of regional growth and innovation systems. In K. Kourtit, P. Nijkamp & B. Stimson (Eds.), (chap. Returns to communication in specialised and diversified US cities). Springer-Verlag.

Lambooy, B. (1998). Polynucleation and economic development: The Randstad. *European Planning Studies*, *4*, 456-466.

Lazear, E. (2009). Firm-specific human capital: A skill-weights approach. *Journal of Political Economy*, *117 (5)*, 914-940.

Leamer, E. (2007). A flat world, a level playing field, a small world after all, or none of the above? *Journal of Economic Literature*, *155*, 83-126.

Lee, S. (2010). Ability sorting and consumer city. *Journal of Urban Economics*, *68*, 20-33.

Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, *96 (3)*, 461-498.

Lewis, E. (2011). Immigrant-native substitutability: The role of language ability. *NBER Working Paper*, *No: 17609*.

Lucas, R. J. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22 (1)*, 3-42.

Marshall, A. (1920). *Principles of economics*. MacMillan, London.

Moretti, E. (2004). Workers' education, spillovers and productivity : Evidence from plant-level production functions. *American Economic Review*, *94 (3)*, 656-691.

Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, *5 (1)*, 65-103.

Moulton, B. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro-units. *The Review of Economics and Statistics*, *72 (2)*, 334-338.

Mouw, T. & Kallenberg, A. (2010). Occupations and the structure of wage inequality in the United States, 1980 to 2000s. *American Sociological Review*, *75 (3)*, 402-431.

Mueser, P. & Graves, P. (1995). Examining the role of economic opportunity and amenities in explaining population redistribution. *Journal of Urban Economics*, *37 (2)*, 176-200.

Patuelli, R., Reggiani, A., Nijkamp, P. & Bade, F. (2010). The evolution of the commuting network in Germany. *Journal of Transport and Land Use*, *2*, 5-37.

Petrongolo, B. & Pissarides, C. (2001). Looking into the black-box: A survey of the matching function. *Journal of Economic Literature*, *39 (2)*, 390-431.

Petrongolo, B. & Pissarides, C. (2006). Scale effects in market search. *Economic Journal*, *116*, 21-44.

Pissarides, C. (2000). *Equilibrium unemployment theory*. MIT Press, Cambridge Mass.

Pollak, A. (2003). Who's reading your X-Ray? *New York Times*, *November 16*.

Rappaport, J. (2007). Moving to nice weather. *Regional Science and Urban Economics*, *37*, 375-398.

Rauch, J. (1993). Does history matter only when it matters little? *Quarterly Journal of Economics*, *108 (3)*, 843-867.

Rohrbach-Schmidt, D. (n.d.). *The BIBB/IAB- and BIBB-BUAuA surveys of the working population on qualification and working conditions in Germany.* (BIBB-FDZ Daten- und Methodenberichte Nr. 1 / 2009)

Rosenthal, S. & Strange, W. (2004). Evidence on the nature and sources of agglomeration economies. *In Handbook of Urban and Regional Economics, Volume 4 (Amsterdam: Elsevier)*, 2119-2172.

Rosenthal, S. & Strange, W. (2008). The attunuation of human capital spillovers. *Journal of Urban Economics*, *64*, 373-389.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. W. Strahan and T. Cadell, London.

Spitz-Oener, A. (2006). Technical change, job tasks and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics*, *24 (2)*, 235-270.

Storper, M. & Venables, A. (2004). Buzz: Face-to-face contact and the urban economy. *Journal of Economic Geography*, *4 (4)*, 351-370.

Strange, W., Hejazi, W. & Tang, J. (2006). The uncertain city: Competitive instability, skills, innovation and the strategy of agglomeration. *Journal of Urban Economics*, *59 (3)*, 331-351.

Tempest, R. (1996). Barbie and the world economy. *Los Angeles Times*, *September 22*.

Ter Weel, B., Horst, A. Van der & Gelauff, G. (2010). *The Netherlands of 2040*. CPB Netherlands Bureau for Economic Policy Analysis.

Teulings, C. (1995). The wage distribution in a model of the assignment of skills to jobs. *Journal of Political Economy*, *103 (2)*, 280-315.

Venables, A. (2011). Productivity in cities: Self-selection and sorting. *Journal of Economic Geography*, *11*, 241-252.

Von Hippel, E. (1994). Sticky information and the locus of problem solving: Implications for innovation. *Mangement Science*, *40 (4)*, 429-439.

Wheeler, A. (2001). Search, sorting and urban agglomeration. *Journal of Labor Economics*, *19 (4)*, 880-898.

SAMENVATTING

(SUMMARY IN DUTCH)

Dalende informatie-, communicatie- en transportkosten stimuleren de ontwikkeling van mondiale productieketens. Het fabriceren van het hoofd van een barbiepop kan op een andere locatie plaatsvinden dan het maken van haar outfit of de marketing rondom haar persoonlijkheid. Steden blijven een belangrijke productieplek en Baldwin & Evenett (2012) refereren zelfs aan steden als de fabrieken van de 21ste eeuw. Waarom betalen bedrijven de hoge prijzen in steden ten tijden van mondiale productieketens en gratis communicatie via media als Skype?

Steden bundelen mensen en hun economische activiteit. Recent ontwikkelde informatie, communicatie- en transporttechnologieën veranderen de verdeling van werk en daarmee de rol van steden in de economie. Lagere kosten maken geografische afstand minder relevant voor productieketens. Verwacht werd dat het belang van steden daarmee ook zou afnemen. Tot nu toe bleven steden een belangrijke en dure vestigingsplek. Sterker nog: de correlatie tussen bevolkingsdichtheid en productiviteit neemt toe. De vraag wat steden faciliteren in het huidige productieproces vormt het onderwerp van deze dissertatie.

De hedendaagse verdeling van werk

Elke afzonderlijke taak van het productieproces kan vandaag de dag uitgevoerd worden op de meest efficiënte locatie. Waar enkele decennia geleden een auto op één locatie werd gefabriceerd is het productieproces van een auto nu verspreid over verschillende werelddelen. Voor vele taken is een stad nog steeds de meest efficiënte plek. Het ontwerpen van Nissan auto's gebeurt bijvoorbeeld onder andere in Londen terwijl de productie plaats vindt in onder andere Marokko en Maleisië. Waarom is dit zo?

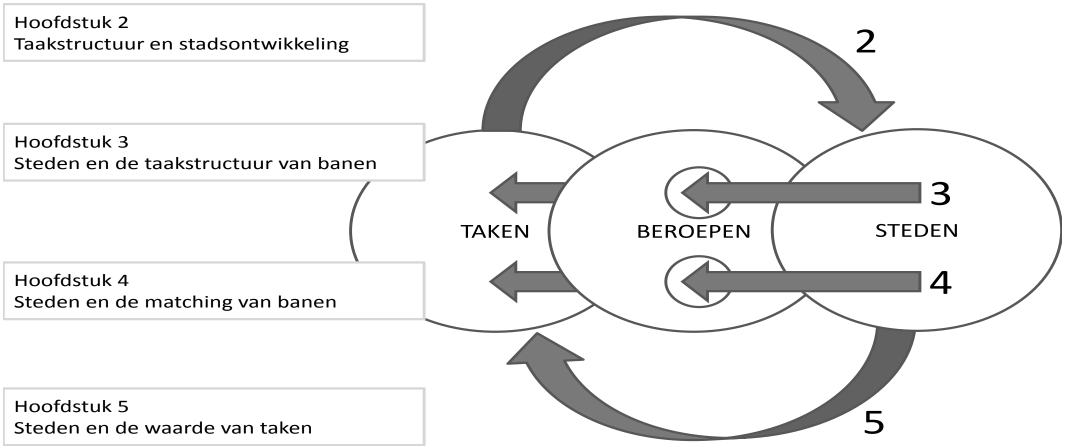
Schaalgrootte en nabijheid van dezelfde (of juist andere) taken, leveranciers en consumenten zijn productievoordelen van steden. Het delen van kennis en faciliteiten en het matchen van werknemers aan banen gebeurt gemakkelijker op dezelfde locatie dan op afstand. Met het opknippen van productieketens is het niveau waarop deze voordelen plaatsvinden veranderd. Waar vroeger agglomeratievoordelen relevant waren voor de locatie van industrieën zijn ze tegenwoordig relevant voor de locaties van het uitvoeren van bepaalde taken (Desmet & Rossi-Hansberg, 2009). Het gaat om de beste plek voor 'ontwerpen', 'repareren' en 'lassen' in plaats van die van de auto-industrie. Vandaag de dag specialiseren steden zich in bepaalde functies in plaats van in bepaalde sectoren (Duranton & Puga, 2005). Vooral taken waarvoor nabijheid en persoonlijke interacties van belang zijn renderen in steden.

Recente technologische ontwikkelingen, zoals de opkomst van de computer, beïnvloeden ook de verdeling van werk binnen en tussen banen. De noodzaak om bepaalde taken te laten uitvoeren door één werknemer neemt af dankzij betere coördinatie en communicatiemogelijkheden. Bovendien vervangen computers bepaalde taken terwijl andere taken juist complementair zijn aan computers. De meeste taken van taxichauffeurs zijn niet vervangen door computers, terwijl vele taken van een secretaresse zijn overgenomen.

Taken, banen en steden

De verdeling van werk tussen banen, bedrijven en locaties verandert. Wat is de rol van steden in mondiale productieketens? De nieuwe comparatieve voordelen van steden vragen om een nieuw perspectief om deze vraag te beantwoorden. Deze dissertatie analyseert de rol van steden vanuit een 'takenperspectief'. Een takenperspectief verduidelijkt wat steden faciliteren in huidige productieketens. Welke taken worden juist wel en welke juist niet uitgevoerd in steden? Zijn banen anders in steden dan buiten steden? De recente literatuur relateert het succes van steden veelal aan de capaciteiten van haar beroepsbevolking. Het is essentieel de capaciteiten van werknemers te onderscheiden van de taken die zij uitvoeren op het werk. Capaciteiten worden ingezet om taken uit te voeren en hiervoor loon te ontvangen, maar zorgen niet direct voor productie. Nieuwe verdelingen van werk zorgen ervoor dat werknemers met bepaalde capaciteiten andere taken zijn gaan uitvoeren. Dit maakt een onderscheid tussen taken en capaciteiten relevant. Deze studie toont verbanden tussen taken, banen en steden. Onderstaande figuur schetst de onderzochte relaties tussen de drie facetten per hoofdstuk.

Schematische weergave van deze studie



Groeiende steden voeren samenhangende taken uit

De interacties tussen taken, banen en steden resulteren in verschillende onderzoeksvragen. De eerste focus van deze dissertatie ligt bij de relatie tussen taken en de ontwikkeling van stedelijke economieën. Hoofdstuk 2 brengt in kaart in hoeverre de taakstructuur van 168 Amerikaanse steden de werkgelegenheidsgroei in deze steden verklaart. Hierbij wordt gekeken naar het belang van de nabijheid van andere taken; de connectiviteit van taken. Bepaalde taken, zoals samenwerken, profiteren sterk van persoonlijke interacties en nabijheid van andere taken. Andere taken, zoals boekhouden, kunnen juist gemakkelijk geïsoleerd uitgevoerd worden. Hoofdstuk 2 laat zien dat steden waarin in 1990 veel taken verricht werden die profiteren van nabijheid, een relatief sterke werkgelegenheidsgroei kenden in de periode 1990 tot 2009. Het takenpakket in deze steden sluit goed aan bij de comparatieve voordelen van Amerikaanse steden in de huidige economie. Begin jaren '90 waren er echter ook steden, zoals Detroit, die taken uitvoerden die relatief gemakkelijk geïsoleerd uitgevoerd kunnen worden. Deze steden kennen een minder harde werkgelegenheidsgroei of zelfs een krimp. Succesvolle steden van gisteren huisvestten industrieën die profiteren van agglomeratie, succesvolle steden van vandaag huisvesten taken die profiteren van agglomeratie. De clustering van de autoindustrie in Detroit resulteerde in een zeer succesvolle economie in de jaren '70. Vandaag de dag floreren echter juist steden zoals Boston en New York die zich richten op het uitvoeren kennis intensieve taken in de financiële wereld, professionele diensten en nieuwe technologie (Glaeser & Ponzetto, 2010).

Banen zijn anders in steden ...

Een takenperspectief draagt tevens bij aan het documenteren van economische activiteit over de ruimte. Takendata geven meer inzicht in de onderliggende geografische variatie in banen en in sectoren. Hoofdstuk 3 gaat na in hoeverre het takenpakket en de gevraagde vaardigheden van banen verschillen tussen locaties. Banen worden gedefinieerd als een beroep in een sector. De takenpakketten van banen variëren substantieel tussen kleine en grote Duitse steden. Ten eerste bevatten de banen een kleiner takenpakket in grote steden dan in kleine steden. Werknemers in grote steden zijn dus meer gespecialiseerd dan werknemers in kleine steden. Ten tweede stijgen de vereiste vaardigheden van een baan met de stadsgrootte waar deze wordt uitgevoerd. Waarschijnlijk hebben de meer gespecialiseerde werknemers in grote steden meer tijd om zich te focussen op hun belangrijkste taken en kunnen zij hierdoor meer vaardigheden voor deze taken ontwikkelen. Hoofdstuk 3 suggereert dat regionale loonanalyses rekening zouden moeten houden met de regionale variatie in de inhoud van banen.

... en matches tussen werknemer en baan zijn beter

De banendichtheid van een stad beïnvloedt naast het takenpakket van banen ook de match tussen werknemerscapaciteiten en hun taken. Werknemers zijn efficiënter wanneer hun takenpakket goed aansluit bij hun vaardigheden. De grote variatie aan zowel vaardigheden van werknemers als takenpakketten maakt een match complex. Voor werknemers en werkgevers zijn deze variaties veelal zichtbaar, maar voor onderzoekers niet. Hoofdstuk 4 maakt gebruik van een dataset met informatie omtrent de aansluiting tussen vaardigheden en taken van werknemers. De kwaliteit van de match tussen vaardigheden en het takenpakket neemt toe met de baandichtheid van de lokale arbeidsmarkt. De grotere keuze maakt werkgevers en werknemers kieskeuriger wat resulteert in een betere match. Deze match is belangrijker voor relatief getalenteerde werknemers en complexe banen omdat zij meer loon of opbrengst verliezen bij een slechtere match. Deze werknemers en banen vestigen zich dan ook onevenredig vaak in arbeidsmarkten met een hoge baandichtheid. Werknemers met een goede match tussen hun capaciteiten en taken ontvangen een hoger loon. Dit suggereert dat baandichtheid resulteert in efficiënte matches.

Maar de voordelen verschillen tussen steden

Tot slot draagt deze dissertatie bij aan het debat over de voordelen van verschillende type steden. Zowel steden met een diverse als steden met een gespecialiseerde sectorstructuur kunnen succesvol zijn. In gediversifieerde steden leren bedrijven van de verscheidenheid aan bedrijven. Bedrijven in gespecialiseerde steden profiteren van de nabijheid van soortgelijke bedrijven. Nabijheid en schaalgrootte zijn de vestigingsvoordelen van steden. Nabijheid drukt communicatie- en coördinatiekosten. Schaalgrootte biedt gelegenheid de kosten van faciliteiten te delen, zorgt voor een gemakkelijke match van werknemers uit dezelfde sector en voordelen van nabijheid van leveranciers en consumenten. Hoofdstuk 5 meet het belang van communicatie en coördinatie in beide type steden. Communicatie verschilt sterk tussen gediversifieerde en gespecialiseerde steden. In gediversifieerde steden vindt communicatie plaats tussen werknemers uit verschillende sectoren, dit stimuleert radicale innovatie en vormt het productieproces van relatief nieuwe producten en bedrijven. De communicatie is echter relatief complex gezien de verschillende achtergronden van de werknemers. In gespecialiseerde steden communiceren werknemers uit soortgelijke bedrijven met elkaar wat resulteert in optimalisatie van het productieproces. Deze communicatie is makkelijker omdat werknemers met eenzelfde achtergrond elkaar beter begrijpen. Hoofdstuk 5 toont aan dat het uitvoeren van communicatietaken in beide soorten steden een positief effect heeft op loon. Dit effect is echter groter in gediversifieerde steden dan in gespecialiseerde steden. De resultaten suggereren dat de economische voordelen van steden wisselen tussen verschillende type steden.

Deze dissertatie toont relaties tussen taken, banen en steden. Het onderzoeken van de takenstructuur van steden staat nog in de kinderschoenen. Hoofdstuk 6 bespreekt empirische en conceptuele uitdagingen van dit onderzoeksveld. De sleutelrol die steden vandaag de dag in de economie spelen maakt inzicht in de verbanden tussen de taken, banen en locaties van het productieproces van Barbie echter (beleids)relevant, zie Ter Weel et al. (2010).