ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Kok, Suzanne

Conference Paper

Town and city jobs: job contents, skills and city size

52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Kok, Suzanne (2012) : Town and city jobs: job contents, skills and city size, 52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/120632

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Town and city jobs: job contents, skills and city size

Suzanne Kok[*]

INCOMPLETE DRAFT VERSION - JUNE 2012

Abstract

This paper examines whether job contents vary with city size. We set out a theoretical model of the division of labour across cities relying on the model of James Baumgardner. The division of labour is limited by the extent of the market. Workers in large cities perform fewer tasks than workers in small cities with the same job. More able workers sort into large cities which results in more complex jobs in large cities. Using individual German task data, we support the predictions of the model. Jobs in large cities consist of other task packages than the same jobs in small cities. Workers in large cities focus more on their core-tasks and perform fewer subtasks than workers in small cities. Jobs require more cognitive skills when they are performed in large cities.

Keywords: Occupations, division of labour, skills, cities
JEL Classification: J24; J44; R23

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. As an ambitious doctor, you would rather work in the big city than in the small town. In the big city you have more chances to become an 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 is labour itself different? Does the job content vary across city size?

Back in 1988 James Baumgardner 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 focus on

[*]University of Groningen and CPB Netherlands Bureau for Economic Policy Analysis (s.j.kok@cpb.nl).

only half the activities. Duranton and Jayet (2011) translate the model of Baumgardner into an assumption 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 (1988*b*) and Garicano and 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 (e.g. Glaeser and Maré (2001) and Combes et al. (2008)). 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 at the occupation and industry level between cities hampers adequate analyses on the mechanism behind agglomeration economies.

In this paper we take a step towards unravelling the efficiency of cities by analysing the variation in job content across cities. Most dataset hinder such an attempt as they lack spatial variation in job content. We make exploit the BIBB/BAuA surveys of the working population in Germany which includes job activities for individuals across German cities. Our main result is that the specialisation level of jobs and the requirement for 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 refers to the model of Baumgardner (1988*a*). The production of a good consists of the performance of a continuum of tasks. The price of each task depends on the output of the task, while the productivity increases with individual 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. Next, we extent this model by introducing heterogeneous workers and heterogeneous tasks. High skilled workers are more productivity given endowment and therefore perform more tasks than low skilled workers. Skilled workers are unequally distributed over space. This results in a uneven distribution of tasks to skill workers across space. The complexity of a subset of tasks rises with the city population.

Second, we test the predictions of our theoretical set up using the BIBB/BAuA surveys of the working population on qualification and working conditions in Germany. In contrast to most information on job tasks, the BIBB 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 the job tasks, occupation, industry, demographic characteristics, education, location and so forth. We construct two measures for job content: the number of subtasks (performed 'sometimes' or 'rarely') and the require of cognitive core tasks. As documented by Duranton and Jayet (2011), scarce occupations occur more often in large German cities than in smaller German cities.

Controlling for the job (occupation - industry combination) average, we find that workers in large cities on average perform 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) than when it is performed in a large city (more than 100,000 inhabitants). Mainly high skilled workers specialise more in large cities than in small cities. Jobs in larger cities also require more cognitive skills than the same jobs in small cities. Again, this spatial variation is substantially larger for high skilled than for other workers. These results are robust to the inclusion of several co-variates and across several sub samples.

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 (1988*a*). The specialisation of workers into certain job tasks increases with market size. Duranton and Jayet (2011) argue that larger markets allow workers to perform in the most efficient way. The micro-foundations of this theoretical work are mostly lacking and the work relies on directly assumed relations. Another strand in the literature (see Becker and 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 paper, we examine whether the extent of the market is relevant for the division of labour.

Empirically, this field is left rather untouched. The empirical work tends to focus on case-studies such as the work of Baumgardner (1988*b*) and Garicano and Hubbard (2009) who study the division of labour across market sizes for respectively doctors and lawyers. Analyses covering the whole economy are scarce and focus on occupational data and not on variation within jobs. Duranton and Jayet (2011) show that scarce occupations occur more 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 across French cities can be explained by differences between occupations 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 job content instead of controlling for worker skills by using fixed effects.

Our work also relates to the literature on spatial sorting of workers. The complementary between cities and skills as documented by Glaeser and Ressenger (2010) results in the sorting of the higher skilled workers into the larger cities. Indeed, empirical evidence for the sorting of workers is found for both the US (Berry and Glaeser (2005)) and France (Combes et al. (2009)). This spatial sorting partly explains the urban wage premia as indicated by Combes et al. (2009), Mion and Natticchioni (2009) and others.

Lastly, our work relates to the empirical work on the (micro-) foundations on the efficiency of cities. For an overview of this very broad and large literature strand we refer to the work of Glaeser and Gottlieb (2009). We only investigate in a possible mechanism behind the efficiency in cities which lack empirical research so far. We do not measure agglomeration economies directly.

The rest of the paper is structured as follows. The next section sets out a simple framework to justify our empirical analyses. Section 3 provides insight in the database construction, the main variables and some descriptive. Section 4 presents the results on the spatial variation in job content. In section 5, some robustness checks are presented. Section 6 concludes.

2 Spatial variation in job content

To justify our empirical setting, we set out a framework for the division of labour across cities. We start with the basic setting of Baumgardner (1988*a*). 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. The division of labour is bound by coordination costs. Next, we extent the model by allowing productivity differences between low and high skilled workers and examine the effects of spatial sorting of more able workers.

2.1 Ingredients

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 of length one and indexed by $s \in [0, 1]$.

The price of each task depends on the output of that task (Q) and a vector of local market demand shifters (b).

$$p(s) = p(Q(s), b) \tag{1}$$

with $p_Q < 0$.

The demand of a task (s) is independent of the output of other tasks. This assumption does not hold for all tasks as some tasks are complementary. For simplicity, we assume that the demand for all tasks of a good is symmetric. All tasks are needed to produce one good. Baumgardner (1988b) finds that the local population is the most important of these local demand shifters. Other examples of local demand shifters are demographic characteristics of the population, education and income.

The more time a worker spends on the production of one specific task, the more specific skills for producing this task he develops (see Becker and Murphy (1992)). As the worker has limited endowment of time, there are increasing returns to input. The more a worker specializes in one task, the more efficient he becomes in producing that specific task. For instance, a doctor who only performs heart surgery will learn more about that surgery than a doctor who also removes appendices. The heart surgery specialist will be more efficient than the general surgeon in performing a heart surgery.

$$q(s) = [x(s)]^2 \tag{2}$$

Again, we assume symmetry in the production technology of tasks. In the model, all tasks in the segment of tasks are equal.

Lastly, the worker is endowed with limited time.

$$E \geq \int_{\delta} x(s)ds \tag{3}$$

δ is the subset of tasks the worker chooses the produce from the total subset of tasks of the good.

2.2 Extent of the local labour market

The division of labour is efficient due to increasing returns to individual task input (equation (2)). For the production of a good, all tasks of the segment should be performed. When workers cooperate, they can divide the tasks, benefit from the increasing returns to individual input and become more efficient. Thus, the subset of tasks of each worker (δ)

becomes smaller with the number of workers in the market (N). The division of labour is however bound by coordination costs C (Becker and Murphy (1992)). When workers divide the complementary tasks of a good, they need to coordinate the production process. Even if workers fully cooperate and do not compete to some extent, information on the tasks will be lost within the coordination process. Tacit knowledge about tasks is difficult to transfer across different workers. The coordination costs hinder workers to perform a unique subset of tasks and fully exploit the increasing returns. [1] A large local population rises demand for certain scarce tasks (Duranton and Jayet (2011)) and makes it possible to exploit the increasing returns to individual input of these tasks. In large cities, tasks can be organised more efficiently over workers and the large demand increases the performance of scarce tasks. Although workers do not perform a unique subset of tasks to reduce coordination costs, the division of labour is more extensive in large cities.

$$\delta = (1/N)C \tag{4}$$

A small local market only houses a few doctors. It is beneficial when these doctors divide the medical activities and specialise in a subset of these activities. In a large market, not all doctors perform a unique set of medical activities. For instance, most large cities employ several psychiatrists whom seem to perform a similar subset of medical activities. For the patient and the doctor it is not always immediately clear what kind of specialist he needs or what the precise issue is. A screening by a more general psychiatrist is therefore important. However, the demand for scarce medical activities rises with population size and with that specialisation possibilities for doctors. After the screening, the patient can be consulted by a specialist.

2.3 Heterogeneous workers

Next, we consider workers with heterogeneous skills. High skilled workers are more productive than middle, low and un-skilled workers. For simplicity we here focus on the productivity and task performance of high versus low skilled workers. When high skilled workers use the input M to produce the subset of tasks Δ, they produce θ times more output than low skilled workers with the same input:

$$M_H^2/\Delta^2 = \theta q \tag{5}$$

with $\theta > 1$.

High skilled workers produce more than low skilled workers with the same endowment. The demand of all tasks is equal so it is inefficient to produce more of a subset of Δs and not on of another subset. To see this, consider two workers, one high skilled (h) and one low skilled (l) both endowed with input M. If both workers produce half of the tasks ($\Delta(s)/2$), the subset of worker h will be produced $\theta[M^2/(\Delta(s)/2)^2]$ times and the subset of worker l $[M^2/(\Delta(s)/2)^2]$ times. As the production of one good requires $\Delta(s)$, the overproduction of worker h is useless. Therefore, worker h performs more tasks such that the output for each s is equal: $\theta[M^2/(\Delta(s_h))^2] = [M^2/(\Delta(s_l))^2]$ with $\Delta(s_h) > \Delta(s_l)$.

[1] It should be noted that in the model of Baumgardner, there is no overlap in the worker's subset of tasks. All workers perform a unique subset of tasks when workers cooperate. This is not realistic as coordination costs affect the division of labour (Becker and Murphy (1992)). For some tasks combinations it is too costly to be performed by multiple workers.

Table 1: Division of tasks - Example

Task	1	2	3	4	5
City A	Very low	Low	Low	Low	High
City B	Very low	Low	Low	Very high	Very high

High skilled workers perform more tasks than low skilled workers. In terms of local demand shifters, the population at which a high skilled worker obtains an optimal specialisation level is higher than for low skilled workers.

2.4 Heterogeneous tasks

For the production of a good a wide variety of tasks is performed. Here, we allow the task content and required skills to vary. In case of heterogeneous tasks, both the number of tasks performed and the content of tasks varies with the skill level of workers. The most complex tasks are performed by highest skilled (or most able) workers while the least complex tasks are performed by the lowest skilled (or least able) workers. Consider a market with N workers. Each worker performs a subset of $\delta = (1/N)C$ tasks. We order tasks $s \in [0, 1]$ so that the complexity of tasks increases. The least skilled worker among the N workers then performs the tasks on the subset $[0, \delta]$. The second lowest skilled worker the subset $[\delta, 2\delta]$ and the highest skilled worker the subset $[1 - \delta, 1]$.

Skilled workers are unequally distributed across space. Workers with high observed and unobserved abilities sort into larger cities for better education, career possibilities, spouse markets and amenities (Glaeser and Maré (2001) and **?**). These sorting patterns have consequences for the division of tasks across skill levels. In cities with a more able work force tasks are performed by workers with higher ability levels. Consider the example of table (1). To produce one good requires 5 tasks (task 1 is the least complex task and task 5 the most complex). City A and city B both employ 5 workers. The most skilled worker performs task 5, the least skilled worker task 1. The skill levels of the workers are different in the two cities. Therefore, task 5 is performed in city A by a worker who has less skills than the worker who performs task 4 in city B.

2.5 Empirical predictions

In summary, a set of tasks need to be performed for the production of one good. The demand price of each task is falling in output of that task and depends on the local market demand shifters. There are increasing returns to scale in production technology. If a worker produces fewer tasks, output per time unit and the marginal product increase. As a counter force, the marginal revenue falls with output per task.

To sharpen attention to the main forces of interest, we empirically analyse only the variation in job content across cities. Section 2.2 shows that cooperation with other workers in the local market enhances possibilities for a more efficient division of labour. The larger local demand induces specialisation and performance of scarce tasks with a relatively low demand. We expect the subset of tasks by worker (δ) to fall with the population of the local market (N):

$$\frac{\partial \delta}{\partial N} < 0 \tag{6}$$

The jobs of workers in large cities are more specialistic than the same jobs in small cities.

High skilled workers are more productive and therefore perform more tasks within the same time unit. The estimations therefore control for the (unequal) distribution of skills over cities.

Another content of jobs is the complexity level of its tasks. Section 2.4 suggest that the unequal distribution of skills over space affects the distribution of tasks over workers. The ability level of a worker is only partly measured with his educational degree. High skilled workers cluster in large cities. If this is the case, tasks in large cities are performed by more able workers than the same tasks in small cities. Local competition will induce the complexity of these tasks and with that the requirement of skills. The complexity (S) of the subset of tasks increases with city size:

$$\frac{\partial S}{\partial N} > 0 \tag{7}$$

The jobs of workers in large cities require more complex tasks than the same jobs in small cities.

Complexity is measured with the require of cognitive skills.

It should be noted that our model discusses one good. The number of required tasks and the complexity level of the tasks varies between goods. For simplicity, we compare the task subsets of different jobs assuming each industry to produce a good with a typical subset of tasks. A job is defined as an occupation within an industry. Thus, we do not compare the tasks of secretaries but the tasks of secretaries within a certain industry. The estimations measure the variation in number of tasks and required cognitive skills across cities for a certain occupation-industry combination.

3 Data, indicators and descriptive statistics

3.1 Data

The empirical predictions demand individual task data for workers across cities of different size. In contrast to most task datasets, BIBB/BAuA surveys of the working population on qualification and working conditions in Germany includes individual information on tasks, occupations and locations. The BIBB is a survey among a representative sample of Germans, here we employ the 2006 wave as this is the most recent wave. The major concerns of the survey are to measure qualifications, career history and detailed job characteristics of the German labour force. The survey of 2006 includes information on demographic characteristics, education, wage, job characteristics and location for 20,000 Germans. Here, we focus on the definitions and construction of our main variables: tasks and local markets.

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 number of different tasks appearing in the BIBB are displayed in table 3. For each task, the survey examines the frequency of appearance 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. The disadvantage of estimations for the whole economy is that possible tasks vary between industries. Therefore, the database does not include all possible tasks. The subset of possible tasks is smaller than in real life. This indicates that we underestimate the variation in job contents.

The dataset contains information on the size of the city of residence. For the largest cities it also includes information whether the workers live in the central or in the peripheral area. For the descriptive data we exploit the variation between the seven different categories. We exclude the information on central versus peripheral area as this indicates the area of residence and not the working area. Table 2 presents the (weighted) number of observations in the dataset for the seven size categories. Besides cities, also 16 regions ('Bundesland') are distinguished in the dataset.

3.2 Job content

A job is defined as an 3 digit occupation and 2 digit industry combination. Throughout the paper the term 'job' refers to an occupation with 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. We focus in our analyses on two specific job contents: the number of (different) tasks a worker performs and the requirement of cognitive skills for the job. The number of different tasks a worker performs indicates the level of specialisation of the work force. The tasks of a job as measured in the BIBB refer largely to required skills. The specialisation level of the job is therefore constructed by the number of subtasks a worker performs instead of core-tasks. The fewer subtasks a worker performs, the higher his specialisation level. Column (1) of table 4 shows the average number of subtasks by age groups, gender, education and native language. On average, workers perform 15 subtasks and 18 core-tasks out of the total of 58 possible tasks. Older workers (above 50 years) perform fewer subtasks than younger workers. Work experience enhances specific knowledge and with that a more specialist 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, with an exception for university education. Management occupations content on average 18 subtasks, elementary occupation only 12 subtasks (see table 5). The variation between subtasks is less strong among industries: workers in other service sectors perform on average 14 subtasks while workers in wholesale trade perform the most sub tasks (17).

The requirement of cognitive skills is measured by the following seven tasks: research, adapt to unforeseen problems, mathematical skills, technical skills, solving new problems, process optimizing and do things you have not learned before. The average job requires 1.8 cognitive skills (see Table 4). Younger workers need more cognitive skills than older workers. Logically, the requirement of cognitive skills increases with education level. The jobs of males on average require 1.90 cognitive tasks while the jobs of females require 1.63 cognitive tasks. Native speakers are less often in a cognitive intensive job than non-native speakers. Concerning broad occupations, agricultural and fishery occupations require the least cognitive skills while professional occupations require the most. Cognitive skills are the most important in the administration and support sector and the least in wholesale trade.

3.3 Scarce occupations

Before we turn to our own estimations we replicate the work of Duranton and Jayet 2011 for German cities and test whether scarce occupations occur more in large cities. Using a logit approach, we estimate the probability that a scarce job (occupation-industry combination) appears in a city. Within each industry, the occupations are classified into four: 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 A describes the full estimation method.

Table 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 obtain valid results, we should control for the division of jobs across cities.

4 Job content across cities

4.1 Division of labour

We start our empirical analysis by examine whether the division of labour is bound by the extent of the market. Do doctors in a large city perform a smaller subset of medical activities than doctors in a town?

The subset of tasks is different for each job. A lawyer performs a different segment of activities than a doctor or a employee in a textile firm. Therefore, we control for the job (occupation - industry combination) of the worker to correct for the average number of subtasks performed in his job. (Column (1) in Table 7). The variation in number of subtasks performed by workers in the same job is substantial: the R^2 of a simple OLS-regression with the job average is only 0.26 and the correlation between number of subtasks performed by a worker and the job average is 0.45 (significant at the 1% level). The local demand shifters seem to cause a large variation in specialisation level between workers, as already suggested by the work of Baumgardner (1988*b*).

Allowing the number of subtasks to vary across city size only slightly increases the explanatory value. We distinguish three sizes 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). As expected, workers in large cities perform fewer subtasks and are more specialised given their job (column (2)). The level of specialisation of a certain job increases linearly with city size. Although the variation in specialisation across city size is significant, the size is only modest. On average a worker in a city with less than 20,000 inhabitants performs 15.73 subtasks. In medium cities, this is 15.65 and in large cities, with more than 100,000 inhabitants, a worker performs on average 15.62 subtasks. Workers in a large city perform 5% of a standard deviation fewer subtasks and workers in a medium city 3%.

Other local demand shifters (b in equation 1) 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 native speakers. Column (3) 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 as described in section 2. Workers with a college or university degree perform 0.80 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 (2). The number of subtasks also significantly varies across other subgroups. 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 variation (see Becker and Murphy (1992)).

Lastly, column (4) adds cross-terms between city size and a holding college or university degree. Analyses including these dummy variables examine whether especially high skilled workers obtain fewer tasks in large cities. The estimates show negative and significant coefficients for these dummy variables. High educated workers in large cities perform - on average - 0.7 less subtasks than low and medium educated workers in small cities. The linear impact of medium cities becomes positive. The linear impact of university degree increases. Thus, the division of labour is only more extensive for high educated workers. As these workers also perform more subtasks, this makes sense.

4.2 Skills in cities

In most countries, the largest cities house the best workers. Larger cities provide more education opportunities, more opportunities to make a career, and are today also the places with the best consumption possibilities. More able workers are drawn to cities for education and career opportunities, a taste for amenities and higher wages caused by agglomeration economies. The sorting of more able workers into larger cities and the existence of agglomeration economies is largely analysed (see Glaeser and Gottlieb (2009) for an overview). Here, we test whether the same job requires more skills in large cities than in small cities. The distribution of tasks according to their complexity varies with the skill distribution of the local workers.

As column (1) in table 8 shows, only a small proportion of the variation in the requirement of cognitive skills is explained by the job. The correlation between the amount of cognitive skills the worker indicates that are required for the job and the national job average is 0.46 and significant at the 1% level.

Allowing the requirement of cognitive skills to vary across city size slightly improves the explanatory power of the model. The coefficients of city population dummies are positive and significant. Workers in large (medium) cities indicate that their job requires 0.05 (0.02) cognitive skills more than workers with the same job in small cities. An average job requires 1.56 cognitive skills with a standard deviation of 1.06. In large and medium cities this requirement lies about respectively 5% and 2% of a standard deviation higher.

Column (3) presents the results when demographic and education information are included. Given the job average, high skilled workers perform fewer cognitive tasks than low skilled workers. The jobs of older, males and native workers require fewer cognitive skills than the same jobs for young, female and non-native workers.

Lastly, in column (4) we add cross-terms between university degree and city size dummies. High skilled workers in large cities perform 0.06 cognitive skills more than low skilled workers in small cities. High skilled workers in medium cities perform fewer cognitive skills than low skilled workers in medium cities.

In summary, the specialisation level and required cognitive skills of jobs increases with city population. This is especially the case for high skilled workers. The variation 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. The fact that we do find significant results suggest substantial spatial variation in the number of subtasks performed.

5 Robustness

Previous estimations include several assumptions which we test in this section with four robustness checks. First, we run estimations with different spatial units to test whether the results are robust for variation in city size and regions. Second, section 5.2 provides separate estimations for the manufacturing and service sectors to indicate whether the variation of job content across city size occurs in both sectors. Next, we perform separate estimations for different skill groups (section 5.3). Lastly, we test the hypothesis that learning and experience could affect the results.

5.1 Spatial units

The extent of the market is defined by the local labour market for demand and supply. The empirical analyses define a local labour market as a city. 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 3 categories: small, medium and large cities. Column (1) of table 9 presents estimates for the level of specialisation in which we distinguish seven city size categories. The ratio between absolute size and the cross-term between size and high skilled varies between the categories. The number of subtasks a high skilled worker performs diminishes with the size of the city of residence. The ratio between the linear impact of city size (the size dummies) and the specific impact for high skilled workers does vary across city categories. Column (2) presents estimates with the same city categories for the requirement of cognitive skills. The importance of cognitive skills increases linearly with city size for high skilled workers and for the full sample of workers (not shown here).

In column (3) and (4) present the results using other 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 than in a region with a low population density. Therefore, population density of German regions ('Bundesland') provides us information on the possibilities to cooperate and divide the tasks among workers. Column (3) 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. High skilled workers in high density areas perform fewer subtasks and are more specialist than low skilled workers in low density areas. The same holds for the full sample (not shown here). High skilled workers in dense areas also obtain more cognitive skills (column (4)).

We cannot rule out that we do not measure the division of labour at the optimal spatial units. However, as our results are robust over several spatial units, we assume that the found relations are not caused by measurement error of the spatial units.

5.2 Manufacturing versus services

Manufacturing and service products posses different production processes. The local market for the demand and supply of service products may be much more local than the local 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 low importance of tacit knowledge enables manufacturing firms to split up their production process easier over workers and even over space. Therefore, the spatial unit of interest might also vary between manufacturing and services. To assess whether these differences indeed occur and whether our results hold for both type of goods, table 10 presents separate results for manufacturing and service sectors. Columns (1) and (2) show that the level of specialisation for manufacturing and service industries is higher in larger cities. Again, the linear impact of city size turns positive when we include cross-terms with high skilled workers. Thus, high skilled workers in the manufacturing and service industry perform fewer subtasks when they are located in a large city. Middle and low skilled workers perform more subtasks when they are located in a large city. Columns (3) and (4) present the same exercise for the requirement of cognitive skills. In large cities, the same manufacturing job requires fewer cognitive tasks than in a small city. Within service industries this is the other way around. These results suggest that workers in the service industry sort into large cities and more able workers in the manufacturing industries do not.

5.3 Ability sorting

The previous results suggest especially high skilled worker are more specialised in large cities. And, in line with the sorting of the more able into large cities, jobs require more cognitive skills in large cities. To more precisely test whether only high skilled workers perform fewer tasks we run separate estimations by skill group. Table 11 presents the results. Columns (1)-(4) show that only high skilled workers perform significantly less subtasks in large cities. Workers with a degree from an operational college even perform more subtasks in large cities. The coefficient for un- and low-skilled are insignificant but negative. It seems to be the case that especially high skilled workers perform fewer subtasks in large cities. We assume the division of labour of their job task is more beneficial than the division of labour of other skill groups. Columns (5)-(8) show that all skill groups indicate that their job requires more cognitive skills when they are located in a large city. Given their education, more able workers sort into larger cities.

5.4 Learning

Older workers are more specialised than younger workers. Experience probably leads to more expert knowledge and with that to more specialisation. Table 12 presents results for separate age groups. Young and old workers perform fewer subtasks when they are located in a large city and more cognitive skills.

6 Conclusion

This paper examines the spatial variation in job contents. The demand for certain activities and the supply of skills vary with city size. In this paper we analyse whether this results in different job contents in terms of specialisation level of the job and required cognitive skills.

The theoretical model of the paper indicates that the division of labour is efficient. When workers focus on less job tasks, they will develop more specific skills for the performance of these tasks and become more efficient. The larger demand in larger cities increases possibilities to cooperate and divide activities over workers. The sorting of the more able workers into large cities furthermore enhances variation in the required skills of jobs. In large cities, a job is performed by more skilled workers than the same job in a small city.

The empirical predictions from the theoretical model are supported with estimations using the BIBB-BAuA surveys. Workers in large cities perform fewer subtasks and focus therefore more on their core-tasks then workers in small cities. Jobs require more cognitive skills when they are performed in large cities. These results are robust to a variety of specifications and other explanations and sub samples.

Our work shows spatial variation in job contents. Likely, this variation affects the underlying mechanism of the efficiency of cities. Analyses at the occupational or industry level lack important variation and with that probably explanatory power.

References

Bacolod, M., Blum, B. and Strange, W. 2009. 'Skills in the City'. *Journal of Urban Economics* 65 (2), 136–153.

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

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

Becker, G. and Murphy, K. 1992. 'The Division of Labor, Coordination Costs and Knowledge'. *Quarterly Journal of Economics* 107 (4), 1137–1160.

Berry, C. and Glaeser, E. 2005. 'The divergenec of human capital levels across cities'. *Papers in Regional Science* 84 (3), 407–444.

Combes, P., Duranton, G. and Gobillon, L. 2008. 'Spatial Wage Disparities: Sorting Matters!'. *Journal of U* 63(2), 723–742.

Combes, P., Duranton, G. and Gobillon, L. 2009. *The Economics of Agglomeration*. Chicago and London: University of Chicago Press. chapter Estimating Agglomeration Economics with History, pp. 15–65.

Combes, P., Duranton, G., Gobillon, L. and Roux, S. 2012. ‘Sorting and the local wage and skill distributions in France’. *Regional Science and Urban Economics* p. Fortcoming.

Duranton, G. and 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.

Garicano, L. and 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.

Glaeser, E. and Maré, D. 2001. ‘Cities and Skills’. *Journal of Labor Economics* 19 (2), 316–342.

Glaeser, E. and Ressenger, M. 2010. ‘The Complementarity between Cities and Skills’. *Journal of Regional Science* 50 (1), 221–244.

Glaeser, E.L. and Gottlieb, J.D. 2009. ‘The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States’. *Journal of Economic Literature* 47, No. 4, 983–1028.

Mion, G. and Natticchioni, P. 2009. ‘The spatial sorting and matching of skills and firms’. *Canadian Journal of Economics* 42 (1), 683–701.

Table 2: Observations seven city categories

Category	Inhabitants	Employees (weighted)
1	1-1.999	1907418
2	2.000-4.999	2651726
3	5.000-19.999	6701806
4	20.000-49.999	4570964
5	50.000-99.999	2199916
6	100.000-499.999	3972548
7	500.000-...	3621968

Employees is the weighted number of observations in the dataset.

Table 3: 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	

7 Tables

Table 4: Job content by subgroup

	Number of	
	sub tasks	cognitive skills
Average	15.61	1.67
Younger than 35	15.74	1.84
35-50	15.83	1.74
Above 50	14.97	1.72
Men	15.01	1.90
Female	16.18	1.63
No education	11.82	1.66
High-school	14.44	1.79
Operational college	15.44	1.59
College	17.68	1.63
University	16.09	2.13
Native speaker	15.66	1.67
Non-native speaker	14.81	1.76

There are 40 subtasks defined. On average workers perform 15.61 subtasks with a standard deviation of 5.79. We distinguish 7 cognitive skills, on average a worker indicates that his job requires 1.76 cognitive skills with a standard deviation of 1.09.

Table 5: Job content by group

	Number of		Number of		
	sub tasks	cognitive skills	sub tasks	cognitive skills	
Managers	17.69	2.03	Manufacturing	15.60	1.53
Professionals	15.81	2.18	Electricity and gas	15.99	1.48
Technicians	16.20	1.81	Water supply	15.35	1.38
Clerks	16.25	1.69	Construction	16.55	1.45
Service workers	15.05	1.66	Wholesale train	16.82	1.25
Agricultural and fishery	15.74	1.24	Transport	15.79	1.59
Operators and assemblers	13.29	1.29	Accommodation and food	15.98	1.68
Elementary	11.53	1.63	Information and communication	14.75	1.56
			Financial	16.50	1.92
			Professional. scientific and technical activities	15.39	1.93
			Administration and support	15.42	2.32
			Education	15.41	1.90
			Arts. entertainment	15.93	1.95
			Other services	13.50	1.69
			Household	14.06	1.70
			International organisations	16.77	1.33
Standard deviation	1.93	0.33		0.92	0.28

Occupations are defined by 1 digit ISCO 1988 codes, industries by 1 digit NACE codes.

Table 6: Estimation results logit for all occupations - 6 city categories

	Scarcity			
	Very High	High	Low	Very Low
Size 1	-1.285***	-1.118***	-0.927***	0
	[0.209]	[0.201]	[0.190]	
Size 2	-1.591***	-1.594***	-1.102***	0
	[0.207]	[0.205]	[0.199]	
Size 3	-1.450***	-1.156***	-1.008***	0
	[0.207]	[0.198]	[0.189]	
Size 4	-0.356	-0.296	-0.402**	0
	[0.219]	[0.209]	[0.197]	
Size 5	-0.810***	-0.744***	-0.628***	0
	[0.200]	[0.196]	[0.187]	
Size 6	0	0	0	0

The estimation method is explained in Appendix A. 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. Size 1 indicates cities with less than 5,000 inhabitants, size 2 between 5,000 - 20,000 , size 3 between 20,000-50,000, size 4 between 50,000 and 100,000, size 5 between 100,000 and 500,000 and size 6 refers to cities with more than 500,000 inhabitants.

Table 7: Level of Specialisation

	Number of subtasks			
	(1)	(2)	(3)	(4)
Job average	1.007***	1.007***	0.966***	0.966***
	[0.000]	[0.000]	[0.000]	[0.000]
Medium city		-0.078***	-0.096***	-0.015***
		[0.002]	[0.002]	[0.003]
Large city		-0.109***	-0.160***	0.090***
		[0.002]	[0.002]	[0.003]
University * medium city				-0.314***
				[0.006]
University * large city				-0.791***
				[0.005]
No education			-1.345***	-1.319***
			[0.007]	[0.007]
Operational college			0.255***	0.290***
			[0.005]	[0.005]
College | University			0.795***	1.173***
			[0.006]	[0.006]
Age			-0.023***	-0.024***
			[0.000]	[0.000]
Female			-0.758***	-0.752***
			[0.002]	[0.002]
Native speaker			0.176***	0.182***
			[0.003]	[0.003]
Observations	15670	15670	15670	15670
R-squared	0.261	0.262	0.273	0.274

Note: individual data. Job average indicates the national average number of subtasks performed by workers in the occupation - industry combination. City size is defined by dummy variables. Small cities house less than 20,000 inhabitants, medium cities house between 20,000 and 100,000 inhabitants and large cities more than 100,000 inhabitants. Small cities are the reference group. Education is measured with dummy variables of the highest finished degree, high school graduation is the reference group. Standard errors in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Table 8: Cognitive skills

	Required cognitive skills			
	(1)	(2)	(3)	(4)
Job average	0.982***	0.978***	0.949***	0.949***
	[0.000]	[0.000]	[0.000]	[0.000]
Medium City		0.018***	0.013***	0.028***
		[0.000]	[0.000]	[0.001]
Large City		0.049***	0.034***	0.020***
		[0.000]	[0.000]	[0.001]
University * Medium City				-0.052***
				[0.001]
University * Large City				0.039***
				[0.001]
No education			0.172***	0.170***
			[0.001]	[0.001]
Operational college			-0.048***	-0.050***
			[0.001]	[0.001]
College | University			0.018***	0.017***
			[0.001]	[0.001]
Age			-0.005***	-0.005***
			[0.000]	[0.000]
Female			0.112***	0.112***
			[0.000]	[0.000]
Native speaker			-0.005***	-0.006***
			[0.001]	[0.001]
Observations	15670	15670	15670	15670
R-squared	0.253	0.254	0.260	0.260

Note: individual data. Job average indicates the national average number required cognitive tasks in the occupation - industry combination. City size is defined by dummy variables. Small cities house less than 20,000 inhabitants, medium cities house between 20,000 and 100,000 inhabitants and large cities more than 100,000 inhabitants. Small cities are the reference group. Education is measured with dummy variables of the highest finished degree, high school graduation is the reference group. Standard errors in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Table 9: Spatial units

	Seven city sizes		Inhabitants per square km	
	Specialisation	Cognitive	Specialisation	Cognitive
	(1)	(2)	(3)	(4)
Job average	0.965***	0.947***	0.967***	0.951***
	[0.000]	[0.000]	[0.000]	[0.000]
Size 2	-0.075***	0.014***	0.011***	0.053***
	[0.004]	[0.001]	[0.003]	[0.000]
Size 3	-0.007*	0.041***	0.326***	0.015***
	[0.004]	[0.001]	[0.005]	[0.001]
Size 4	-0.174***	0.026***		
	[0.005]	[0.001]		
Size 5	0.017***	0.017***		
	[0.004]	[0.001]		
Size 6	0.078***	0.042***		
	[0.004]	[0.001]		
University * size 2	-0.661***	0.117***	-0.311***	0.027***
	[0.007]	[0.001]	[0.005]	[0.001]
University * size 3	-0.703***	0.012***	-1.167***	0.114***
	[0.008]	[0.001]	[0.009]	[0.002]
University * size 4	-0.756***	0.039***		
	[0.009]	[0.002]		
University * size 5	-0.972***	0.096***		
	[0.008]	[0.001]		
University * size 6	-1.449***	0.129***		
	[0.008]	[0.001]		
No education	-1.322***	0.169***	-1.340***	0.171***
	[0.007]	[0.001]	[0.007]	[0.001]
Operational college	0.288***	-0.050***	0.280***	-0.050***
	[0.005]	[0.001]	[0.005]	[0.001]
College | University	1.585***	-0.056***	1.062***	-0.005***
	[0.008]	[0.001]	[0.006]	[0.001]
Age	-0.024***	-0.005***	-0.023***	-0.005***
	[0.000]	[0.000]	[0.000]	[0.000]
Female	-0.751***	0.112***	-0.757***	0.111***
	[0.002]	[0.000]	[0.002]	[0.000]
Native speaker	0.170***	-0.004***	0.189***	-0.002***
	[0.003]	[0.001]	[0.003]	[0.001]
Observations	15670	15670	15670	15670
R-squared	0.275	0.261	0.274	0.261

Note: individual data. Job average indicates the national average number of subtasks performed by workers in the occupation - industry combination. City size is defined by dummy variables, sizes are defined as in table 2 for columns (1) and (2). In columns (3) and (4) size refers to the density level of the region ('Bundeslande'): 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. Education is measured with dummy variables of the highest finished degree, high school graduation is the reference group. Standard errors in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Table 10: Manufacturing versus Services

	Specialisation		Cognitive tasks	
	Manufacturing	Services	Manufacturing	Services
	(1)	(2)	(3)	(4)
Job average	0.964***	0.966***	0.983***	0.932***
	[0.000]	[0.000]	[0.001]	[0.001]
Medium city	0.193***	-0.305***	-0.021***	0.087***
	[0.004]	[0.004]	[0.001]	[0.001]
Large city	0.062***	0.058***	-0.027***	0.075***
	[0.004]	[0.004]	[0.001]	[0.001]
University * medium city	-0.422***	-0.070***	-0.016***	-0.102***
	[0.009]	[0.007]	[0.002]	[0.001]
University * large city	-0.683***	-0.771***	-0.006***	0.026***
	[0.009]	[0.006]	[0.002]	[0.001]
No education	-1.156***	-1.627***	0.263***	0.054***
	[0.010]	[0.010]	[0.002]	[0.002]
Operational college	0.116***	0.490***	-0.025***	-0.076***
	[0.008]	[0.007]	[0.001]	[0.001]
College | University	1.220***	1.115***	0.025***	0.026***
	[0.010]	[0.008]	[0.002]	[0.002]
Age	-0.026***	-0.022***	-0.005***	-0.004***
	[0.000]	[0.000]	[0.000]	[0.000]
Female	-0.724***	-0.819***	0.092***	0.138***
	[0.003]	[0.003]	[0.001]	[0.001]
Native speaker	-0.002	0.444***	0.023***	-0.040***
	[0.005]	[0.005]	[0.001]	[0.001]
Observations	6583	9087	6593	9087
R-squared	0.279	0.270	0.253	0.223

Note: individual data. Job average indicates the national average number of subtasks performed by workers or required cognitive skills in the occupation - industry combination. City size is defined by dummy variables. Small cities house less than 20,000 inhabitants, medium cities house between 20,000 and 100,000 inhabitants and large cities more than 100,000 inhabitants. Small cities are the reference group. Education is measured with dummy variables of the highest finished degree, high school graduation is the reference group. Standard errors in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Table 11: By education level

	Specialisation				Cognitive			
	No educ.	High-school	Op. college	Coll | Uni	No educ.	High-school	Op. college	Coll | Uni
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Job average	0.974***	0.998***	0.996***	1.001***	0.895***	0.952***	0.933***	0.985***
	[0.001]	[0.001]	[0.000]	[0.000]	[0.002]	[0.002]	[0.000]	[0.001]
Medium City	0.032***	0.130***	-0.081***	-0.334***	0.182***	0.020***	0.016***	-0.025***
	[0.009]	[0.009]	[0.003]	[0.004]	[0.002]	[0.003]	[0.001]	[0.001]
Large City	-0.005	-0.013	0.081***	-0.588***	0.163***	0.095***	0.004***	0.052***
	[0.009]	[0.008]	[0.003]	[0.004]	[0.002]	[0.002]	[0.001]	[0.001]
Age	-0.034***	0.016***	-0.024***	-0.015***	-0.007***	-0.001***	-0.004***	-0.006***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Female	-0.438***	-0.817***	-0.735***	-0.330***	0.088***	0.250***	0.106***	0.124***
	[0.007]	[0.007]	[0.003]	[0.003]	[0.002]	[0.002]	[0.000]	[0.001]
Native speaker	-0.266***	-0.283***	0.062***	0.576***	-0.233***	-0.029***	0.035***	0.027***
	[0.008]	[0.008]	[0.005]	[0.005]	[0.002]	[0.002]	[0.001]	[0.001]
Observations	511	608	9045	5506	511	608	9045	5506
R-squared	0.631	0.671	0.305	0.389	0.226	0.265	0.225	0.276

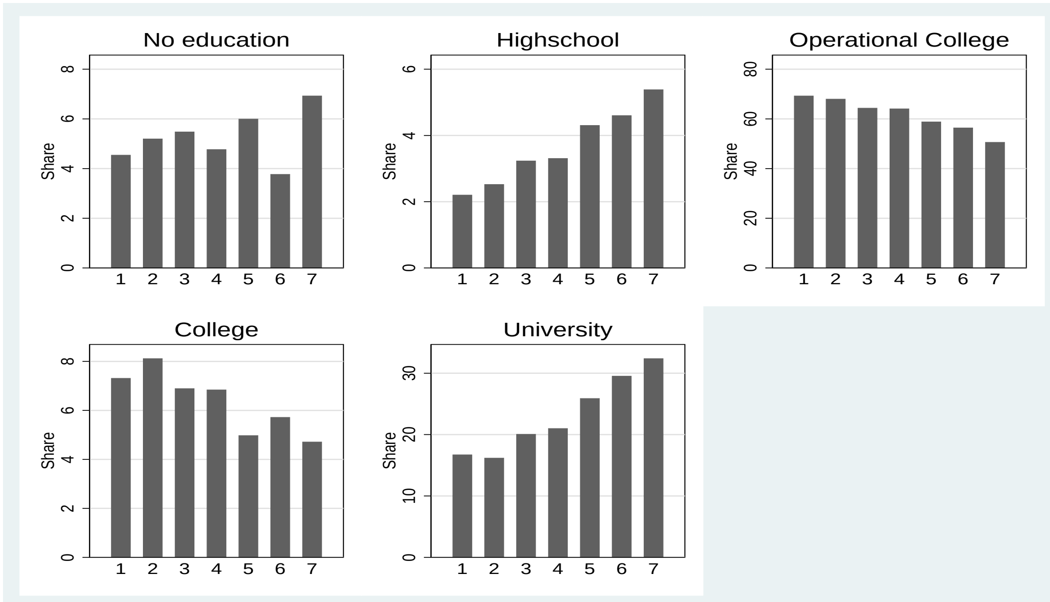
Note: individual data. Job average indicates the national average number of subtasks performed by workers in the occupation - industry combination. City size is defined by dummy variables. Small cities house less than 20,000 inhabitants, medium cities house between 20,000 and 100,000 inhabitants and large cities more than 100,000 inhabitants. Small cities are the reference group. Standard errors in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Table 12: Age groups

	Specialisation			Cognitive skills		
	< 35	35-50	>50	< 35	35-50	>50
	(1)	(2)	(3)	(4)	(5)	(6)
Job average	0.931***	0.960***	1.005***	0.975***	0.955***	0.912***
	[0.001]	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]
Medium city	0.095***	-0.026***	-0.131***	0.035***	0.025***	0.021***
	[0.006]	[0.004]	[0.006]	[0.001]	[0.001]	[0.001]
Large city	-0.013**	-0.119***	0.657***	0.090***	0.008***	-0.027***
	[0.006]	[0.004]	[0.006]	[0.001]	[0.001]	[0.001]
University * medium city	-0.158***	-0.603***	0.272***	0.139***	-0.097***	-0.073***
	[0.013]	[0.008]	[0.010]	[0.002]	[0.001]	[0.002]
University * large city	-0.091***	-0.790***	-1.268***	0.070***	0.045***	0.031***
	[0.011]	[0.007]	[0.010]	[0.002]	[0.001]	[0.002]
No education	-1.257***	-0.720***	-2.174***	0.387***	0.039***	0.225***
	[0.013]	[0.010]	[0.014]	[0.003]	[0.002]	[0.003]
Operational college	-0.238***	0.860***	-0.291***	0.011***	-0.117***	-0.004*
	[0.009]	[0.008]	[0.013]	[0.002]	[0.001]	[0.002]
College | University	0.358***	1.700***	0.745***	0.007***	-0.036***	0.066***
	[0.012]	[0.009]	[0.014]	[0.002]	[0.002]	[0.002]
Age	-0.046***	-0.006***	-0.038***	0.010***	-0.009***	-0.003***
	[0.001]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]
Female	-0.532***	-0.799***	-0.856***	0.092***	0.093***	0.181***
	[0.004]	[0.003]	[0.004]	[0.001]	[0.001]	[0.001]
Native speaker	0.217***	0.215***	-0.029***	-0.022***	0.011***	0.001
	[0.006]	[0.005]	[0.009]	[0.001]	[0.001]	[0.002]
Observations	3600	8518	3552	3600	8518	3552
R-squared	0.245	0.270	0.307	0.266	0.256	0.267

Note: individual data. Job average indicates the national average number of subtasks performed by workers in the occupation - industry combination. City size is defined by dummy variables. Small cities house less than 20,000 inhabitants, medium cities house between 20,000 and 100,000 inhabitants and large cities more than 100,000 inhabitants. Small cities are the reference group. Education is measured with dummy variables of the highest finished degree, high school graduation is the reference group. Standard errors in parentheses, * significant at the 10% level,** significant at the 5% level, *** significant at the 1% level.

Figure 1: Employment by education and city size



A Scarce occupations - Method

Following Duranton and Jayet, section 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 u and occupation k combination $s^j_{u,k}$:

$$s^j_{u,k} = a_0 + a_1(1/s^j_k) + a_2 N_u + a_3 N^j_u + a_4(1/s^j_k) * N_u + \epsilon^j_{u,k} \tag{8}$$

in which s^j_k represents the average employment share of occupation k in sector j. N_u is a city size dummy and N^j_u 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_k), for each sector and city size combination (b^j_u) and for each city size and scarcity level combination ($d_{m(u),r(j,k)}$). Scarcity is defined as the scarcity of occupation k within sector j, we measure this in terms of quartiles.

$$s^j_{u,k} = a^j_k + b^j_u + d_{m(u),r(j,k)} + \epsilon^j_{u,k} \tag{9}$$

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:

$$\pi^j_{u,k} = \frac{exp(Y^j_{u,k})}{\sum_{i=1,...U} \sum_{l=1,...K} exp(Y^j_{i,l})} \tag{10}$$

with: $Y^j_{u,k} = \alpha^j_u + \beta^j_k + \xi_{m(u),r(j,k)}$
For more detailed information we refer to the work of Duranton and Jayet (2011).