CPB Netherlands Bureau for Economic Policy Analysis

CPB Discussion Paper | 246

Town and city jobs: your job is different in another location

Suzanne Kok

Town and city jobs: your job is different in another location

Suzanne Kok*

May 2013

Abstract

This paper shows that a job contains a different task package in a large city than the same job in a small city. We set out a theoretical model of the division of labour across cities, which shows that both the division of labour and the skill demand increase with city size. Most datasets hinder an empirical analysis of such a model as they lack spatial variation in job content. Using individual German task data, we are able to empirically estimate our model and analyse spatial variations in task content of jobs. The estimations 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 demand more cognitive skills when they are performed in large cities. This spatial variation in job contents likely bias regional wage equations.

Keywords: Occupations, division of labour, tasks, 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. 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 (1988*a*) modelled the idea of Adam Smith that the division of labour is bound by the extent of the market. Cooperation in a larger

*University of Groningen and CPB Netherlands Bureau for Economic Policy Analysis (s.j.kok@cpb.nl).
I thank Steven Brakman, Harry Garretsen, Yannis Ioannides, Jan Rouwendal, William Strange and Bas ter Weel and participants at the ERSA conference 2012, the Nederlandse Economendag 2012 and the NARSC conference 2012 for useful comments and insights.

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 and 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 (1988*b*), 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), 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 paper 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 (1988*a*). 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 and 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 (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 more efficient. 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 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 (1988*b*) and Garicano and 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 and 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 and Dorn (Forthcoming) and Bacolod et al. (2010). Autor and 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 and Schönberg (2010) and Dustmann et al. (2009).

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 statistics. The empirical strategy is discussed in Section 4. Section 5 presents the results on the spatial variation in job content. In Section 6, further sensitivity analyses are presented. Section 7 concludes.

2 Spatial variation in job content

This sections 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.

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:

$$E_i = \int_{t_i=\delta_1}^{t_i=\delta_n} x_{i,t} dt. \tag{1}$$

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:

$$S_{i,t} = cx_{i,t}, \tag{2}$$

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 and 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:

$$x_t = \frac{a}{S_{i,t}} = \frac{a}{cx_{i,t}}, \tag{3}$$

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

[1] The consequences of this assumption for our empirical strategy are discussed in Section 4.

worker i is determined by both the time the worker devotes to the performance of the task and the amount of time it takes to produce one unit:

$$q_{i,t} = \frac{x_{i,t}}{x_t} = \frac{x_{i,t}}{a/S_{i,t}} = \frac{cx_{i,t}^2}{a}. \tag{4}$$

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:

$$q_i = \int_{t_i=\delta_1}^{t_i=\delta_n} q_{i,t} dt. \tag{5}$$

A worker is the 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:

$$Q = \int_0^1 q_t dt, \tag{6}$$

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]

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 (2) to (4) into the output function (equation(6)) it follows that:

$$Q = \int_0^1 \int_{i=1}^{i=I} \frac{cx_{i,t}^2}{a} di, \tag{7}$$

subject to equation (1). There are increasing returns to worker task input $x_{i,t}$. Worker productivity decreases with the subset of tasks they perform:

$$\frac{\partial q_i}{\partial \delta_i} < 0. \tag{8}$$

Full specialisation into one task may however be hindered by the production function (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:

$$\delta_i = \frac{T}{I}. \tag{9}$$

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

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

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 and Murphy (1992)). Section 4 discusses the consequences of coordination costs.

Figures 1 and 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 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 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 shows the relation between number of workers and produced tasks by worker for this example.

Figure 1: Market with 1 worker

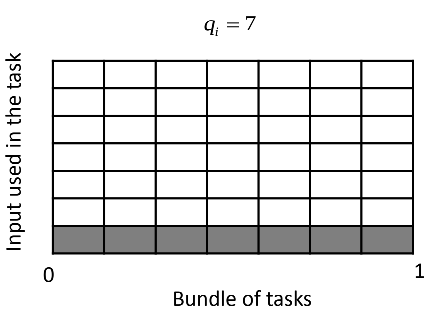
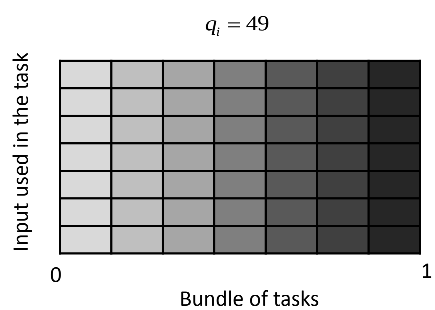


Figure 2: Market with 7 workers



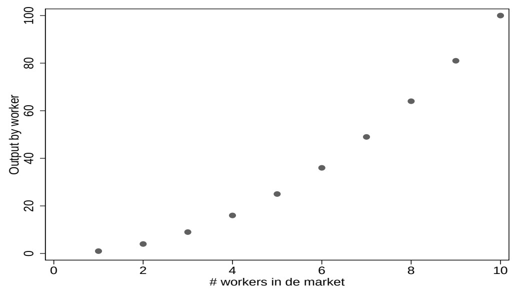
Each block represents a time unit, each worker is endowed with 7 time units. For simplicity $a = 2$ and $c = 2$. The different shades of grey indicate the time units of the different workers. The horizontal axis divides the good into 7 tasks, the vertical axis represents the time units a worker spends on each task. q_i is measured by equation (5).

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 (intermediate) 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.

Figure 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.

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 Data, indicators and descriptive statistics

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 1. The full list of tasks is displayed in Appendix A. 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

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

Table 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 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

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 and Schönberg (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 extensive discussion about the disadvantages of task information we refer to the work of Autor and Handel (Forthcoming), Acemoglu and 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 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.

[4] Section 4 discusses whether this biases our estimates.

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 optimizing 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 6.1 tests the sensitivity of the results towards the choice of indicators for job content.

3.3 Descriptive statistics

Tables 3 to 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 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 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 indicate.

Table 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

Table 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 15 in Appendix A. 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.

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 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 4 to 7 present the spatial distribution of jobs regarding their specialisation level and demanded skills. Figure 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 distribution. Workers in the large cities perform slightly fewer subtasks than workers in medium and small cities. The differences are however only modest. Figure 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 6 and 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 the large cities than in the medium and small cities.

Table 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 15 in Appendix A 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.

Lastly, we test whether scarce occupations are more likely to be performed in large cities and replicate the analysis of Duranton and Jayet (2011) for German cities. Cities produce many (intermediate) products with various demand threshold. Duranton and 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 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.

Figure 4: Distribution number of performed subtasks

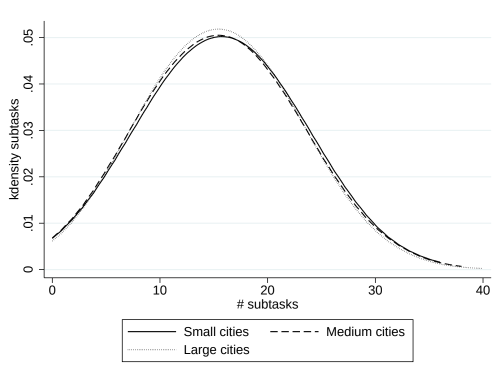


Figure 5: Distribution number of performed subtasks - high-skilled workers

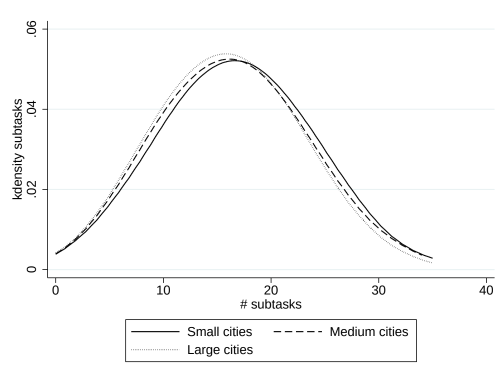


Table 5: Summary statistics - Industrial groups

	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 15 in Appendix A 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 depend on natural resources.

Table 6: Estimation results logit 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 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.

Figure 6: Distribution of demanded cognitive skills

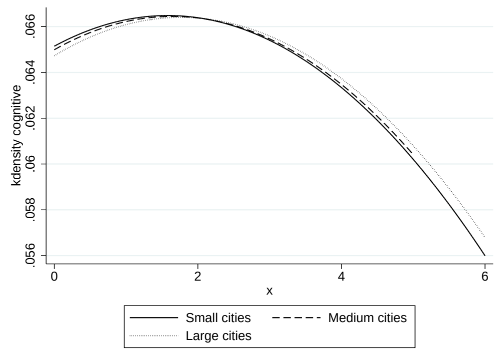
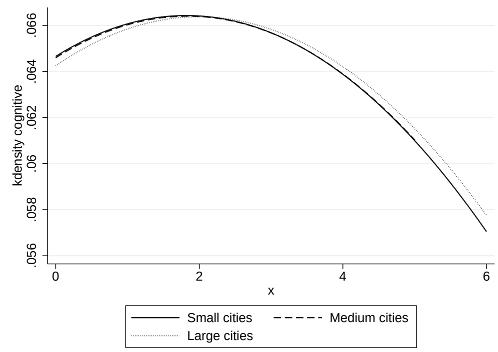


Figure 7: Distribution of demanded cognitive skills - high-skilled workers



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:

$$C_{i,o,j,l} = \alpha_1 + \alpha_2 C_{o,j} + \alpha_3 I_l + \alpha_4 S_i + \alpha_5 (I_l \cdot S_i) + \alpha_6 V_i + \epsilon_{i,o,j,l}. \tag{10}$$

$C_{i,o,j,z}$ 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 occupational level (Moulton (1990)). The observations are weighted by the size of occupations.

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 2 and the estimation as described in equation (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 6 displays the spatial variation of job contents across separate broad occupational groups. Furthermore, Section 6 shows the 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 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 (e.g. Glaeser and Maré (2001), Berry and 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 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 and 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. The coordination costs hinder workers to perform a unique subset of tasks and fully exploit the increasing returns. The model in Section 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 and 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 (1988*b*)). Unfortunately, we do not observe other city characteristics than size. To test the impact of industrial structure, analyses in Section 6 make the distinction between manufacturing and service sectors.

5 Job contents across cities

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^2 of 0.26. The correlation between number of subtasks performed by a worker and the job average is 0.45 (significant at the 1 percent level). We confirm the notion of Autor and

Handel (Forthcoming) that measures of task composition at the occupation 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 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. Although 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 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 15 in Appendix A 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 (see Becker and 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 dummy's 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 probably results from a large measurement error.

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 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 indicate. The jobs of young 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 to focus the estimation on the spatial variation within jobs. 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,

Table 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 15 in Appendix A 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.

the largest cities house the best workers (as discussed in Section 4). We expect the sorting of the more skilled workers to affect the skill demand as well.

6 Further analyses

Previous estimates include several assumptions which we test in this section with five robustness checks. A 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 6.1). Second, we deal with the regional measurement error in Section 6.2 and run estimations with different spatial units. Third, the possible impact of sorting of more able workers is discussed in Section 6.3. Section 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]

6.1 Indicators for specialisation and cognitive skills

Measuring task packages of jobs is challenging. As described in Section 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 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 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 of job tasks which likely results in a underestimation of the spatial variation in job tasks as well. The estimates in Table 9 indicate that measurement error does not drive our results.

6.2 Spatial units

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 three categories: small, medium and large cities. Column (3) of Table 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 using 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 ('Bundesland') provides us information on the possibilities to cooperate

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

Table 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. Job average indicates the national average number of subtasks or demanded cognitive skills in the occupation–industry combination. City size is defined by dummy variables, sizes are defined as in Table 2 for columns (3) and (4). In columns (5) and (6) 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. Table 15 in Appendix A 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.

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)).

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 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 sensitive 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 9 we include the job's specialisation level into the estimation for 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 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 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.

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

Table 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	-0.008	-0.057**	-0.047**	-0.047**	-0.005	0.055**	0.043**	0.043**
	[0.036]	[0.024]	[0.021]	[0.021]	[0.039]	[0.025]	[0.021]	[0.021]
Large city	-0.088**	-0.067**	-0.070***	-0.070***	0.018	0.092***	0.079***	0.079***
	[0.036]	[0.026]	[0.023]	[0.023]	[0.036]	[0.027]	[0.024]	[0.024]
Unskilled	-0.007***	-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]
Medium skilled	-0.241***	-0.189***	-0.198***	-0.198***	0.189***	0.195***	0.195***	0.195***
	[0.050]	[0.030]	[0.027]	[0.027]	[0.048]	[0.037]	[0.033]	[0.033]
High skilled	0.084	0.110**	0.104***	0.104***	0.001	-0.032	-0.022	-0.022
	[0.063]	[0.047]	[0.038]	[0.038]	[0.054]	[0.047]	[0.039]	[0.039]
Age	-0.155	-0.308**	-0.282***	-0.282***	0.121	0.003	0.029	0.029
	[0.101]	[0.122]	[0.092]	[0.092]	[0.082]	[0.083]	[0.065]	[0.065]
Female	0.219***	0.086*	0.109**	0.109**	-0.057	-0.144**	-0.127**	-0.127**
	[0.071]	[0.052]	[0.045]	[0.045]	[0.075]	[0.060]	[0.051]	[0.051]
Native speaker	0.487***	0.206***	0.254***	0.254***	0.003	-0.058	-0.046	-0.046
	[0.085]	[0.057]	[0.050]	[0.050]	[0.081]	[0.064]	[0.054]	[0.054]
Constant	0.122	-0.017	0.010	0.010	-0.051	0.445***	0.350***	0.350***
	[0.127]	[0.092]	[0.078]	[0.078]	[0.109]	[0.093]	[0.076]	[0.076]
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	0.032	0.017	0.019	0.019	0.010	0.016	0.015	0.015

Note: individual data. Table 15 in Appendix A 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.

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 low importance of tacit knowledge enables manufacturing firms to split up their production process easier between workers and even across space (Glaeser and 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 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 and Ponzetto (2010)).

Table 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 and Puga (2001) and Desmet and 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.

6.5 Learning and Experience

The task packages of workers vary with age (Autor and 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 specialisation. Here, we test whether the found results hold for all age groups.

Table 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.

Table 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 15 in Appendix A 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.

Table 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 15 in Appendix A 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.

The spatial variation in cognitive skill demand is the strongest for young workers in the beginning of their career.

Table 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 15 in Appendix A 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.

7 Concluding remarks

This paper 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.

References

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

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

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

Autor, D. and Handel, M. Forthcoming. 'Putting tasks to the test: Human capital, job tasks and wages'. *Journal of Labor Economics* .

Autor, D, Levy, F. and 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. and Strange, W. 2009. 'Skills in the city'. *Journal of Urban Economics* 65 (2), 136–153.

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

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 divergence 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 Urban Economics* 63 (2), 723–742.

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* 42 (6), 913–930.

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

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.

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

Dustmann, C., Ludsteck, J. and Schönberg, U. 2009. 'Revisting the German wage structure'. *Quarterly Journal of Economics* 124 (2), 809–842.

Eeckhout, J., Pinheiro, R. and 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.

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.

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

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

Glaeser, E. and Ponzetto, A. 2010. *Agglomeration economics*. The University of Chicago Press. chapter Did the death of distance hurt Detroit and help New York?, pp. 303 – 337.

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

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

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.

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

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.

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.

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

A Included tasks and variables

Table 14: 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	Is one and the same work cycle repeated
Repairing	Legal	Confronted with new problems that remain to be understood
Purchasing, selling	Project management	Process optimization / trying out new things
Transporting, storing, shipping	Medical or custodial	Are you failed or disturbed
Promoting, marketing, public relations	Lay-outing, designing, visualizing	Required output is stipulated in the minutest details
Organizing, 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	
Recognizing 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 15: 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 3.2	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 variables:	0.94	0.23
Job	Three-digit occupation and two-digit industry combination		

Table 16: 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 15 displays the definitions of the variables. P-values are in parentheses.

B Replication estimates of Duranton and Jayet (2011) for Germany

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 l and occupation o combination $E^j_{o,l}$:

$$E^j_{l,o} = a_0 + a_1(1/E^j_o) + a_2 N_l + a_3 N^j_l + a_4(1/E^j_o) * N_l + \epsilon^j_{l,o} \tag{11}$$

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.

$$E^j_{l,o} = a^j_o + b^j_l + d_{m(l),r(j,o)} + \epsilon^j_{l,o} \tag{12}$$

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:

$$\pi^j_{l,o} = \frac{exp(Y^j_{l,o})}{\sum_{i=1,...L} \sum_{l=1,...O} exp(Y^j_{i,l})} \tag{13}$$

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 and Jayet (2011).

Publisher:

CPB Netherlands Bureau for Economic Policy Analysis
P.O. Box 80510 | 2508 GM The Hague
T (070) 3383 380

June 2013 | ISBN 978-90-5833-603-3