CIRCLE

LUND UNIVERSITY

Paper no. 2014/06

The Sources of the Urban Wage Premium by Worker Skills

Martin Andersson (martin.andersson@circle.lu.se)
CIRCLE, Lund University and the Department of Industrial Economics and Management, Blekinge Institute of Technology

Johan Klaesson (johan.klaesson@ihh.hj.se)
Department of Economics, Finance and Statistics, Jönköping International Business School

Johan P Larsson (johan.p.larsson@ihh.hj.se)
Department of Economics, Finance and Statistics, Jönköping International Business School

This is a post-print version of a paper that has been accepted for publication in Papers in Regional Science. Please cite journal article.

This version: May 2014

Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE)
Lund University
P.O. Box 117, Sölvegatan 16, S-221 00 Lund, SWEDEN
http://www.circle.lu.se/publications

WP 2014/06

The Sources of the Urban Wage Premium by Worker Skills

Martin Andersson, Johan Klaesson and Johan P Larsson

Abstract

We estimate the respective importance of spatial sorting and agglomeration economies in explaining the urban wage premium for workers with different sets of skills. Sorting is the main source of the wage premium. Agglomeration economies are in general small, but are larger for workers with skills associated with non-routine job tasks. They also appear to involve human capital accumulation, as evidenced by the change in the wage of workers moving away from denser regions. For workers with routine jobs, agglomeration economies are virtually non-existent. Our results provide further evidence of spatial density bringing about productivity advantages primarily in contexts when problem-solving and interaction with others are important.

JEL codes: J24, J31, R12

Keywords: spatial sorting, selection, learning, non-routine skills, spatial wage disparities, density, agglomeration economies, innovation

Disclaimer: All the opinions expressed in this paper are the responsibility of the individual author or authors and do not necessarily represent the views of other CIRCLE researchers.

THE SOURCES OF THE URBAN WAGE PREMIUM BY WORKER SKILLS

- spatial sorting or agglomeration economies?

Martin Andersson[∅], Johan Klaesson[⊕] and Johan P Larsson[⊗]

ABSTRACT

We estimate the respective importance of spatial sorting and agglomeration economies in explaining the urban wage premium for workers with different sets of skills. Sorting is the main source of the wage premium. Agglomeration economies are in general small, but are larger for workers with skills associated with non-routine job tasks. They also appear to involve human capital accumulation, as evidenced by the change in the wage of workers moving away from denser regions. For workers with routine jobs, agglomeration economies are virtually non-existent. Our results provide further evidence of spatial density bringing about productivity advantages primarily in contexts when problem-solving and interaction with others are important.

JEL: J24, J31, R12
Keywords: spatial sorting, selection, learning, non-routine skills, spatial wage disparities, density, agglomeration economies, innovation

[∅] Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE), Lund University and the School of Management, Blekinge Institute of Technology (BTH); e-mail: martin.andersson@circle.lu.se
[⊕] Department of Economics, Finance and Statistics, Jönköping International Business School (JIBS), Jönköping, e-mail: johan.klaesson@ihh.hj.se
[⊗] Department of Economics, Finance and Statistics, Jönköping International Business School (JIBS), Jönköping; e-mail: johan.p.larsson@ihh.hj.se

Acknowledgements: We are grateful for constructive comments from three anonymous reviewers which improved the paper, as well as from seminar participants at the Max Planck Institute in Jena, CIRCLE at Lund University and the Jönköping International Business school. We also wish to thank Katarina Nilsson Hakkala (Aalto University), Fredrik Heyman (IFN) and Fredrik Sjöholm (Lund University) for sharing their data on non-routine job tasks by occupations. Martin Andersson acknowledges financial support from the Swedish research council FORMAS (Dnr 2011-80), as well as the Swedish Research Council (Linnaeus Grant No. 349200680) and the Swedish Governmental Agency for Innovation Systems (Grant agreement 2010-07370).

1. INTRODUCTION

Workers in urban areas of high spatial economic density earn higher wages than their counterparts in rural and more sparsely populated regions. Glaeser and Maré (2001) report that wages of urban workers in the United States are about 33% higher than their non-urban counterparts. Combes et al (2008) show that average wages in Paris are about 15% higher compared to other large French cities, 35% higher than in mid-sized cities and as much as 60% higher than in the rural areas of France. The empirical regularities of this kind are generally referred to as the 'urban wage premium' (UWP).

While the UWP is established as a general phenomenon, less is known about its sources and particularly whether it differs across workers with different sets of skills. This paper deals directly with these issues. We quantify the UWP for workers with different degrees of non-routine skills, respectively, and estimate the relative importance of spatial sorting and agglomeration economies in explaining the spatial wage disparities for each type of worker. The analyses in the paper provide empirical evidence on which type of skills are rewarded by density, and bear on the broader question of the contexts in which agglomeration is important.

1.1 Background and motivation

Recent research on the UWP has focused on two main lines of inquiry. One puts the issue of untangling the sources of the density wage premium at center stage, where a main question regards the respective importance of non-random spatial sorting of workers and agglomeration economies (Melo et al 2009, Combes et al 2011).[1] Spatial sorting refers to selection and explains the gap by more productive workers being more prone to locate in denser regions. This explanation involves no causal effect of spatial density on worker productivity. The existence of agglomeration economies, on the other hand, implies that density boosts worker productivity, for example through more efficient matching or faster human capital accumulation due to knowledge spillover phenomena (cf. Duranton and Puga 2004).[2] A general finding in this literature is that spatial sorting of workers is the main source of the UWP (Combes et al 2008).

The other line of inquiry focuses on differences in the magnitude of the UWP across workers with different sets of skills. Bacolod et al (2009) show that the UWP is not uniform across workers, but depends on workers skills. They maintain that the empirical literature on spatial wage differentials tends to equate skills with education levels, which does not capture horizontal differentiation of skills,

[1]There are several recent papers on this issue, such as Glaeser and Maré (2001), Combes et al (2008), Gould (2007), Yankow (2006), Wheeler (2006), Melo et al. (2009), Puga (2010), Combes et al (2011), Baum-Snow and Pavan (2012).
[2]Duranton and Puga (2004) discuss three families of micro-foundations of agglomeration economies – sharing, matching and learning.

such as cognitive, people and motor skills. The horizontal dimension of skills is yet important, they argue, as it may condition the ability to learn from the environment as well as the extent to which one benefits from matching and interaction with others – i.e. how much one benefits from agglomeration. Consistent with this, they show that it is primarily workers with jobs in which cognitive and people skills are important that enjoy an UWP.[3]

Skills that make workers better apt to benefit from agglomeration should yet not only be reflected in workers having a higher UWP, but also with regard to the importance of agglomeration economies as a source of the wage premium. Though a higher UWP among workers with certain skills could in principle be due to them being more prone to self-select towards urban regions than other groups of workers, the argument that learning and matching effects are stronger for workers with skills related to problem-solving and interaction with others is indeed not about spatial sorting. It is instead an argument emphasizing interactions between workers and their local environment that lead to productivity gains, i.e. agglomeration economies. The implication is that for workers with problem-solving and interaction skills, agglomeration economies should quantitatively be a more important source of the density wage premium compared to other types of workers. We test this prediction, thus bridging the two lines of inquiry on the UWP.

Available evidence on the magnitude and sources of the UWP by worker skills is limited. Bacolod et al (2009) employ data on a sample of US workers and estimate the effect of agglomeration on the hedonic price of cognitive, people and motor skills, respectively. While the question of the sources of the UWP for the various sets of skills is not spelled out explicitly in their paper, they isolate agglomeration economies by controlling for measures of worker ability as well as unobserved worker heterogeneity. Our approach is different in terms of both the measure of skills and identification strategy.

1.2 Measuring non-routine skills

We employ a longitudinal matched employer-employee dataset covering the full population of Swedish private sector workers over a seven-year period (2002-2008). These data do not include any direct information on worker skills, but do inform about the occupation according to the ISCO-88 classification scheme. To differentiate between skills we make use of a job-task classification scheme developed by Becker et al (2009), which reports the fraction of non-routine job tasks associated by each ISCO-88 occupation. Their original classification is based on a German work survey, which reports answers to 81 questions regarding workplace tool use by occupation. Tools are codified

[3]Gould (2007) as well as Möller and Haas (2003) also find that the UWP is significantly larger for better educated workers, and Baum-Snow and Pavan (2012) show that large cities foster human capital accumulation, especially for more highly skilled workers. None of these studies yet consider the horizontal dimension of skills emphasized by Bacolod et al (2009).

according to whether or not the use of a tool indicates non-routine tasks. Becker's et al (2009) classification is similar to that of Autor et al (2003) and Spitz-Oener (2006) in that occupations are linked to the involved share of routine versus non-routine tasks. We thus measure workers' non-routine skills by the extent of non-routine job tasks involved with the occupation of that worker.[4]

Autor et al (2003) define non-routine job tasks as tasks that cannot be performed by computers. In Becker et al (2009) non-routine tasks are defined as tasks characterized by non-repetitive work methods.[5] Such non-routine job tasks typically involve problem-solving and a lack of deductive rules and codifiable information (Becker et al 2009, Hakkala et al 2008). This corresponds to the way in which Autor et al (2003) conceptualize non-routine tasks. In relation to Bacolod's et al (2009) types of skills, cognitive and people skills are surely more important for non-routine job tasks, and we expect that workers with skills associated with non-routine job tasks should benefit more from density.

1.3 Identification strategy: spatial sorting and agglomeration economies

As we seek to quantify the sources of the wage premium of different workers, a key issue in our analysis is identification of spatial sorting and agglomeration economies, respectively. Recent work by e.g. Mion and Naticchioni (2009) and Combes et al (2008) illustrates that quantification of spatial sorting of workers depends crucially on the ability to account for worker heterogeneity, and that spatial sorting on unobservable skills account for a large fraction of spatial wage disparities. We quantify the importance of spatial sorting as a source of wage disparities by first estimating raw wage-density elasticities and then study their sensitivity to the inclusion of observable and unobservable (time-invariant) worker characteristics. Our data allow us to assess the role of several observable worker, employer and regional characteristics, as well as permanent worker heterogeneity. Agglomeration economies are indirectly quantified as a residual wage gap after accounting for spatial sorting. A significant wage-density elasticity that remains after controlling for spatial sorting of workers on observable and unobservable skills should in principle capture agglomeration economies (cf. Combes et al 2008).

The empirical strategy is straightforward: if sorting is important, we should observe that the (raw) wage premium drops significantly as we account for worker heterogeneity. The importance of agglomeration economies is instead reflected by the magnitude of the remainder wage-density elasticity. By undertaking these analyses for workers with skills associated with high and low fractions of non-routine job tasks, respectively, we empirically assess the magnitude and sources of the UWP for workers with different degrees of non-routine skills.

---

[4]This should reflect worker skills in the sense that workers with a job requiring a large fraction on non-routine tasks should have skills associated with non-routine work.

[5] Details of the classification as well as the correspondence between this and the job task classification in Spitz-Ooener (2006) can be found in Becker et al (2009).

To further probe our analysis of the sources of the UWP, we follow Glaeser and Maré (2001) and identify workers that move from urban to rural regions. The idea behind this is that agglomeration economies capture different effects, such as matching and learning. Learning implies that workers in cities may enjoy faster human capital accumulation, for instance through knowledge spillover phenomena (Glaeser 1999, Rauch 1993). Because accumulated human capital stays with the worker, the advantages of having worked in a larger dense region should remain while moving away. Static agglomeration economies, on the other hand, should be lost upon moving away from the agglomeration (cf. De La Roca and Puga 2012).[6] To test the argument that workers with non-routine skills are more apt to learn from their environment, we identify routine as well as non-routine workers that move away from dense agglomerations and test, for each category of worker, if their wage drops or remains upon moving. This is a straightforward and simple test of whether learning by workers depends on the skills, and the hypothesis is that non-routine workers show stronger learning.[7]

The paper includes some additional features further separating it from previous studies. Many of the analyses of the UWP separate urban from rural regions with a dichotomous variable or employ continuous measures of regional density that only account for the internal density of regions. The analysis in this paper recognizes the message emphasized by Irwin et al (2010), i.e. that there is interdependence across regions that produces a continuum from dense urban regions to more remote rural ones. Our measure of density is access to economic 'mass', as measured by each region's exponentially travel time-distance-weighed access to total wage earnings inside the region as well as to all other regions. The total density of a location is decomposed into three spatially distance-weighed components: (i) municipal, (ii) regional and (iv) extra-regional. This decomposition allows us to obtain a parameter estimate for each aggregation level, making it possible to assess the importance of each component, such as the relative importance of the municipal and the regional density. With these measures the total density of a region is not only dependent on its internal characteristics, but also on the characteristics of surrounding regions and its travel time-distance to those regions. This captures interdependence between regions. Moreover, most of the existing analyses have been conducted on countries hosting large metropolitan areas, such as the US (Glaeser and Maré 2001, Gould 2007), Germany (Möller and Haas 2003) and France (Combes et al 2008). Sweden is a small and generally sparsely populated country (around 9 million inhabitants on a total land area of about 410 000 km^2). Most cities and urban areas in the country are small in an international context, and only three cities

[6]Models of matching effects in thick markets suggest that the average quality of each match is higher in agglomerations (cf. Hesley and Strange 1990, Kim 1990). Such an agglomeration economy surely does not follow workers.

[7] We are yet cautious in drawing strong conclusions from the analyses of the wages of movers as we are not able to fully account for the endogeneity issue raised by Gould (2007); changes in wages form moving may be correlated with changes in the quality of opportunities in different regions.

may, with a generous standard, be labeled metropolitan.[8] An analysis of Sweden thus constitutes a conspicuous contrast to existing analyses on countries with big urban areas such as New York and Paris.

1.4 Main findings

We find sharp differences between workers with non-routine and routine skills in terms of the magnitude of the spatial wage disparities as well as their sources. Workers with skills associated with non-routine job tasks enjoy an unadjusted wage-density elasticity of about three percent. For these workers, agglomeration economies are significant, though quantitatively of much smaller importance than spatial sorting. After controlling for observed and unobserved worker heterogeneity we find that a doubling of either municipal or regional density yields a wage increase in the order of .5 percent. Non-routine workers also appear to be better apt to accumulate human capital, as evidenced by that workers that move away from denser regions keep (or increase) their wage upon moving. For workers with skills associated routine job tasks on the other hand, agglomeration economies appear to be non-existent.

1.5 Outline

The rest of the paper is organized as follows: Section 2 presents the data, defines variables and also provides the big picture regarding wages, education levels and skills in the economic geography. Section 3 describes our empirical strategy, focusing on how we empirically assess the relative importance of spatial sorting and agglomeration economies as sources of the UWP. Section 4 presents the results and section 5 concludes.

2. DATA, VARIABLES AND DESCRIPTIVES

2.1 Data

We use a matched employer-employee audited register dataset, maintained by Statistics Sweden. The data comprise all employees in Sweden during the period 2002 to 2008. By construction of the data, employees are assigned to their work establishment (and thus sector, occupation and location) in the month of November each year. Though the data span all sectors of the economy, we exclude all public sector employees and workers in the agriculture and mining industries. This isolates workers whose wage formation is determined by market outcomes and workers in sectors whose locations are not directly linked to natural resources. As we are interested in labor income, we also exclude workers whose primary income comes from self-employment. Workers in our data are in the age interval 20-64.

[8] If we for instance apply the 'big city' classification in Yankow (2006), only one metropolitan area in Sweden (Stockholm) would barely pass the bar.

This leaves us with a panel containing about 2.4 million employees with a mean population size of just short of 2 million yearly observations. The discrepancy between the number of individuals and the number of observations per year is an effect of the cut-off values created by the age interval and to a lesser extent by increased labor force participation in later stages of the reporting period.

The data inform about several characteristics of each employee and their employer. For employees we have information such as education (length and specialization), sex, age, wage income and immigrant status. Employee characteristics include basic observables such as sector and employment size.

2.2 Variables and classification of non-routine job tasks

Density

Our variable of main interest is spatial economic density. Many studies of the UWP distinguish urban dense areas from rural ones by a dichotomous indicator variable based on some threshold value of e.g. population size. Alternatively, they consider a continuous indicator measuring the internal density of each region, commonly employment per square kilometer. The density measure employed in this paper is different.

We define density in a way akin to Harris' (1957) classic measure of market potential. The basic spatial unit in our analysis is the municipality of which there are 290 in Sweden. Specifically, the data inform about in which municipality each workers' employer is situated. These spatial units are in general of limited size and there is significant commuting and other types of interaction across municipal borders. Many of the spillover effects alluded to in the literature on agglomeration economies and human capital spillovers are thus likely to transcend municipal borders, especially as they may be mediated by labor market mobility (cf. Andersson and Thulin 2013). The same applies from the viewpoint of spatial sorting. When workers choose where to operate in space, they most likely consider characteristics of an integrated labor market, which in general comprises more than one municipality. We may thus expect interdependencies between municipalities, such that it is not only the internal density of municipalities that matter, but also the surroundings. On these grounds we employ an accessibility approach. One can think of the total density of a municipality *r* as the sum of municipal, regional and extra-regional accessibility to total wage-earnings, *W*:

$$De_r^{Tot} = De_r^M + De_r^R + De_r^E \tag{1}$$

$De_r^M = W_r \exp\{-\lambda_M t_{rr}\}$, municipal accessibility to total wage earnings of municipality *r*

$De_r^R = \sum_{k \in R_r} W_k \exp\{-\lambda_R t_{rk}\}$, regional accessibility to total wage earnings of municipality *r*

$De_r^E = \sum_{l \notin R_r} W_l \exp\{-\lambda_E t_{rl}\}$, extra-regional accessibility to total wage earnings of municipality *r*

Total wage earnings reflect the magnitude of economic activities (or economic mass) and accessibility to economic activity is our measure of spatial economic density. Municipal density is simply each municipality's total wage earnings weighed exponentially with travel time-distances by car between zones within the municipality. Regional accessibility is defined in a similar way but here we sum the municipality's access to every other municipality belonging to the same local labor market region.[9] Extra-regional accessibility is the sum of its accessibility to all municipalities outside the region. The distance-decay parameter λ takes on three different values for municipal, regional and extra-regional accessibility, respectively. These parameter values are based on observed commuting behavior of workers, and are estimated for Swedish municipalities by Johansson et al (2003) using doubly constrained gravity models.

The accessibility approach recognizes that the density of a municipality is built up through a geographic continuum where the contribution of other places' economic activities falls as travel-time distances increase.[10] Thereby, the measure is consistent with Tobler's (1970) '1st law of geography': everything is related, but near things are more related than distant things. Because of the nature of the exponential distance-decay function, the contribution of municipalities far away is small but remains positive. In terms of an urban-rural dichotomy, the accessibility formulation recognizes interdependence across places where there is a continuum from dense urban regions to more remote rural ones (cf. Irwin et al 2010).

In the empirical analysis we include De_r^M, De_r^R and De_r^E as three distinct independent variables. This allows us to assess which type of density that matters. In general we expect density effects to primarily pertain to the local labor market region in which the workers work, i.e. De_r^M and De_r^R.

Controls – observable characteristics

We control for several characteristics of workers and employers that may influence a worker's wage. The observable characteristics that we include in the analysis are presented and defined in Table 1. Experience and its squared value are standard control variables and in accordance with previous literature we expect that wages increase with experience but at a diminishing rate. Years of schooling is assumed to have a positive influence on a worker's wage.

[9] Local labor market regions comprise a number of municipalities forming an integrated labor market, and are delineated based on the intensity of inter-municipality commuting flows.
[10] This also alleviates potential problems with spatial autocorrelation (Andersson and Gråsjö 2009).

Table 1. Variables, definitions and expected sign.

Variable	Definition	Expected sign
Wage	The total wage earnings of a worker during a year[11]	n.a
Experience	The employee's *age* minus *years of schooling* minus *6*. This definition follows Rauch (1993).	+
Experience squared	Same as above but squared.	-
Schooling	Theoretical years of schooling.	+
Education specialization	Dummies for different education specializations, defined according to the 1-digit SUN2000 classification, which is based on ISCED 1997.	n.a
Immigrant	A dummy which is 1 if the worker is a first generation immigrant, 0 otherwise.	-
Sex	A dummy which is one if the worker is male, 0 otherwise.	+
Tenure	The number of years the worker has been employed at her current workplace. Max tenure is the observational year minus 2001, as we have no information prior to 2001.	+
Number of prior employers	The number of different employers the worker has had since 2001.	-
Job change	A dummy which is 1 if the worker changed occupation between year *t* and *t*-1.	-
Log of number of employees	The natural logarithm of the total number of employees at the workplace at which the employer is employed.	+
Sector affiliation	Dummies for different sectors at the level of 2-digit NACE sectors.	n.a
Municipal density	Exponentially distance-weighed accessibility to wage sums in the municipality the worker works in.	+

[11]The individuals included are workers who are primarily wage laborers, but like most other studies using audited full population register data where wage incomes are drawn from tax declarations, we lack information on the number of hours worked. While this represents the best information available, we recognize that using yearly wages may be a source of bias in an OLS setting under the assumption that workers in dense areas systematically work longer hours than workers in sparse areas and consequently make higher yearly wages. In a fixed effects setting, this is a smaller problem. The reason is that a bias can in this case only arise if workers in dense areas systematically work increasingly longer hours, relative to workers in sparse areas during the reporting period, or that workers moving to more dense regions increase their working hours by moving. In the empirical analyses that follow, every model specification further includes region-year effects, which means that any systematic region-specific trends by which workers in certain regions increase working hours over time is picked up.

Regional density	Exponentially distance-weighed accessibility to wage sums to all municipalities in the local labor market region the municipality belongs to.	+
Extra-regional density	Exponentially distance-weighed accessibility to wage sums to all municipalities in Sweden except those belonging to the municipality's local labor market region.	+

Note: All variables are based on audited register data maintained by Statistics Sweden. Accessibility calculations based on travel time distances by car between municipalities. Travel time distances by car are obtained from the Swedish Road Administration.

We also include a set of dummy variables reflecting the educational specialization of the worker. These are defined at the 1-digit SUN2000 classification system in Sweden, which corresponds to the 1997 International Standard Classification of Education (ISCED). This leaves us with nine dummy variables reflecting the educational specialization of each worker. We have a priori no clear idea of how different educational specializations may influence a worker's wage, but we acknowledge that they reflect potentially relevant characteristics of the workers. The analysis further includes immigrant and sex dummies. The former is one if the worker is a first generation immigrant and the latter is 1 if the worker is male. The general finding in the literature is that immigrants have lower average wages whereas males have higher average wages than females.

Tenure is an important variable in labor market analyses and is assumed to reflect the quality of the match between the worker and her workplace (Farber 1994). On these grounds, we expect that tenure is positively associated with a worker's wage. We define tenure as the number of years the worker has stayed with her current workplace. Due to data availability reasons, max tenure is the observation year minus 2001 because we have no information prior to 2001. In addition to tenure we also include the number of prior employers and a dummy for whether the worker switched jobs between year t and t-1. Both these variables may reflect workers in search of a good match in the labor market, why we expect them to be negatively associated with wages.

The employment size of the establishment at which the workers are employed is another important determinant of wages. Ample studies in labor market economics show that larger firms pay higher wages (Oi and Idson 1999).[12] We expect that establishment size has a positive influence on wages. Furthermore, we include dummy variables to account for the possibility that wages may depend on the sector in which a worker is employed. The analysis includes one sector dummy for each 2-digit sector

[12]This is often explained by larger firms being better equipped than smaller firms in terms of resources and productivity, as well as by behavioral arguments. The latter includes that larger firms may be more apt to adopt discretionary wage policies and paying efficiency wages to deter shirking.

amongst NACE sectors 15-74.[13] The sector of a worker is determined by the sector affiliation of the establishment he or she is employed by.

Measuring non-routine job tasks

The data on the fraction of non-routine job tasks by occupation originate from Becker et al (2009) and details on the construction of the data as well as their various robustness checks are documented therein.[14] They classify answers in a German qualification and career survey for 1998/1999, undertaken by the German Federal Institute for Vocational Training and the research institute of the German Federal Labor Agency. It tracks the usage of 81 different tools in a multitude of occupations. Becker et al (2009) classify different tools according to their relation to non-routine tasks (non-repetitive work methods). The different tasks are then mapped to ISCO-88 standardized occupations. For each 2-digit occupation, the degree of non-routine tasks is then computed as the ratio between the average number of non-routine tasks in the occupation and the maximum number in any occupation, and the numbers are then standardized so that the fraction of non-routine tasks in an occupation varies between 0 and 1.

In Table 2 we follow Hakkala et al (2008) and present the fraction on non-routine job tasks for each occupation at the 2-digit ISCO-88.[15] The general picture is that science-based, engineering and corporate management occupations have the highest fraction of non-routine tasks. A low degree of non-routine job tasks are found in occupations related to agriculture, fishing, extraction sectors and simpler transport services. The patterns reported in the table confirm that non-routine job tasks typically involve problem-solving with a general lack of deductive rules and codifiable information (Hakkala et al 2008). The occupations with high fractions of non-routine tasks are also jobs in which cognitive and people skills should be important (cf. Bacolod et al 2009).

[13]In the analyses presented in the sequel, we have also tested if the results depend on the level at which the sector dummies are defined. Our results are robust to using sector dummies at the 2, 3, 4 or 5 digit level.

[14]They also classify jobs according to the extent it involves interaction. There is yet considerable overlap between the two classifications, where non-routine tasks tend to involve interaction tasks. In all analyses presented in the sequel, we have also tested this classification and results are robust. We choose the non-routine classification as it emphasize jobs in which cognitive and people skills should be important.

[15]Hakkala et al (2008) use the task data mapped to ISCO-88 occupations developed by Becker et al (2009) in analysis of how multinational activities influence demand for different job tasks.

Table 2. The fraction of non-routine tasks in different 2-digit occupations according to ISCO-88.

Occupation title	Fraction non-routine tasks (%)
Physical, mathematical and engineering science professionals	100.0
Life science and health professionals	90.4
Physical and engineering science associate professionals	79.7
Corporate managers	78.4
Other professionals	63.0
Teaching professionals	61.2
Life science and health associate professionals	56.3
Legislators and senior officials	54.4
Other associate professionals	52.7
Office clerks	52.1
General managers	46.6
Stationary-plant and related operators	43.6
Metal, machinery and related trades workers	41.6
Precision, handicraft, printing and related trades workers	39.8
Teaching associate professionals	36.1
Personal and protective services workers	32.0
Customer services clerks	27.1
Extraction and building trades workers	21.4
Machine operators and assemblers	18.8
Other craft and related trades workers	17.7
Market-oriented skilled agricultural and fishery workers	10.8
Models, salespersons and demonstrators	8.1
Drivers and mobile-plant operators	6.3
Laborers in mining, construction, manufacturing and transport	2.5
Agricultural, fishery and related laborers	0.9

Note: Based on Hakkala et al (2008) using task data developed by Becker et al (2009).

2.3 Wages, education levels and skills in the Swedish economic geography

Table 3 presents the mean wage, fraction of graduates, mean experience and the fraction of workers working in any of the three largest regions in Sweden for all workers as well as for occupations with high and low fractions of non-routine job tasks, respectively.[16] About one third of all workers in the population work in the three largest regions and about 15 percent are university graduates. Workers with jobs requiring more non-routine tasks are much better educated and are the ones most prone to work in a metropolitan area. Roughly 36 percent of all workers with non-routine jobs work in a metropolitan area compared to 19 percent for workers with less non-routine tasks in their job. The mean wage of workers with jobs associated with high fractions of non-routine tasks is also higher than for other types of jobs.

[16]High fraction non-routine jobs are those occupations with fraction non-routine tasks above the mean fraction across all occupations. Low fraction non-routine jobs are those whose fraction of non-routine tasks is below the mean.

Table 3. Key figures divided by fraction of non-routine work tasks.

Job type	Mean wage (EUR)	Graduate share	Mean experience	Metropolitan share
All types of professions	*29 698*	*15%*	*22*	*27%*
High fraction non-routine tasks	36 683	28%	23	36%
Low fraction non-routine tasks	23 088	3%	21	19%

Note: Graduate share is the fraction of workers with a university education of at least three years. Metropolitan share is the fraction of workers that work in three biggest labor market regions: Stockholm, Göteborg and Malmö. Wages converted to EUR using the 2008 exchange rate between SEK and EUR of 9.68. High (low) fraction non-routine jobs are those with fraction non-routine tasks above (below) the mean fraction across all occupations (see Table 2).

The unadjusted wage differential between metropolitan and non-metropolitan workers overall and for jobs with high and low fractions of non-routine tasks is presented in Table 4. For the private sector as a whole, the raw wage differential between metropolitan and non-metropolitan workers amounts to just over 20 percent.

The urban-rural wage gap yet appears to depend crucially on the type of job. The difference is substantially larger for occupations with high fraction non-routine tasks (20%) whereas the same 'raw' wage differential is negative but small for occupations with low fractions of non-routine tasks. These patterns are broadly consistent with the recent literature (e.g. Bacolod et al 2009, Gould 2007), and suggest that spatial sorting with regard to type of jobs is one reason for the (unadjusted) overall UWP.

Table 4. Mean wages (2008) and unadjusted wage gap between metropolitan and non-metropolitan workers.

Job type	Metropolitan wage (EUR)	Non metropolitan wage (EUR)	Wage differential
All types of professions	*34 417*	*27 926*	*23%*
High fraction non-routine tasks	41 024	34 245	20%
Low fraction non-routine tasks	22 634	23 195	-2%

Note: The metropolitan areas are defined as the three biggest labor market regions: Stockholm, Göteborg and Malmö. Wages converted to EUR using the 2008 exchange rate between SEK and EUR of 9.68. High (low) fraction non-routine jobs are those with fraction non-routine tasks above (below) the mean fraction across all occupations (see Table 2).

One reason for the described wage differences between workers in metropolitan and non-metropolitan regions may of course be that better educated workers are more inclined to move to bigger cities. Indeed, highly educated individuals tend to agglomerate in cities, for instance since specialized workers are better matched with employers where markets are thick (Strange 2009) and since highly educated individuals may self-select to cities where consumption amenities are abundant (Lee 2010).[17] Workers with higher education levels indeed have higher wages, and the graduate share in the metropolitan areas was 28 percent in 2008, while it was 13 percent in other areas.

[17]There is a large literature on the extent to which the location of educated workers is driven by amenities or productivity (e.g. Moretti 2008) but this issue is not the main focus and beyond the scope of this paper.

The subsequent empirical analysis focuses on the relationship between density as measured by accessibility to total wage earnings and workers' wages. Figure 1 plots the logarithmic relationship between mean wages in our population and our (summed up) density measure. It is clearly the case that workers in denser municipalities have higher average wage.

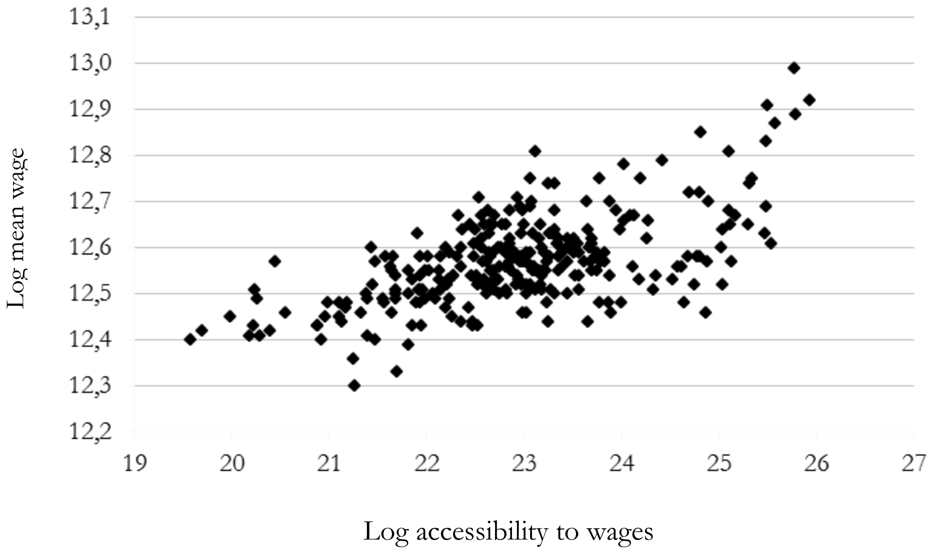

Figure 1. *The relationship between mean wages (log) and accessibility to total wage earnings (log) across Swedish municipalities in 2008.*

A simple OLS estimation of the log of density on the log of average wages across municipalities in Sweden using the data in Figure 2 yields the following results (*t*-values beneath parameter estimates, *N*=290):

$$\ln w_r = \underset{74.4}{11.4} + \underset{13.7}{.05} \ln De^{Tot}_r + \varepsilon_r \tag{2a}$$

$$\ln w_r = \underset{58.7}{6.6} + \underset{10.9}{.04} \ln De^{M}_r + \underset{3.9}{.01} \ln De^{R}_r + \underset{3.3}{.01} \ln De^{E}_r + \varepsilon_r \tag{2b}$$

where the estimation in (2b) separates between the three components of the total density of municipalities (see equation 1). The estimates in (2a) show that 10% higher density is associated with about .5% higher wages. The decomposition of the total density in (2b) shows that the municipal density is responsible for the bulk of this relationship with an estimated coefficient of 0.04. The density of the local labor market region and extra-regional density contribute with a significantly smaller share amounting to about 0.01 each.

3. EMPIRICAL STRATEGY

The baseline empirical model is as follows:

$$\ln w_{irt} = \alpha + \beta_1 \ln De_{rt}^{M} + \beta_2 \ln De_{rt}^{R} + \beta_3 \ln De_{rt}^{E} + \dots$$
$$\dots + \sum_{R=1}^{81} \gamma_R D_R + \sum_{t=1}^{T} \lambda_t D_t + \sum_{tR=1}^{TR} \sigma_{tR} \left(D_t \times D_R \right) + \mathbf{Z}'\boldsymbol{\gamma} + \varepsilon_{irt} \qquad (3)$$

where w_{irt} is the wage earnings of individual *i* at time *t* working in municipality *r*. De_{rt}^{M}, De_{rt}^{R} and De_{rt}^{E} represent municipal, regional and extra-regional accessibility to wage earnings, respectively. The baseline model always includes year dummies (D_t), dummies for local labor market regions (D_R) as well as time dummies interacted with dummies for local labor market regions $\left(D_t \times D_R \right)$. Year dummies are intended to account for general business cycle effects, and region dummies are included to capture region-specific effects. The region-year effects account for any region-specific time-varying shocks shared by all workers in the same local labor market region.[18] Previous work, e.g Moretti (2004) emphasizes the importance of accounting for both region and region-year effects. $\mathbf{Z}$ is a matrix of control variables. ε_{irt} is an error term. Our main interest is in the β parameters.

To quantify the importance of spatial sorting we start by estimating 'raw' wage-density elasticities, indicating how wages of private sector workers correlate overall with our three density measures. We obtain these raw elasticties by estimating the model in (3) using pooled OLS without any controls besides year, region and region-year dummies. We then estimate four additional models, while keeping the raw wage-density elasticities as points of reference.

In the first estimation we add standard Mincerian observable worker characteristics in the form of years of schooling, experience, sex, immigrant status as well as dummies reflecting different education specializations (Mincer 1974). The second estimation adds labor market information of each worker, i.e. tenure, number of prior employers, a dummy for whether the worker's current occupation is new for the worker and employer size. This second specification also includes two-digit NACE industry dummies to capture differences in general wage levels across industries. As the reference estimation, these two specifications are estimated with pooled OLS. This means that identification of the wage-density elasticities is based on differences across workers in municipalities of varying densities, while

[18]To be precise, the region-year dummies account for shocks over time that are common for all employees working in municipalities belonging to the same local labor market region *R*. We choose the local labor market region as aggregation level for the region-specific shocks as the labor market regions represent integrated local labor markets and comprise several municipalities connected through intense commuting flows. There are 81 local labor market regions in Sweden.

controlling for observable worker and employer characteristics as well as time, industry and region-year effects.

The two additional specifications exploit the panel structure of the data and add worker fixed effects (FE). These worker FE fully absorb any permanent heterogeneity at the worker, employer, industry or municipality level. Due to the within transformation of the FE estimator, identification of the wage-density elasticities is now based on changes over time in the three density measures. As the within variation of each respective density measure is limited, the parameters of the density variables are primarily identified based on workers who over years move between municipalities of varying densities.[19] The first FE model is the basic model in equation (3) augmented with worker FE but excluding any other controls besides year and region-year dummies. The second one adds time-varying worker and employer characteristics, including industry dummies. The inclusion of FE worker effects means that these observables are also identified from changes over time.[20]

This empirical set-up allows us to quantify how sensitive the estimated wage-density elasticities are to spatial sorting on observable and unobservable worker characteristics. In view of previous research such as Combes et al (2008), we expect that the wage-density elasticities are significantly reduced when accounting for worker characteristics, especially unobservable permanent worker heterogeneity. Any remainder significant wage-density elasticities should reflect agglomeration economies.

We further isolate workers who move from high to low density regions. In our empirical context we accomplish this in a straightforward manner by identifying workers who move from any of Sweden's three metropolitan regions (Stockholm, Göteborg and Malmö) to any other place in Sweden. We then estimate whether they reduce or keep their wage upon leaving a metropolitan region, using both pooled OLS and FE models. The idea behind this is to test for learning effects in the form of human capital accumulation effects (Glaeser and Maré 2001, De La Roca and Puga 2012): if workers gain human capital in cities, the advantages of having worked in a larger and denser city should remain while moving away.

We systematically apply the empirical strategy described above for workers with occupations associated high and low fractions of non-routine job tasks, respectively. We thus split the sample of workers in two groups; one with workers having occupations with a fraction of non-routine tasks above the mean fraction for all occupations, and one with a non-routine job task fraction below the

[19]For each of density variable, the within variation is substantially smaller than the between variation. For municipal, regional and extra-regional density the between variation is about 2.7, 3.7and 2.9 times larger than the within variance, respectively.

[20]In a similar way, the regions-specific effects (D_R) are identified from workers that move between local labor market regions over time.

mean (see Table 2). This allows us to identify differences in the importance of spatial sorting and agglomeration economies between the two groups in a straightforward way.[21]

4. RESULTS

Table 6 presents estimation results for all private sector workers in Sweden. Starting from the left, the first three specifications are pooled OLS estimations and the last two are panel estimations with worker fixed effects. The municipal and regional densities are significant and positive in all specifications. Workers earn more in denser regions. It is thus not only the density of the municipality that matter, but also the density of the wider local labor market region in which the worker operates. This is in line with expectations as labor market regions represent integrated labor markets and consist of municipalities between which there is intense interaction. The extra-regional density is negative and significant, indicating that if the surroundings of a labor market region grow it has a negative impact on wages in the region, all else equal. This may be understood as an effect from lagging behind the surroundings.

The raw unadjusted wage-density elasticity is about .03 for municipal and regional density, respectively. Taken together, they correspond broadly with the estimates reported by Ciccone and Hall (1996), who find that a doubling of density is associated with about six percent higher productivity. The wage-density elasticities are also rather insensitive to observable worker characteristics. In the Mincerian model which adds years of schooling, experience as well as dummies for sex, immigrants and education specialization, the estimated wage-density elasticties for municipal and regional density only falls marginally – from .03 to about 0.2.

The estimated parameters change only slightly from adding indicators for labor market status and employer characteristics (full OLS model). These patterns suggest that spatial sorting of workers on basic observable worker and employer characteristics is not a quantitatively important source of the raw wage-density relationship.

The picture yet changes as we control for permanent worker heterogeneity with worker fixed effects. The second column from the right shows the results with the raw specification with worker fixed effects, i.e. excluding any other controls besides year and region-year dummies. A comparison of the wage-density elasticties in this specification with the ones obtained with 'raw OLS' (second column from the left) illustrates what worker fixed effects means for the magnitude of the estimated wage-density elasticities.

[21] An alternative strategy would be to include the fraction of non-routine job tasks as a separate independent variable. The pooled OLS estimations would then identify its effect through differences across workers, whereas identification with the FE estimator would be based on workers that shift occupations over time (the fraction of non-routine job tasks of an occupation is time-invariant). We have considered this strategy as well and the findings reported in the sequel are robust to this alternative approach.

Table 6. The relationship between spatial economic density and wages, all private sector workers.

	Raw OLS	**Mincerian OLS**	**Full OLS**	**Raw with worker FE**	**Full with worker FE**
Municipal density (log)	0.0326***	0.0218***	0.0205***	0.00773***	0.00538***
	(0.00322)	(0.00224)	(0.00123)	(0.000242)	(0.000242)
Regional density (log)	0.0335***	0.0218***	0.0195***	0.00790***	0.00522***
	(0.00777)	(0.00425)	(0.00641)	(0.000518)	(0.000514)
Extra-regional density (log)	-0.0323*	-0.0221	-0.0248***	-0.0127***	-0.00797***
	(0.0185)	(0.0139)	(0.00788)	(0.000679)	(0.000674)
Years of schooling		0.0930***	0.0823***		0.117***
		(0.00468)	(0.00300)		(0.0190)
Experience		0.0503***	0.0408***		0.0587***
		(0.00326)	(0.00279)		(0.0190)
Experience^2		-0.000781***	-0.000635***		-0.000745***
		(5.42e-05)	(4.64e-05)		(2.84e-06)
Immigrant (dummy)		-0.136***	-0.108***		
		(0.00737)	(0.00409)		
Male (dummy)		0.351***	0.330***		
		(0.00947)	(0.00425)		
Tenure			0.0176***		-0.0109***
			(0.000458)		(0.000127)
Number of prior employees			-0.0120***		-0.0189***
			(0.00286)		(0.000176)
New occupation (dummy)			-0.0930***		-0.0268***
			(0.00222)		(0.000366)
Employer size (log)			0.0257***		0.0183***
			(0.00225)		(0.000175)
Year dummies	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes
Region*Year effects	Yes	Yes	Yes	Yes	Yes
Education type dummies	No	Yes	Yes	No	Yes
Industry dummies	No	No	Yes	No	Yes
Observations	12,367,700	12,367,700	12,367,700	12,367,700	12,367,700
Individuals	2,681,164	2,681,164	2,681,164	2,681,164	2,681,164
R-squared	0.031	0.248	0.288	0.059	0.078

Note: The table reports estimates of wage-density elasticities for private sector workers in Sweden 2002-2008. Raw refers to the wage equation in equation (3) without any further controls. The Mincerian model adds years of schooling, experience and its squared value as well as dummies for immigrants, males and education specialization. The full specification further adds variables reflecting labor market status and employer characteristics of each worker. OLS refers to the pooled OLS estimator and FE to a panel estimator with worker fixed effects. All variables are defined in Table 1. The full FE model excludes immigrant and sex dummies as these reflect time-invariant worker characteristics. All models include year and region dummies as well as region-year dummies, where the latter account for any region-specific time-varying shocks shared by all workers in the same local labor market region. The dependent variable is the natural logarithm of wage earnings. Robust standard errors are presented in brackets. *** p<0.01, ** p<0.05, * p<0.1.

As is evident from the table, the inclusion of worker fixed effects induces the wage-density elasticities to drop sharply. Both the municipal and the regional densities drop from about .03 to .008. The raw OLS estimates are thus almost four times as a high as the estimates obtained with worker fixed effects. The estimates show that after accounting for worker fixed effects, a doubling of either municipal regional density is associated with about .8 percent higher wages. This result suggests that spatial sorting on unobservable worker characteristics is indeed an important source of the wage-density relationship. After controlling for sorting there remains a small but significantly positive wage-density elasticity, indicating the existence of a small agglomeration effect. Sorting effects yet dominate, and these patterns are broadly in line with the findings by Combes et al (2008) on worker-level data for France.

Turning to the control variables we find that the estimated influence of years of schooling is positive throughout, and the magnitude of the estimates are roughly in accordance with results reported in previous studies (cf. OECD 1998). Moreover, more experienced workers earn in general better though the positive effect falls off as experience rises. Immigrants earn less on average whereas male workers earn more than females. In the OLS specifications, tenure and the number of prior employees are positive, though the latter estimate is statistically insignificant. That longer tenure is positive is in line with the hypothesis that tenure signals match quality (cf Farber 1994). In the fixed effects specification, however, tenure is negative and significant. This may be explained in two ways. First, those with long tenure represent a select group which may have lower career aspirations. Second, the fixed effects model may capture those that switch employer and make a career move after a number of years of accumulation of experience with the same employer.[22] Employer size is positive and significant throughout which is an established result in the literature (Oi and Idson 1999).

The main aim of this paper is yet to test whether the magnitude and sources of the wage density premium vary across workers with different sets of skills. We split the population of workers in two groups: one with jobs with high fraction non-routine tasks and one with jobs with low fraction non-routine tasks. Table 7 reports results obtained for the first group – workers with jobs associated with high fraction non-routine tasks.

The results are similar as those reported in Table 6. The raw OLS estimates are around .03 for municipal and regional density, while the extra-regional density is negative (though not significant). Controlling for observable worker and employer characteristics reduces the estimates for municipal and regional density to about .02. Also for workers with non-routine jobs, spatial sorting effects dominate. Including worker fixed effects in the raw model reduces the estimated wage-density elasticties substantially. For municipal and regional density the difference between the raw OLS and the raw model with worker fixed effects amounts to a factor of almost four. There is a general tendency that the estimates with worker fixed effects are larger for workers with jobs in which non-routine tasks are important, but the differences to Table 6 are still marginal.

[22]Such effects are more likely to be captured when the estimates are based on within variance, as workers are here followed over years.

Table 7. The relationship between spatial economic density and wages for workers with occupations associated with high fractions of non-routine job tasks.

	Raw OLS	**Mincerian OLS**	**Full OLS**	**Raw with worker FE**	**Full with worker FE**
Municipal density (log)	0.0317***	0.0253***	0.0250***	0.00810***	0.00655***
	(0.00206)	(0.00260)	(0.00141)	(0.000345)	(0.000346)
Regional density (log)	0.0364***	0.0263***	0.0240***	0.00868***	0.00618***
	(0.00734)	(0.00523)	(0.00829)	(0.000772)	(0.000769)
Extra-regional density (log)	-0.0271	-0.0231	-0.0253**	-0.0124***	-0.00834***
	(0.0232)	(0.0183)	(0.0114)	(0.00105)	(0.00105)
Years of schooling		0.0797***	0.0766***		0.133*
		(0.00312)	(0.00289)		(0.0691)
Experience		0.0556***	0.0513***		0.0926
		(0.00204)	(0.00185)		(0.0691)
Experience^2		-0.000862***	-0.000794***		-0.000737***
		(3.95e-05)	(3.54e-05)		(4.05e-06)
Immigrant (dummy)		-0.0419***	-0.0371***		
		(0.00285)	(0.00260)		
Male (dummy)		0.353***	0.351***		
		(0.00319)	(0.00306)		
Tenure			0.0105***		-0.00644***
			(0.000400)		(0.000172)
Number of prior employees			0.000754		-0.00741***
			(0.000741)		(0.000237)
New occupation (dummy)			-0.0703***		-0.0151***
			(0.00209)		(0.000502)
Employer size (log)			0.0259***		0.0149***
			(0.00314)		(0.000236)
Year dummies	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes
Region*Year effects	Yes	Yes	Yes	Yes	Yes
Education type dummies	No	Yes	Yes	No	Yes
Industry dummies	No	No	Yes	No	Yes
Observations	5,986,454	5,986,454	5,986,454	5,986,454	5,986,454
Individuals	0.038	0.258	0.280	0.061	0.074
R-squared	1,388,166	1,388,166	1,388,166	1,388,166	1,388,166

Note: The table reports estimates of wage-density elasticities for private sector workers in Sweden 2002-2008 with occupations associated with high fractions of non-routine job tasks (see Table 2). Raw refers to the wage equation in equation (3) without any further controls. The Mincerian model adds years of schooling, experience and its squared value as well as dummies for immigrants, males and education specialization. The full specification further adds variables reflecting labor market status and employer characteristics of each worker. OLS refers to the pooled OLS estimator and FE to a panel estimator with worker fixed effects. All variables are defined in Table 1. The full FE model excludes immigrant and sex dummies as these reflect time-invariant worker characteristics. All models include year and region dummies as well as region-year dummies, where the latter account for any region-specific time-varying shocks shared by all workers in the same local labor market region. The dependent variable is the natural logarithm of wage earnings. Robust standard errors are presented in brackets. *** p<0.01, ** p<0.05, * p<0.1

Results for workers with jobs associated with low fractions of non-routine tasks are presented in Table 8. The first apparent result is that the raw OLS estimates suggest a non-existent or negative wage-density relationship for jobs with low fraction non-routine tasks. The estimated coefficient for municipal density is negative and significant, whereas the remaining densities are positive but insignificant. There are thus no clear patterns that workers with these jobs earn more in denser areas.

Instead, it appears that workers with routine jobs in denser municipalities earn less than their more rural counterparts.

Table 8. The relationship between spatial economic density and wages for workers with occupations associated with low fractions of non-routine job tasks.

	Raw OLS	**Mincerian OLS**	**Full OLS**	**Raw with worker FE**	**Full with worker FE**
Municipal density (log)	-0.0124*** (0.00339)	-8.97e-05 (0.00217)	0.00714*** (0.000639)	0.00311*** (0.000363)	0.00215*** (0.000363)
Regional density (log)	0.00917 (0.00651)	0.00483 (0.00445)	0.00623*** (0.00224)	0.00358*** (0.000741)	0.00176** (0.000736)
Extra-regional density (log)	0.00673 (0.0111)	0.00244 (0.00944)	-0.00807** (0.00365)	-0.00555*** (0.000950)	-0.00318*** (0.000944)
Years of schooling		0.0212*** (0.00324)	0.0208*** (0.00252)		0.0627*** (0.0206)
Experience		0.0351*** (0.00152)	0.0252*** (0.00145)		0.0479** (0.0206)
Experience^2		-0.000583*** (2.02e-05)	-0.000431*** (2.10e-05)		-0.000614*** (4.16e-06)
Immigrant (dummy)		-0.128*** (0.00870)	-0.0968*** (0.00505)		
Male (dummy)		0.366*** (0.0146)	0.319*** (0.00914)		
Tenure			0.0247*** (0.000979)		-0.0124*** (0.000195)
Number of prior employees			-0.0220*** (0.00320)		-0.0265*** (0.000262)
New occupation (dummy)			-0.114*** (0.00302)		-0.0521*** (0.000558)
Employer size (log)			0.0202*** (0.00146)		0.0208*** (0.000282)
Year dummies	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes
Region*Year effects	Yes	Yes	Yes	Yes	Yes
Education type dummies	No	Yes	Yes	No	Yes
Industry dummies	No	No	Yes	No	Yes
Observations	6,565,670	6,565,670	6,565,670	6,565,670	6,565,670
Individuals	0.018	0.177	0.223	0.043	0.061
R-squared	1,659,886	1,659,886	1,659,886	1,659,886	1,659,886

Note: The table reports estimates of wage-density elasticities for private sector workers in Sweden 2002-2008 with occupations associated with low fractions of non-routine job tasks (see Table 2). Raw refers to the wage equation in equation (3) without any further controls. The Mincerian model adds years of schooling, experience and its squared value as well as dummies for immigrants, males and education specialization. The full specification further adds variables reflecting labor market status and employer characteristics of each worker. OLS refers to the pooled OLS estimator and FE to a panel estimator with worker fixed effects. All variables are defined in Table 1. The full FE model excludes immigrant and sex dummies as these reflect time-invariant worker characteristics. All models include year and region dummies as well as region-year dummies, where the latter account for any region-specific time-varying shocks shared by all workers in the same local labor market region. The dependent variable is the natural logarithm of wage earnings. Robust standard errors are presented in brackets. *** p<0.01, ** p<0.05, * p<0.1

In the Mincerian specification which adds education, experience and other basic worker characteristics, none of the density variables is statistically significant. Only in the full OLS model, which further adds employer characteristics and indicators of the labor market status of the workers, the wage-density elasticities become statistically significant and have the same sign as in previous tables. They are yet substantially smaller than in Tables 6 and 7. The estimated parameters for municipal and regional density are about .007 and .006, respectively, and the same estimates for jobs with high-fraction non-routine tasks are in the order of .02, i.e. a difference of a factor of more than three. This pattern does not change from the inclusion of worker fixed effects. The last two columns in the table show that adding worker fixed effects reduces the estimated elasticities for municipal and regional density to about .002. Again, the same elasticities for non-routine jobs (Table 7) are about three times as large.

We draw three main conclusions from these patterns. First, the wage-density relationship is much weaker for jobs with low fractions of non-routine tasks. The wage premium from operating in denser regions is much smaller for these jobs – a pattern that is robust across specifications. Second, the effects from spatial sorting on basic observable worker characteristics go in the opposite direction for these jobs as compared with non-routine jobs. Conditioning on Mincerian variables as well as indicators of workers' labor market status increases rather the decreases the estimated wage-density elasticities. Yet, spatial sorting on permanent worker heterogeneity go in the same direction as for jobs with high fraction of non-routine tasks. Third, the wage-density relationship attributable to agglomeration economies is significantly smaller for jobs with low fraction of non-routine tasks. The remainder wage-density elasticities after controlling for observable and unobservable worker heterogeneity is about three times larger for jobs with high fraction non-routine tasks, .006 compared to .002. In economic terms these effects are nevertheless small, as they imply that a doubling of density yields about .6 and .2 percent higher wages, respectively.

Our estimates thus imply that operating in dense regions primarily generates benefits for workers with skills associated with non-routine job tasks, and that these types of jobs are also more likely to be found in denser regions in the first place (see Table 3). The wage-density elasticities found for the full sample of Swedish workers (Table 6) are primarily driven by workers with skills associated with non-routine jobs. These results correspond to the analyses by Bacolod et al (2009) who find that an urban wage premium predominantly applies to jobs in which cognitive and people skills are important.

The results in Tables 6-8 do show evidence of agglomeration economies in the sense that significant wage-density elasticities remain after controlling for worker characteristics, observable as well as unobservable. But the mere existence of remainder wage-density elasticities reported in Tables 6-7 does not inform about the type of agglomeration economy. As explained in the previous sections, to further probe the results, and to cautiously get at learning effects, we also estimate wage premiums for workers that move away from a dense metropolitan area. If density fosters human capital accumulation, this means that benefits remain with the worker upon moving away from dense agglomerated regions. Such learning should yet primarily pertain to workers with non-routine skills who are more apt to learn from the environment.

Table 9 presents the estimated coefficient of a dummy variable which identifies workers that move from any of Sweden's three main metropolitan labor market regions (Stockholm, Göteborg and Malmö) to anywhere else in the country. These estimations include the full set of variables as the 'full worker fixed effects' estimations in Tables 6-8. Due to the inclusion of worker fixed effects, the coefficient estimate shows whether the wage of a worker remains unaffected, increase or decrease upon moving away from a metropolitan region. An insignificant or positive parameter estimate lends support for human capital, as it means that the worker at least retains his or her wage upon moving away from a larger agglomeration. We present results obtained for all workers as well as for workers with jobs associated with high and low fraction non-routine tasks, respectively.

Table 9. Wage premium for workers moving away from a metropolitan region to the rest of the country, by fraction of non-routine job tasks.

	All private sector workers	**High fraction non-routine tasks**	**Low fraction non-routine tasks**
Dummy for moving away from metropolitan region	0.00286 *** (0.0014)	0.01085*** (0.0020)	-0.00055 (0.0021)
Model	*Full with worker fixed effects*	*Full with worker fixed effects*	*Full with worker fixed effects*

Note: The table reports the coefficient estimate of a dummy variable reflecting a move from any of Sweden's three metropolitan labor market regions (Stockholm, Göteborg and Malmö) to anywhere else in Sweden. The underlying model is a panel estimator with worker fix effects including the full set of additional control variables reported in the 'Full with worker FE' specification in Tables 6-7. Complete estimation results are obtained from the authors upon request. The dependent variable is the natural logarithm of wage earnings. Robust standard errors are presented in brackets. *** p<0.01, ** p<0.05, * p<0.1

We find a small significant premium for those workers that move away from a metropolitan region among the full sample of Swedish private sector workers. For workers with jobs associated with high fraction non-routine tasks the premium is positive, statistically significant and substantially larger compared to the full sample of workers. For workers with jobs associated with low fraction of non-routine tasks, however, the estimated coefficient is negative albeit insignificant. These patterns are consistent with density fostering human capital accumulation that remains with the workers upon moving away from the agglomerations. Moreover, these effects appear to be particularly strong for workers with jobs with non-routine tasks. These are jobs requiring problem-solving and more interaction with others implying not only that learning is more important, but also greater opportunities for learning.

One way to appreciate the positive parameter estimate of the dummy for moving away is that workers' may value the greater variety of consumption-based amenities and the greater thickness of the local labor markets in the metropolitan regions, and thus want to be compensated when moving away from these regions (cf. Roback 1982). Yet, such compensation can only be motivated if the workers bring human capital that is valued by the employers. From this perspective, one may argue that it is only workers with jobs associated with high fraction non-routine tasks that 'learn enough' in the city to motivate such compensation. A caveat should, however, be noted: among other possible sources of endogeneity, the workers that move may be a self-selected minority. For instance, Gould (2007) emphasizes that the decision to move may be endogenous as the change in the wage associated with moving may be correlated with changes in the quality of the opportunities in the different locations. As we do not fully account for such potential endogeneity here, the results in Table 9 should be interpreted as somewhat restrained empirical support for the learning proposition.

5. CONCLUSIONS

The main conclusion from this paper is that the benefits from agglomeration are not uniform across workers and activities. Agglomeration yields productivity gains primarily in contexts in which problem solving and interaction with others are important. This conclusion is derived from an analysis of how the magnitude and sources of the wage density premium differs across workers with different degrees of skills associated with non-routine job tasks. Non-routine job tasks typically involve problem solving, lack of deductive and codifiable information, as well as interaction with others.

The analyses in the paper demonstrate that the relationship between wages and spatial economic density is significantly stronger for workers with skills pertaining to non-routine tasks, and such workers are more concentrated to denser regions in the first place. A main finding is that agglomeration economies, i.e. productivity gains from interactions between workers and their local environment, are quantitatively a more important source of the density wage premium for non-routine workers. The bottom line is that skills associated with non-routine job tasks are better rewarded in denser regions. For workers with routine job tasks, however, agglomeration economies are virtually non-existent.

In a broad sense, these results reinforce the idea of large city regions as 'innovation environments', fostering and rewarding activities related to face-to-face interaction, knowledge, ideas and development of new products, designs, organizational routines and technology blueprints. Innovation is indeed a prominent example of a context in which skills associated with non-routine tasks are imperative. The literature on 'geography of innovation' has for a long time made the argument that cities matter more for innovation, but the kind of micro-based evidence presented in this paper, where the question of which sets of skills and job tasks are better rewarded in agglomerations is directly addressed, provides an improved understanding of these issues. After all, innovation processes are essentially linked to workers' skill sets and the nature of their jobs and tasks.

As regards the source of spatial wage disparities in general, our analyses line-up with the growing evidence suggesting that who you are is more important than where you live in explaining spatial wage disparities. The main reason why workers in denser regions earn more is simply that they are different from the workers in more rural regions. Spatial sorting on permanent unobserved worker heterogeneity is the main source of the density wage premium.

Further work on these issues may take a variety of directions. One is to untangle the various sources of agglomeration economies. For instance, to what extent are the stronger agglomeration economies of non-routine workers driven by matching, learning and sharing mechanisms, respectively? Another route is to focus on the location processes of workers with different skills and abilities. This applies to both theoretical and empirical work. Since a large part of the density wage premium is due to spatial sorting (even for non-routine workers), the question of why denser areas are more attractive places for workers with different skill sets and their migration patterns appear as a particularly relevant line of inquiry.

REFERENCES

Andersson M and U Gråsjö (2009), "Spatial dependence and the representation of space in empirical models", *Annals of Regional Science* 43: 159-180.

Andersson M and P Thulin (2013) "Does spatial employment density spur inter-firm job switching?", *Annals of Regional Science*, 51(1), 245-272

Autor, D.H, F. Levy and R.J Murnane (2003), "The skill content of recent technological change – an empirical exploration", *Quarterly Journal of Economics*, 118 (4), 1279-1333

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

Baum-Snow, N and R Pavan (2012), "Understanding the city size wage gap", *Review of Economic Studies* 79: 88-127

Becker S O, K. Ekholm and M-A Muendler (2009), "Offshoring on the onshore composition of tasks and skills", CEPR Discussion paper 7391

Ciccone A and R E Hall (1996), "Productivity and the density of economic activity", *American Economic Review* 86: 54-70

Combes P, G Duranton and L Gobillon (2008), "Spatial wage disparities: sorting matters!", *Journal of Urban Economics* 63: 723-742

Combes P, G Duranton and L Gobillon (2011), "The identification of agglomeration economies", *Journal of Economic Geography*, 11: 253-266.

De La Roca J and D Puga (2012), "Learning by working in big cities", Mimeo

Farber H.S (1994), "The analysis of inter-firm worker mobility", *Journal of Labor Economics* 12:554–593

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

Glaeser E.L (1999),"Learning in cities", *Journal of Urban Economics* 46: 254-277

Glaeser E.L (2008), *Cities, Agglomeration and Spatial Equilibrium*. Oxford University Press, Oxford

Gould E.D (2007), "Cities, workers and wages: a structural analysis of the urban wage premium", *Review of Economic Studies* 74: 477-506

Hakkala, K. F Heyman and F Sjöholm (2008), "Multinational firms and job tasks", IFN working paper No. 781

Harris C.D (1954), "The market as a factor in the localization of industry in the U.S", *Annals of the Association of American Geographers* 44: 315-348

Hesley, R and W Strange (1990), "Matching and agglomeration economies in a system of cities," *Regional Science and Urban Economics*. 20: 189–212

Irwin E.G, A.M Isserman, M Kilkenny and M Partridge (2010), "A century of research on rural development and regional issues", *American Journal of Agricultural Economics* 92: 522-553

Johansson B, J Klaesson and M Olsson (2003), "Commuters' non-linear response to time distances", *Journal of Geographical Systems* 5: 315-329

Kim S (1990), “Labor heterogeneity, wage bargaining and agglomeration economies”, *Journal of Urban Economics* 28: 160–177

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

Melo P.C, D.J Graham and R.B Noland (2009), “A meta-analysis of estimates of urban agglomeration economies”, *Regional Science and Urban Economics* 39: 332-342

Mincer J (1974), *Schooling, Experience and Earnings*. National Bureau of Economics Research (NBER), New York

Moretti E. (2004), “Workers' education, spillovers and productivity: evidence from plant-level production functions”, *American Economic Review* 94: 656-690.

Moretti, E (2008), “Real wage inequality”, CEPR Discussion paper 6997

Möller J and A Haas (2003), “The agglomeration differential reconsidered: an investigation with German micro data 1984-1997”, in Bröcker J, D Dohse and R Soltwedel (Eds.), *Innovation Clusters and Interregional Competition*, Springer. Berlin

OECD (19989, *Human capital investment – an international comparison*, OECD, Paris

Oi W.Y and T.L Idson (1999), “Firm size and wages”, in *Handbook of Labor Economics, vol. 3*. North-Holland, Amsterdam

Puga, D (2010),”The magnitude and causes of agglomeration economies”, *Journal of Regional Science* 50: 203-219

Rauch J. E (1993), “Productivity gains from geographic concentration of human capital: evidence from the cities”, *Journal of Urban Economics*, 34: 380-400

Roback J. (1982), “Wages, rents and the quality of life, *Journal of Political Economy* 90: 1257-1278

Splitz-Oener, A (2006), “Technical change, job tasks and rising educational demands – looking outside the wage structure”, *Journal of Labor Economics*, 24 (2), 235-270

Strange W.C (2009), “Viewpoint: agglomeration research in the age of disaggregation”, *Canadian Journal of Economics / Revue Canadienne d'Economique* 42: 1-27

Tobler W.R (1970), “A computer movie simulating urban growth in the Detroit region”, *Economic Geography* 46: 234-240

Wheeler C.H (2006), “Search, sorting, and urban agglomeration”, *Journal of Labor Economics* 19: 879-899

Yankow J.J (2006), “Why do cities pay more? - an empirical examination of some competing theories of the urban wage premium”, *Journal of Urban Economics* 60: 139-161

APPENDIX

VARIABLE	MEAN	MEDIAN	ST.DEV
Yearly wage (log)	7.80	7.88	.660
Municipal density (log)	22.96	22.87	1.668
Regional density (log)	21.79	22.51	4.734
Extra-regional density (log)	21.18	21.18	1.093
Years of schooling	12.05	12	2.079
Experience	21.96	21	11.88
Immigrant (dummy)	.13	0	.3343
Male (dummy)	.6504	1	.4768
Tenure	3.25	3	2.022
Number of prior employees	.44	0	.933
New occupation (dummy)	.17	0	.3753
Employer size (log)	4.06	3.91	1.868

N=12 367 700