Cities and tasks

Koster, Hans R.A.; Ozgen, Ceren

published in
Journal of Urban Economics
2021

DOI (link to publisher)
10.1016/j.jue.2021.103386

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

Journal of Urban Economics 126 (2021) 103386

Contents lists available at ScienceDirect

Journal of Urban Economics

journal homepage: www.elsevier.com/locate/jue

Cities and tasks[☆]

Hans R.A. Koster[a], Ceren Ozgen[b,c,*]

[a] Department of Spatial Economics, Vrije Universiteit Amsterdam, Netherlands
[b] Department of Economics, University of Birmingham, University House 1105, Birmingham B15 2TT, UK
[c] IZA; GLO, UK

ARTICLE INFO

JEL classification:
R30
R33

Keywords:
Routinisation
Tasks
Agglomeration economies
Employment density
Skills mismatch

ABSTRACT

This paper explores the relationship between routine-biased technological change and agglomeration economies. Using administrative data from the Netherlands, we first show that in dense areas, jobs are less routine-task intensive (*i.e.* less repetitive and automatable), meaning that jobs cover a larger spectrum of *tasks*. We then explore how the routine intensity of jobs affects the urban wage premium. We find that the urban wage premium is higher for workers performing non-routine tasks, particularly analytic tasks, while it is absent for workers in routine task intensive jobs. These findings also hold *within* skill groups and suggest that routinisation increases spatial wage inequality within urban areas. We further provide suggestive evidence that a better matching of skills to jobs and increased learning opportunities in cities can explain these findings.

1. Introduction

Mainly through automation, the asymmetric impact of technological developments in labour markets – known as Routine-Biased Technological Change (RBTC) (Autor et al., 2003) – has led to an increase in employment in analytic and manual task intensive occupations, and to a decrease in those occupations requiring routine tasks. Workers performing analytic tasks, in addition to enjoying a comparative advantage in adapting to new technologies, benefit from higher returns to analytic tasks (Acemoglu and Restrepo, 2018). Following Acemoglu and Autor (2011), in this paper, tasks are defined as a unit of work activity required to produce certain outputs, while skills refer to a formal degree of education obtained by a worker.[1]

The complementarity between analytic tasks and technological change is likely to further disadvantage workers undertaking routine tasks once the location of economic activity is taken into account. This occurs because denser areas, which provide a large supply of specialised, high-skilled workers, offer relatively fewer jobs that are routine task intensive (see Autor, 2019; Davis et al., 2020, on the geography of polarisation in the U.S. and in France, respectively). Michaels et al. (2016), for example, show that non-routine occupations are much more likely to be performed in metro areas in 2000 in the U.S (which is confirmed by Ehrl and Monteiro Monasterio 2016 for Brazil and by Grujovic 2018 for Germany). Similar to Michaels et al. (2016) and others, in this study, we confirm a clear negative relationship between employment density and routine task intensity of occupations in the Netherlands. In our analysis, the least routine task intensive occupations, *e.g.* professional services managers and teaching professionals, are exclusively those requiring *analytic tasks* – encompassing analysing data, thinking creatively, interpretation of information for others and requiring complex personal interactions. By contrast, the most routine task intensive occupations are those that require repetitive and routine tasks such as vehicle and laundry cleaning workers or textile machine operators.

There are three main reasons why occupations and firms that require analytic tasks, in particular, are concentrated in dense urban areas. First, workers that are able to perform analytic tasks may sort themselves into

☆ Ceren Ozgen gratefully acknowledges the Marie-Skłodowska Curie Individual Grant for MAStErS project (H2020-MSCA-IF-2015, No. 705366) from the European Commission. Hans Koster acknowledges the support of the HSE University Basic Research Program. A part of this paper was written while Koster was an IAS Vanguard Fellow at the University of Birmingham. We thank Donald Davis and two anonymous referees for providing insightful comments. We thank Matt Cole, Rob Elliott, and Jos van Ommeren, as well as participants of Workshop on Dynamics of Skills Supply and Demand in Maastricht (2018) for discussions and useful suggestions.

* Corresponding author.
E-mail addresses: h.koster@vu.nl (H.R.A. Koster), c.ozgen@bham.ac.uk (C. Ozgen).

[1] The previous literature sometimes confusingly refers to tasks as 'skills'. To avoid such confusion in the paper, we make a distinction between tasks and skills proxied by (formal) education.

https://doi.org/10.1016/j.jue.2021.103386
Received 26 June 2020; Received in revised form 7 August 2021
Available online 14 August 2021

cities. Combes et al. (2008) show that a large part of the variation in spatial wages can be explained by worker characteristics. Davis and Dingel (2019, 2020) show that larger cities are skill-abundant with greater specialisation in skill-intensive activities. Davis et al. (2020) further show that larger cities have a higher share of highly paid jobs, which at the same time are often more analytic. Second, more productive entrepreneurs and firms may self-select into cities. Behrens et al. (2014) argue that entrepreneurial profit increases with city size; hence, more productive less routine task intensive firms sort themselves into urban areas. Moreover Combes et al. (2012) show that in cities competition is tougher, allowing only the most productive firms to survive. Third, agglomeration economies may foster new idea generation and complementarity among resources, implying that occupations in cities are more task-diverse and less routine task intensive (Bacolod et al., 2009; Lin, 2011; Davis and Dingel, 2019). This relates to the literature which shows that particularly diverse cities are engines of innovation and entrepreneurial activities (Glaeser et al., 1992; Henderson et al., 1995; Duranton and Puga, 2001; Davis and Dingel, 2020).

In line with the descriptive evidence that urban areas attract more non-routine, particularly analytical workers, we aim to investigate the *implications of routinisation for agglomeration economies*.[2] We hypothesise that cities should offer higher returns to analytic task intensive jobs through higher wages and better employment opportunities. Specifically, we argue that the urban wage premia should be heterogeneous across different levels of the routine task intensity index. This is because learning is faster and matching is better among non-routine task intensive jobs. Although our analysis is static in nature, it benefits from the works of Violante (2002) and Beaudry et al. (2016) to explain the firms' and workers' potential pace of adaptation to technological change. They argue that with the technological advances taking place, workers at varying levels of task complexity are expected to adapt to the technological improvements at a different pace. The reason for the level of technology adaptation to be variable across jobs is due to the fact that not every firm is at the same level of technology, even though they may be producing similar goods and services. These productivity differentials should then lead to a variation in wage returns as workers undertaking analytic task intensive jobs are more likely to adapt to the technological advances quicker.

Our main contributions are as follows. First and foremost, we combine the literature on agglomeration economies with the literature on routine-biased technological change. Accordingly, we construct a routine task intensity index by using rich linked employer-employee data (LEED) from the Netherlands in 2006–2012. Following Autor and Dorn (2013), Goos et al. (2014), but adapting the O*NET based task measure from SOC to ISCO classification in the Netherlands, we work with a very refined routine task intensity index at the *4-digit* ISCO level.[3] Using semi-parametric estimation techniques, we then let agglomeration economies depend on the routine task intensity of jobs.

Second, by using the aggregate distribution of commuting times in the Netherlands, we calculate the number of jobs within the relevant commuting time from the home location of the workers and hence construct a commuting time-weighted employment density measure. The standard measures of density have the disadvantage of being commonly defined on the basis of some legal demarcation rather than labour market dynamics and are subject to the modifiable-areal unit problem (see Briant et al., 2010).

Third, we improve on the identification of agglomeration effects by identifying the effects within labour market areas and controlling for historic sorting of high-skilled workers more than century ago into certain areas within the labour markets. Furthermore, using fine-grained data on the skill composition of jobs in 1909, we determine the share of high-skilled, medium-skilled and low-skilled workers within commuting distance in 1909 and include those as extra controls in the regressions. To investigate whether any remaining ability bias is an issue, we utilise a unique parent-child linkage dataset and hence focus on siblings in our sample to mitigate the issue of unobserved innate ability of workers. This enables us to identify agglomeration effects on the basis of siblings by employment density.

We find that controlling for sorting and mitigating endogeneity issues, there is a sizeable negative effect of employment density on the routine task intensity of jobs, meaning agglomeration economies reduce the routine task intensity of jobs, *e.g.* through creating complex tasks while also expanding new combinations of existing tasks, hence leading to a strong spatial concentration of complex tasks in denser areas. This aligns with the literature arguing that cities are places where new ideas are created and innovations take place (see *e.g.* Davis and Dingel, 2020; Duranton and Puga, 2001, 2004; Lin, 2011). We show that, depending on the complexity of the tasks performed, there is a large variation in the urban wage premium. For example, the wage-density elasticity is approximately 0.15 for workers that are in non-routine jobs, while it is essentially zero for workers in routine task intensive jobs. Moreover, observationally similar workers performing analytic tasks receive a much higher urban wage premium, regardless of their education level. This suggests that only using skills to explain differences in urban premia will be inadequate.

To explain these findings, we provide suggestive evidence on the potential channels, such as better skills matching (*i.e.* skills to jobs) and the enhanced learning opportunities cities offer (*i.e.* through work experience in the (local) labour market). We show that better matching is only relevant for workers performing analytic tasks. Moreover, learning externalities accrue only to workers in non-routine task intensive jobs with more years of experience in the local labour market. We interpret the latter as circumstantial evidence that learning effects are important in explaining why mostly workers in non-routine task intensive occupations receive density premia.

Related literature. Our paper first ties into a vast literature on agglomeration economies, that assumes higher productivity and wages in dense places (Ciccone and Hall, 1996; Ciccone, 2002; Combes et al., 2008; Melo et al., 2009). Traditionally, the density premium is assumed to percolate across all skill groups proportional to city size, hence independent of workers' skill or tasks they perform.[4] This proposition suggests that the urban wage premium will homogeneously increase in talent with city size and predicts a spatially *invariant* skill premium (Behrens and Robert-Nicoud, 2015; Davis and Dingel, 2019). This traditional framework commonly uses the observed degree of education as an approximation of unobserved skills (*i.e.* a degree premium), rather than a premium that directly relates to the market value of the skills the worker possesses. However, widening wage and skill inequalities within occupational groups – even between observationally similar workers *within* urban areas – warrant further scrutiny to understand the heterogeneity in the urban wage premium, as well as to revisit whether the approximation of workers' talent with skills sufficiently explains this heterogeneity.[5]

Davis and Dingel (2019) rightfully argue that the canonical spatial-equilibrium model falls short as talent-homogeneous cities cannot rationalise the spatial variation in skill premia. Similarly

[2] Note that in this paper we focus on the economies of density (Ahlfeldt et al., 2016), rather than localisation economies or the economies of diversity (Glaeser et al., 1992; Duranton and Puga, 2001).

[3] O*NET stands for The Occupational Information Network developed with the support of the U.S.Department of Labor/Employment and Training Administration (USDOL/ETA); SOC stands for Standard Occupational Classification system used in the U.S.

[4] The density premium is commonly explained by the availability of better technology in cities complementing higher human capital stock, and being able to learn from other high ability workers in the vicinity (Glaeser and Maré, 2001; Wheeler, 2006; D'Costa and Overman, 2014; De la Roca and Puga, 2017). However, as Duranton and Kerr (2018) highlight, until recently the lack of data at the firm-worker level hindered an exploration of wage returns by type of workers.

[5] Autor (2019) defines occupation groups as nine exhaustive mutually exclusive occupational categories rank-ordered by the level of mean log wage.

Black et al. (2009) show that only with homothetic preferences, skill premia can be location-invariant, while with more realistic non-homothetic preferences, skill premia are lower in cities with higher house price. Deviating from the canonical spatial-equilibrium model Davis and Dingel (2019) instead distinguish between tradables and non-tradables producing workers. Tradables producers gain from learning externalities through frequent local interactions (*i.e.* idea exchanges). The resulting spatial sorting of tradables producers and their productivity-increasing idea exchanges in cities provide a strong basis for the polarisation of economic activity across space and higher skill premia observed in denser areas. Note that their model differentiates itself from the earlier spatial equilibrium models by not defining skills by a level of ability linked to an educational degree. This theoretical framework implicitly supports studying *tasks* rather than skills to analyse heterogeneous productivity gains and is therefore consistent with our findings that the urban wage premium only exists for non-routine task workers.[6]

Our paper is further related to the works of Bacolod et al. (2009) and Grujovic (2018) that have explicitly tested whether agglomeration economies vary with the task content of jobs. Bacolod et al. (2009) find that large cities host more complex jobs than small cities, but only to a modest degree. Unfortunately, their definition of skills required to perform certain tasks is, due to data restrictions, somewhat aggregate. Grujovic (2018) estimates heterogeneous urban wage premium for task content of jobs and ranks the wage returns to tasks on the basis of their complexity.[7]

Another strand of literature that aligns strongly with our findings above is the literature on RBTC, which tries to explain the changing wage structure between skill groups. It focuses on two major sources: *(i)* automation that led to a hollowing out of the employment in middle-skilled occupations through routine-biased technical change (see Adermon and Gustavsson, 2015; Autor et al., 2015; Autor and Acemoglu, 2011; Goos et al., 2014; Oesch, 2013); and *(ii)* skill-biased technological change (SBTC) that assumes an advantage for high-skilled workers through a strong complementarity between skills and technology (see Acemoglu and Autor, 2011). The so-called canonical model, which does not make an explicit distinction between skills and tasks, implies that the improvements in technology should naturally increase the demand for skilled workers. In recent decades, though, there has been a strong polarisation of jobs through the growth in the share of employment in high and low skilled occupations leading to an increased wage dispersion between and within skill groups. Autor (2019) further refines the implications of job polarisation on the reshaping of work in urban areas: *(i)* cities have always been more intensive in high-skilled work; *(ii)* although the share of low-skilled work is typically lower in cities, in the last two decades its employment share increased considerably in the densest areas (*e.g.* in the U.S. and many other developed countries); *(iii)* since the early 80s there has been a sharp attenuation in the fraction of middle-skilled work in the densest areas. Looking at these patterns in detail, the compositional shift within education groups indicates a profound reallocation of medium-skilled workers from middle to low-skilled work such as services, transportation and labourer occupations, but only in the *densest areas*.

Recently, Van der Velde (2017) shows that occupations where tasks performed complement newer technologies exhibit higher wage dispersion. When focused only on the skill levels of workers as opposed to the tasks they perform, the relative differences between wages across occupations remain largely unexplained by the observables, despite controlling for education. This is at odds with the empirical evidence since the technology has impacted the skill groups variably across time and space, leading to job polarisation.[8]

These changes in the structure of work have *two* implications for our paper. Firstly, as Autor (2019) suggests, the link between polarisation, the changing structure of work and wages across geographic locations is to be explored within a framework that allows for the uneven folding of occupational structure in the labour market across locations. Secondly, the downward pressure exerted due to polarisation between occupational skill groups, particularly from the middle to the bottom of the occupational distribution likely gives rise to skills mismatch (Beaudry et al., 2016).[9]

Indeed the empirical studies (Autor and Acemoglu, 2011; Autor et al., 2015) that adopt a task approach to analyse the work content of jobs inevitably allows larger variation in worker productivity, hence leading to a larger wage dispersion, such that even for observationally identical workers wages can vary significantly due to the tasks they perform. We confirm this in our paper as predominantly workers in non-routine task-intensive jobs profit from urban density, regardless of their skill level (*i.e.* their formal education).

Given the implications of the literature on RBTC, we try to understand the mechanisms why agglomeration externalities may be task-specific. Duranton and Puga (2004) distinguish between three mechanisms why agglomeration economies arise: learning, matching and sharing.[10] We, therefore, expect learning opportunities to be greater in cities (in line with Davis and Dingel, 2019, and RBTC), even more so for non-routine workers due to the diffusion of technologies creating demand for non-routine tasks. However, De la Roca and Puga (2017) show that learning externalities may not be instantaneous but may increase with experience (*i.e.* proxied by the time spent in large cities). We will show in Section 4 that agglomeration economies are absent for inexperienced workers even if they are performing analytic intensive tasks. We interpret this as suggestive evidence that time spent in the labour market or in cities is important to exploit learning effects.

Furthermore, as the literature suggests, a better matching of jobs-skills is another externality workers can benefit from in urban areas (Boualam, 2014; Berlingieri, 2019). Since non-routine task-intensive jobs and workers are more clustered in denser areas, we expect cities to provide better matching opportunities, particularly for non-routine workers. To investigate this, in line with Beaudry et al. (2016), we consider two dimensions of skills mismatch, namely vertical mismatch (*i.e.* overqualification) and horizontal mismatch.[11]

[6] While several papers analyse the relationship between agglomeration and skills (see *e.g.* Glaeser and Maré, 2001; Wheeler, 2006; D'Costa and Overman, 2014; De la Roca and Puga, 2017), these studies are not able to go beyond using college degree as an aggregate proxy for worker skills; therefore they cannot explain within (skill) group variation in returns to human capital endowments. We aim to improve on this literature by measuring the routine task intensity of occupations through a continuous measure of task complexity.

[7] Our analytical framework deviates from this work because our analysis and theoretical discussion focus on the importance of tasks, rather than skills (we control for skills proxied by education throughout the paper). We show that tasks are not a sub-dimension of the skill distribution; by contrast, the task approach is a cross-cutting phenomenon across all skill levels. More specifically, our starting point is Acemoglu and Restrepo (2018), who focus on RBTC rather than SBTC.

[8] Therefore, SBTC falls short to guide this empirical evidence based on the monotonicity assumption that technology will increase the demand for skilled workers (*e.g.* documented by Autor and Dorn, 2013; Böhm, 2020; Goos and Manning, 2007). One reason for us focusing on the RBTC over SBTC is that the nature of the canonical model of SBTC lacks skill-replacing technologies and only allows for substitution and complementarity between skill groups (see Acemoglu and Autor, 2011, for a more detailed discussion). Therefore, SBTC does not necessarily provide a satisfactory understanding of the changes in earnings and employment distributions.

[9] Autor (2019) shows that polarisation in urban labour markets has contributed to: *(i)* the middle-skilled being pushed into performing traditionally low-skilled work; *(ii)* the middle-wage employment being disproportionately depressed in urban labour markets, hence average middle-skilled wages and the urban wage premium for this group have decreased; *(iii)* the creation of an excess supply of less-educated workers that depress middle-skilled wages across occupations and *geographic areas*.

[10] As our analysis is at the individual level and we do not directly focus on firms reducing production costs due to clustering, exploring sharing externalities in urban areas falls beyond the remit of this paper.

[11] *Vertical mismatch* is defined as education-occupation mismatch which occurs when a worker performs in an occupation that requires a degree lower than then the worker holds. In other words, the workers falling into these groups of mismatch are overqualified. *Horizontal mismatch* occurs when a worker performs an occupation that requires a field of education that is different from the one she/he obtained. However, independent of how well people feel they are matched to their current jobs (here what we mean now is what

We measure matching through overqualification due to the potential of technological change to increase overqualification.[12] It is argued that due to automation and offshoring hollowing out certain types of jobs (regardless of skill level though mostly those in the middle-skilled occupations), overqualification is likely to increase. This is because the qualified workers who lose their jobs or have fewer employment opportunities and also have less opportunity to upskill instantaneously, either exit the labour market or go down the occupational ladder. For instance, Autor (2019) points out that the compositional shift within education groups indicates a profound reallocation of middle-skilled workers in urban metros from middle to work requiring low skills, such as services, transportation and labourer occupations in the *densest areas*. This downward pressure exerted by higher-skilled occupation workers towards lower skilled ones is a likely mechanism to increase the overqualification in labour markets.[13] Given the compelling nature of routinisation towards rising skills mismatch, we analyse statically whether thick urban labour markets help improve matching. The effect of employment density on overqualification is indeed negative, suggesting better skills-jobs matching in cities, but only for non-routine workers.

RBTC implies that the wage gap between non-routine and routine task intensive jobs become wider once technological change is taken into account. We show that this effect will be reinforced in cities, and will be most pronounced for workers performing analytic jobs through a density premium. This wage-density premium may be explained by better skills-to-jobs matching and learning externalities for workers in non-routine task intensive jobs. Hence, wage inequalities seem to be wider in cities due to routine task intensity of jobs. Therefore, the heterogeneous effect of agglomeration economies does not necessarily operate through skill levels, but more so via task complexity.

The paper continues as follows. In Section 2 we outline our research framework, including a discussion on the data used, descriptive statistics, and the econometric framework. Section 3 reports the results, while Section 4 studies the mechanisms why the urban wage premium only pertains to workers in non routine task intensive jobs. We draw conclusions in Section 5.

2. Research framework

In this section we discuss the data used, report descriptive statistics and outline our econometric framework to measure the effects of density on routine task intensity of jobs and wages.

2.1. Data and variables

In this paper, we utilise administrative data combined with secondary data, including some historical series from the beginning of the 20th century. The administrative data are obtained from *Statistics Netherlands*.[14] These datasets include detailed information on work and residential locations of employees; the characteristics of employers; demographic and job characteristics of employees.

Our estimations are based on the LEED data from 2006 to 2016. To create this LEED data, we link several administrative datasets where *Dutch Labour Force Surveys* (*LFS*) constitute the core of the analysis. The construction of the data and the study period have two main restrictions. Firstly, the information on the location of the firms at the postcode level is only available from 2006 onwards. Hence, we limit the study period to 2006–2016. A postcode is a fine-grained spatial unit and covers about 15-20 addresses so it is comparable in size to a U.S. census block. Secondly, in the Netherlands, one can only observe the education and occupation levels of the employees from the *LFS*. Therefore, although using the *LFS* comes at the expense of not analysing the universe of employees, it provides us with the necessary information on skills mismatch, education and occupation as well as a wealth of information on demographic characteristics.

We retrieve information on employers' characteristics, annual earnings (employers' declaration of annual earnings before tax) of employees, job spells and exact days worked per job spell of each employee from *Tax Registers. Tax Registers* include the population of employees in the Netherlands based on employers' annual tax declaration and contain a unique job identifier, which is a combination of job, employer and time period. This allows us to identify each person by employer by job. We can then correctly link each employee in the *LFS* spells to a job, hence to his/her employer, *at the time of* the *LFS* interview. This linking removes the employees in the *LFS* who declare to be working but cannot be found in the tax registers, thus for these employees neither an employer nor a workplace can be identified. Finally, for the analysis addressing ability bias through focusing on siblings, we use a parent-child linkage dataset. The dataset, namely *Kindoudertab*, provides administrative records of the full population in the Netherlands that link children to their legal parents.

Our measure of the routine task intensity of occupations rests on adapting the SOC level measure used in Autor and Dorn (2013) to Dutch occupations at the highest possible resolution, which is 4-digit ISCO (ISCO'08). This level corresponds to ISCO subdivisions of minor groups and allows us to measure the degree of routinisation at the lowest level of occupational breakdowns. The SOC-ISCO cross-walk made available by the *U.S. Bureau of Labor Statistics*, facilitates a good match of all the occupations observed in the Dutch *LFS*. The degree of routine tasks in these occupations is quantified based on the importance scores used in the O*NET database. Based on the frequency of occurrence of certain tasks, hence the importance scores of these tasks per occupation, the tasks are categorised into commonly used categories of routineness. These categories, from routine to non-routine are: routine cognitive ($\mathcal{RC}$), routine manual ($\mathcal{RM}$), non-routine manual ($\mathcal{NRM}$), non-routine analytic ($\mathcal{NRA}$) and non-routine interactive ($\mathcal{NRI}$). Following Autor and Dorn (2013), let $\mathcal{R}_{ot}$ be the routine task intensity of an occupation o in year t:

$$\mathcal{R}_{ot} = \mathcal{RC}_{ot} + \mathcal{RM}_{ot} - \mathcal{NRM}_{ot} - \mathcal{NRA}_{ot} - \mathcal{NRI}_{ot}, \tag{1}$$

To reduce the data dimensionality, these five indicators (at the four-digit ISCO level) are combined into a single composite measure: the routine task intensity index (RTI). We normalise $\mathcal{R}_{ot}$ to have mean zero and unit standard deviation. Unlike Bacolod et al. (2009) and Grujovic (2018), who include a range of indices, we account for routine task intensity of occupations by constructing a single continuous RTI. Because the distinction between task groups across occupations is not mutually exclusive – meaning every occupation has a degree of all tasks as we show in Fig. 2 – the results based on multiple indices are somehow difficult to interpret due to the overlap between the task groups. Therefore, our analysis uses a single continuous metric of routinisation.[15]

they do by choice, and not based on the definitions explained above), from a welfare point of view, particularly for overqualification, mismatch is arguably a welfare loss as it is an educational investment with lower returns. The empirical evidence strongly suggests that the overqualified workers have relatively lower earnings than rightly matched workers and they are likely to experience lower job satisfaction due to under-utilisation of job-specific skills (Sanchez-Sanchez and McGuinness, 2015).

[12] One of the early contributions to link overqualification to technological change is by Mendes de Oliveira et al. (2000) and partly Violante (2002), while more recent and sophisticated work is presented by Beaudry et al. (2016).

[13] Indeed, Autor (2019) clearly describes the increased prevalence of middle-skilled occupation workers in low-skilled jobs, pointing out their significantly rising share in low-skilled occupation jobs, while upgrading to high-skilled occupations has been negligible.

[14] The datasets require a confidentiality agreement with *Statistics Netherlands* and are subject to special access conditions.

[15] Several papers in the literature have estimated wage returns to analytic, routine and manual tasks workers. Measuring 'routine' intensity of tasks through 3 separate indices is not straightforward. This approach is, for example, taken by Grujovic (2018) who estimates wage-density elasticities for task groups. She ranks them to benefit from agglomeration externalities in the order of analytic, routine and manual task workers, respectively. We show the robustness of our results to this categorisation. More specifically, in Appendix C.5 we include interactions of density (and the control variables) with indices capturing the intensity of manual tasks, analytic tasks and routine tasks.

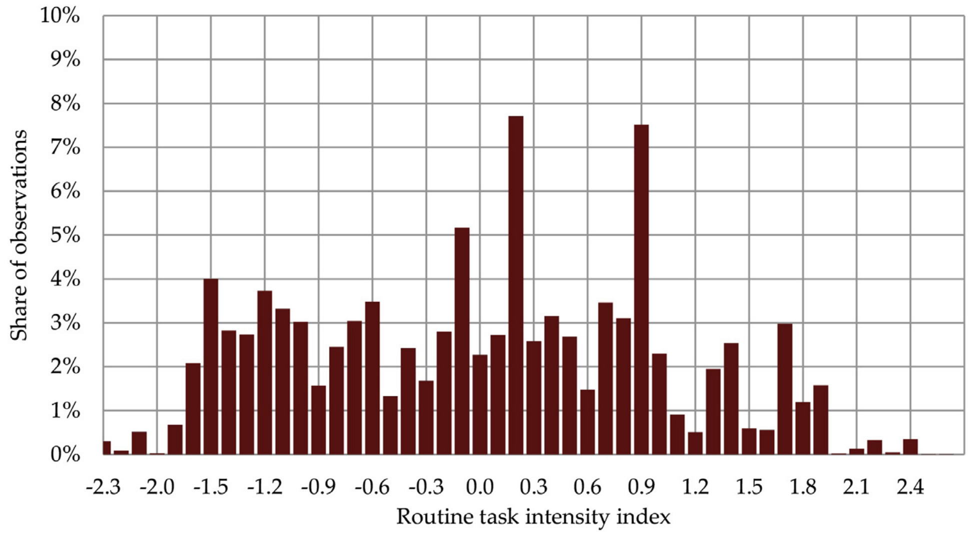


Fig. 1. Histogram of the routine task intensity index.

Table 1
Summary statistics.

	mean	std. dev.	min	max
Routine task intensity index	0	1	-2.301	2.440
Yearly wage *(in €)*	34,451	34,005	1003	1,550,000
Employment density *(weighted by commuting time)*	353,335	207,743	5427	998,533
Age	43.80	10.55	25	64
Female	0.495	0.500	0	1
Native born	0.908	0.290	0	1
Married	0.648	0.478	0	1
Divorced or widowed	0.0881	0.283	0	1
Single	0.264	0.441	0	1
Education – elementary skilled	0.0214	0.145	0	1
Education – low skilled	0.166	0.372	0	1
Education – medium skilled	0.433	0.495	0	1
Education – high skilled	0.353	0.478	0	1
Tenure	9.471	8.730	0.060	48.38
Work days	212.7	63.12	1	366.0
Firm size	6259	19,746	1	207,511
Employment in 1909 within commuting distance	29,886	52,951	0.00351	458,946
Share employment in 1909 in elementary occupations	0.108	0.0756	0.0108	0.368
Share employment in 1909 in low-skilled occupations	0.588	0.0687	0.343	0.790
Share employment in 1909 in medium-skilled occupations	0.255	0.0744	0.0453	0.518
Share employment in 1909 in high-skilled occupations	0.0492	0.0279	0.00375	0.116
Year of observation	2011	3.235	2006	2016

Notes: The number of observations is 473,322. For confidentiality reasons, the min and max refer to values where the top 10 and bottom 10 observations are excluded.

To construct our measure of potential accessibility to jobs, we make use of different data sources. First, we use information from *VUGeo-Plaza* on travel times by car between 4,033 Dutch neighbourhoods. We combine that with information on 4-digit postcode location which corresponds to neighbourhood-level. We then calculate the number of workers in a given distance (area) based on the travel times as *LFS* is a random sample of Dutch employees. Following Gaigné et al. (2017), we weight the number of workers by commuting time as below:

$$\mathcal{E}_{jt} = \frac{N_t}{\sum_{k=1}^{J} n_{kt}} \sum_{k=1}^{J} F[\tau_{jk}] n_{kt}, \tag{2}$$

where $\mathcal{E}_{jt}$ is the weighted number of workers at location j in year t. τ_{jk} is the travel time between home place j and work places $k = 1, \dots, J$, $F[\tau_{jk}]$ is the share of people who commute at least τ_{jk} minutes in the sample (excluding people who commute more than 2 hours).[16] n_{kt} is the number of workers at k in the *LFS* data. As we use survey data, from the *LFS* to calculate Eq. (2), we weight total number of workers in each *LFS* wave to match the total employment in the Netherlands, N_t.[17] This measure has the advantage of reflecting the actual accessible jobs for an employee, given the distribution of home-to-work commuting time and the road network, and mitigates methodological issues due to the arbitrary choice of spatial units (Briant et al., 2010).

We construct historic variables based on the 1909 census. For each of the 1,121 municipalities in 1909 we observe the number of workers in 1,571 occupations, divided into two classes: apprentice and master. For each occupation, we match the required level of education (in 4 classes), by relying on the variable in the *LFS* which determines the required education level for occupations. Inevitably, determining the level

[16] We display $F[\tau_{jk}]$ in Fig. A.2 in Appendix A.1.

[17] Because *LFS* is a survey, we may have some measurement error if n_{kt} is not random; however, this error is expected to be very small. Ancillary data on establishment-level employment is obtained from *ABR Regio*, which contains the universe of firm establishments and their exact postcode locations. We then calculate the number of accessible workers using the *ABR Regio* data only. The correlation with $\mathcal{E}_{jt}$ using *LFS* data is 0.98, suggesting that any selection bias is small.

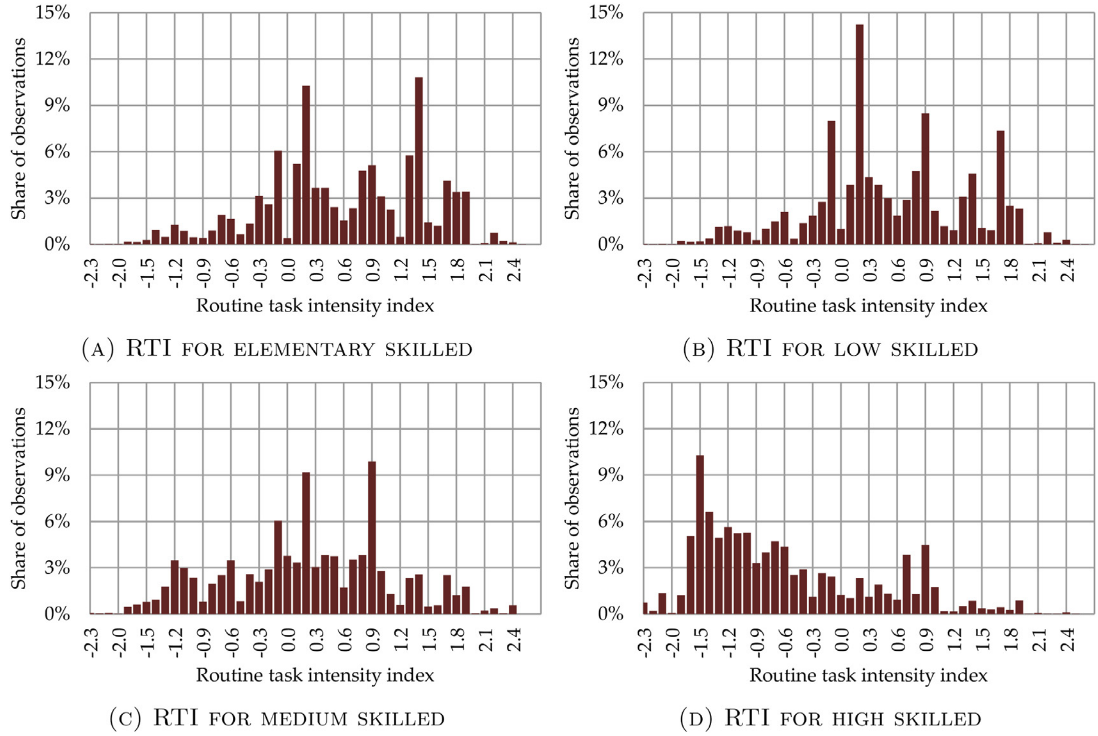


Fig. 2. Histogram of the routine task intensity index by skill level.

of education for each occupation in 1909 involves a degree of discretion. However, to the extent this measurement error is uncorrelated with the current spatial distribution of employment, this should not be a major concern.

We further use information on built-up areas from Knol et al. (2004) and information on employment at the municipality level in 1909 from the census.[18] Using information on the 1900 railway network from Gaigné et al. (2017) we calculate the number of jobs accessible within commuting time, given the cumulative distribution of commuting times:

$$\mathcal{E}_{j1909} = \sum_{k=1}^{J} F[\tilde{\tau}_{jk}] n_{k1909}, \tag{3}$$

where $\mathcal{E}_{k1909}$ is the number of jobs in 1909. $F[\tilde{\tau}_{jk}]$ is the share of people who commute at most $\tilde{\tau}_{jk}$ minutes, where $\tilde{\tau}_{jk}$ denotes the travel time using the railway network in 1900.

2.2. Descriptive statistics

We report descriptive statistics of our sample in Table 1. We have 473,322 observations between 2006 and 2016. By construction, the average routine task intensity index (RTI) has a mean of zero and unit standard deviation. We report a histogram of this variable in Fig. 1. The variable has few outliers. It appears that the lowest values of the RTI (*i.e.* RTI<−1) almost exclusively correspond to the occupations that require a high degree of analytic and cognitive tasks. Indeed, the correlation coefficient between the index analytic task intensity and the RTI is −0.89, while it is 0.30 for non-routine manual task intensity; and 0.70 for the routine task intensity of occupations.

We further show the distribution of the RTI within each skill group. Fig. 2 illustrates that regardless of the skill level, workers perform a degree of each task type. Unsurprisingly though, for elementary and low skill levels, many jobs are routine task intensive, and only very few have RTI levels lower than −1.

In particular, we find a high concentration of high-skilled workers in non-routine task intensive jobs. The highest concentration of high-skilled workers is for RTI values below −1.

In line with these observations Fig. 3 reflects a secular change in the Dutch labour market such that the share of non-routine task intensive occupations are increasing at the expense of those with routine-intensive tasks. More specifically, the share of non-routine (routine) task intensive jobs in 2000 was 0.43 (0.57) for 2000, while it was 0.48 (0.52) in 2016.

The median yearly wage is € 34,451 with significant variation. In our sample, 2.1% of the workers have elementary education, 16% are low skilled, 43% are medium-skilled and 35% have a bachelor's degree or higher.

During the study period 2006–2016, wages changed by 14.6%. Wage change is somewhat different for different levels of routine task intensive occupations. For workers performing routine tasks (RTI>0) the wage increase between 2006 and 2016 was 10.8%, while it was 14.7% for workers executed non-routine task (RTI<0). The wage change for high-skilled (*i.e.* employees holding a bachelor's degree or higher) *and* non-routine task intensive jobs in our sample was 15.6%, while it was only 7.4% for those in high-skilled routine task intensive jobs.

Our employment density measure is based on the number of jobs that are within commuting distance and ranges from just over 5 thousand (in the remote Wadden Islands) to almost a million in and around Rotterdam. The correlation with more standard measures of employment density is relatively low. For example, the correlation between the log of commuting-time weighted employment density and the log of employment density in the own neighbourhood is just 0.524. In Fig. 4 we plot *commuting time-weighted employment density* by the routine task intensity index (RTI) per occupation in the Netherlands. As shown in Fig. 4 high levels of density are associated with lower levels of the RTI, meaning that denser areas host workers in occupations that are less routine.

[18] We assume a uniform employment density within (historic) municipalities, which may overstate employment density slightly at the edges of the municipalities. However, municipalities in 1909 were much smaller and equal in size to large neighbourhoods. Hence, the implied measurement error will be small.

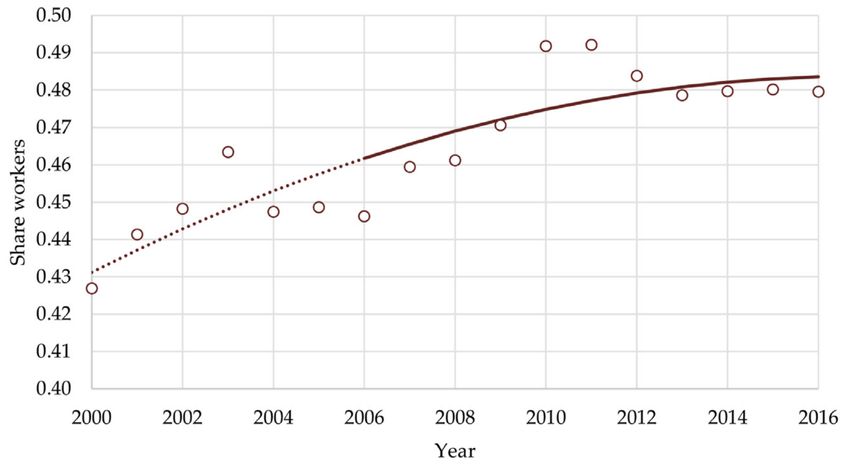


Fig. 3. Share of non-routine task intensive jobs. *Notes*: The dotted line is based on data that is not part of the main analyses. We consider jobs to be non-routine task intensive when RTI<0.

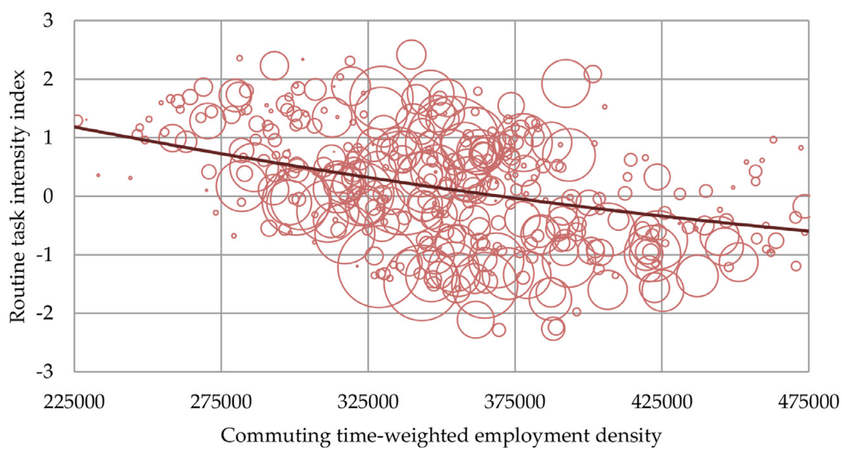


Fig. 4. Routine task intensity and employment density. *Notes*: Each circle represents the mean of Routine Task Intensity for a given level of employment density in each 4-digit occupation. The size of the dot is proportional to the number of workers in each occupation.

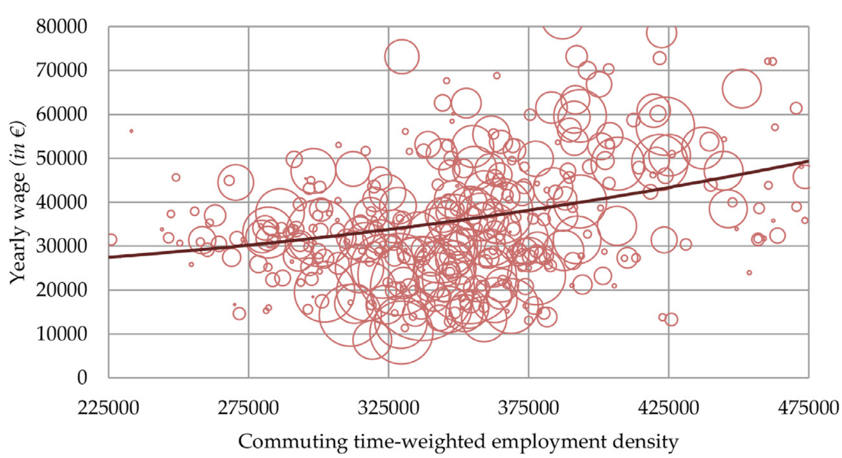


Fig. 5. Wages and employment density. *Notes*: Each circle represents the mean of wages for a given level of employment density in each 4-digit occupation. The size of the dot is proportional to the number of workers in each occupation.

In Fig. 5 we further plot the relationship between yearly wages and employment density by 4-digit occupation. We observe the familiar unconditional positive relationship between wages and employment density. That is, occupations that are concentrated in denser areas are usually better paid.

Table 1 also reports descriptives for our historical variables. Employment density in 1900 was much lower than it is during our sample period; on average the number of jobs that could be reached within the same commuting time was only 10% of the number of jobs that could be reached over the period 2006-16. Moreover, the skill distribution of occupations was more skewed to the right, implying that there were more low-skilled occupations. More specifically, the share of high-skilled people in 1909 was approximately 5%, while it is 35% today. Furthermore, the share of people in elementary occupations was somewhat higher: 10% in 1909 and 2.1% in our sample.

2.3. Econometric framework and identification

2.3.1. Parametric regressions – the effects of density

We hypothesise that the routine task intensity of jobs and wages are impacted by the number of accessible jobs within commuting time. That is, we expect to see that workers are more productive when they are surrounded by others from whom they can learn and complement their skills.

We observe a worker i residing in neighbourhood j in year t. We have two dependent variables: the routine task intensity index, $\mathcal{R}_{ijt}$, as well as the log of yearly wage, $\mathcal{W}_{ijt}$.

Further, let $\mathcal{E}_{jt}$ be the number of jobs a worker can reach within the commuting time from the home location, as defined by Eq. (2). The basic equation to be estimated then yields:

$$\{\mathcal{R}_{ijt}, \mathcal{W}_{ijt}\} = \beta \log \mathcal{E}_{jt} + \gamma X_{ijt} + \theta_t + \epsilon_{ijt}, \tag{4}$$

where β and γ are parameters to be estimated, X_{ijt} are worker, workplace, and neighbourhood characteristics, and θ_t are year fixed effects.

In line with a large literature on agglomeration economies (see *e.g.* Ahlfeldt et al., 2016; Combes et al., 2008; Melo et al., 2009), several endogeneity issues thwart a causal interpretation of β. The first issue is that omitted consumption amenities are correlated with $\mathcal{E}_{jt}$. As Gaigné et al. (2017) show, consumption amenities may disproportionately attract high-skilled workers that in turn may earn higher wages. Furthermore, there may be unobserved locational endowments that could be correlated with both $\mathcal{E}_{jt}$ and $\{\mathcal{R}_{ijt}, \mathcal{W}_{ijt}\}$. For example, certain regional policies may disproportionately attract certain firms that in turn require workers with certain skill levels. A third issue is that more productive workers and firms may sort themselves into dense urban areas (Behrens et al., 2014).

To mitigate endogeneity issues, we first include a wide range of control variables X_{ijt}. These controls include worker, neighbourhood and historical characteristics of the locations. Workers' characteristics include the level of education, age, gender, household size and composition, marital status and whether the worker is foreign-born. These worker characteristics should control for the fact that certain people (*e.g.* single adults without children) disproportionately sort themselves in dense areas. Job characteristics – such as the total days worked, tenure, and firm size, for example – control for the fact that workers may work in different types of jobs in cities and work, for example, more hours (see Rosenthal and Strange, 2008). All regressions include industry fixed effects $\lambda_{j\in s}$, and in the wage regressions, we further include ISCO 2-digit occupation fixed effects.

Detailed neighbourhood characteristics that control for locational quality include the share of land in historic districts, the share of open space and the share of water bodies in the neighbourhood. Controlling for amenities is important because certain workers may accept lower wages in urban areas because of higher amenity levels (Roback, 1982; Glaeser and Maré, 2001). We further control for sorting by calculating the share of young and elderly people, the share of foreigners, and household composition in each neighbourhood.

We also include travel-to-work area (COROP regions in the Netherlands) fixed effects $\eta_{j\in a}$. Essentially, labour market area fixed effects should also control for differences in skill composition between these regions. This implies that we identify agglomeration effects *within* regional labour markets.[19] We then have:

$$\{\mathcal{R}_{ijt}, \mathcal{W}_{ijt}\} = \beta \log \mathcal{E}_{jt} + \gamma X_{ijt} + \lambda_{j\in s} + \eta_{i\in a} + \theta_t + \epsilon_{ijt}. \tag{5}$$

In specifications in which wages are the dependent variable, we also include routine task intensity $\mathcal{R}_{ijt}$ to control for the fact that routine task intensity generates a task premium.

To further address endogeneity issues, and following Ciccone and Hall (1996); Ciccone (2002) and Combes et al. (2008) among others, we exploit historic data from 1909. The idea is that unobserved locational shocks are unlikely to be (strongly) correlated with our dependent variables over a century, whilst there is a strong autocorrelation of employment density over time. Hence, we use $\mathcal{E}_{j1909}$ as an instrument for current employment density.

We emphasise that we do not include worker fixed effects in these specifications, which is a common way to address the unobserved ability bias. Including worker fixed effects, though, is not without problems as the identification comes from workers that move between residential locations, who cannot be considered as a random subset of the population (Groot et al., 2014). Additionally, our data are based on a pooled cross-section of *Labour Force Surveys* therefore we can only trace very few workers over time, even fewer of whom would have changed residential location.

We, therefore, address sorting bias in other ways. Note that we use employment density in 1909 as an instrument for current employment density. The main criticism to such an instrument is that it is correlated with sorting patterns in 1909, which in turn may be correlated to current sorting patterns. For example, the capital of Amsterdam has attracted high-skilled workers for centuries. We then exploit the unique characteristic of our historic data – that is, we calculate the share of workers employed in high-skilled, medium-skilled, low-skilled and primary jobs within commuting time in 1909 and include those as control variables. This should address the issue that historic employment density is correlated with the sorting of more able workers in 1909. Note that we do not give a causal interpretation to these variables as they in fact may further capture local endowments of the area. Also, by controlling extensively for individual characteristics, by including industry, and travel-to-work-area fixed effects, and by instrumenting for employment density we think it is unlikely that any remaining ability bias is quantitatively important.

We also propose an alternative strategy that allows us to mitigate innate ability bias without needing workers to be mobile. We benefit from a parental-child linkage data, for the entire population in the Netherlands, which combines every worker in our sample to their legal parents (both mother and father). The literature shows that there is a strong genetic component to the abilities one embodies especially for cognitive intelligence (Haworth et al., 2009). Therefore, by constructing parent-child (worker) pairs in our sample, we are able to focus on siblings who share the same genetic ability traits. This approach has the advantage of exploiting the variation in density and wages based on siblings who reside in areas with different employment densities.

2.3.2. Semi-parametric regressions

Our primary interest is to identify how employment density effects differ for jobs with varying levels of routine task intensity. We hypothesise that jobs with low values of the RTI (*i.e.* analytic task intensive jobs) are particularly likely to reap the benefits of agglomeration, translating into a higher density premium with respect to wages and better skill matches.

Let us consider the following more flexible regression equation:

$$\mathcal{W}_{ijt} = f_{\mathcal{R}}\left(\log \mathcal{E}_{jt}, X_{ijt}\right) + \lambda_{j\in s} + \eta_{i\in a} + \theta_t + \epsilon_{ijt}, \tag{6}$$

where $f_{\mathcal{R}}(\cdot)$ implies that the impact of log employment density and control variables is a flexible function of the routine task intensity index.

We specify $f_{\mathcal{R}}(\cdot)$ by a locally linear function $f_{\mathcal{R}}(\cdot) = \beta_{\mathcal{R}} \log \mathcal{E}_{jt} + \gamma_{\mathcal{R}} X_{ijt}$. Hence, each 'local' coefficient is dependent on each unique value of the RTI. The simple alternative would be to estimate a separate regression for each value of the RTI. However, this would lead to very imprecise estimates. We therefore use kernel regressions to exploit the correlation between similar values of the routine task intensity index.

[19] One may argue that travel-to-work area fixed effects may partly absorb agglomeration effects. We show in Appendix C.7 that the estimated effects are essentially the same if we include lower resolution fixed effects.

For notational simplicity, let us suppress the fixed effects for now. The estimator is then:

$$(\hat{\beta}_{\mathcal{R}}, \hat{\gamma}_{\mathcal{R}}) = \underset{\beta_{\mathcal{R}},\gamma_{\mathcal{R}}}{\arg\min} \sum_{\ell=1}^{N} K\left(\frac{\mathcal{R}_{ijt} - \mathcal{R}_{\bar{\imath}jkt}}{h}\right) \times (\mathcal{W}_{ijt} - \beta_{\mathcal{R}} \log \mathcal{E}_{jt} - \gamma_{\mathcal{R}} X_{ijt})^2. \quad (7)$$

We specify $K(\cdot)$ to be a Gaussian kernel function:

$$K\left(\frac{\mathcal{R}_i - \mathcal{R}_\ell}{h}\right) = \frac{1}{\sqrt{2h\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{\mathcal{R}_i - \mathcal{R}_\ell}{h}\right)^2}. \quad (8)$$

Hence, the kernel function determines the vector of weights for a worker ℓ, which is between 0 and 1. It is 1 when another worker i has the same value for the RTI. The bandwidth h determines how 'smooth' the function to be estimated is. When $h \to \infty$, Eq. (7) collapses to a standard linear regression function. By contrast, if $h \to 0$ we estimate for each value of the RTI a separate (unweighted) regression, which would be inefficient. We employ the multivariate generalisation of Silverman's (1986) rule-of-thumb bandwidth, proposed by Li and Racine (2007) which is given by $h = 1.06N^{-\frac{1}{4+M}}$, where N is the number of observation in the sample and M the number of variables included in the non-parametric function to be estimated.

A last issue which needs attention is that Eq. (6) is a partially linear equation, where the fixed effects $\lambda_{j\in s}$, $\eta_{i\in a}$, and θ_t are linearly related to the dependent variable. We choose to employ Robinson's (1988) procedure to estimate the parameters of the model. This procedure separately regresses $\mathcal{W}_{ijt}$ and the dummies for the industrial sector, regions, and years on the non-parametric variables $\{\mathcal{E}_{jt}, X_{ijt}\}$, using local linear regressions. We then generate residuals for the dependent variable and dummies. The residuals of the dependent variable are then regressed on the dummy residuals using OLS, which identifies $\hat{\lambda}_{j\in s}$, $\hat{\eta}_{i\in a}$, and $\hat{\theta}_t$.[20] In the second part of the procedure, we replace the dependent variable $\mathcal{W}_{ijt}$ in Eq. (7) by $\mathcal{W}_{ijt} - \hat{\lambda}_{j\in s} - \hat{\eta}_{i\in a} - \hat{\theta}_t$ to obtain the coefficients of interest (*i.e.* $\hat{\beta}_{\mathcal{R}}, \hat{\gamma}_{\mathcal{R}}$).

3. Results

The results section first shows the effects of employment density on the task intensity of jobs in Section 3.1. Then, in Section 3.2 we proceed by investigating the effect of employment density on wages. Section 3.3 shows the key findings where we estimate the effect of employment density on wages depend on the routine task intensity of jobs. In Section 3.4 we provide various robustness checks.

3.1. Routine task intensity and employment density

We first document the relationship between tasks and employment density. That is, we hypothesised that agglomeration economies, proxied by employment density, should foster new ideas and innovativeness. This should create a more diverse array of jobs involving less routine tasks. Hence, we expect employment density to be inversely related to the routine task intensity of a job.

The results are reported in Table 2. In column (1) we estimate a naive specification where we regress the RTI on employment density and year fixed effects. We find a negative effect of employment density on the RTI, in line with Fig. 4. Doubling employment density leads to a decrease in the RTI of $(\ln 2 - \ln 1) \cdot 0.0980 = 0.068$ standard deviations. Hence, in line with Michaels et al. (2016), we find that non-routine tasks are more concentrated in cities. Let us investigate whether this result is either an agglomeration effect – meaning denser areas tend to produce jobs with higher levels of complexity – or is simply explained by the sorting of highly able workers and certain types of occupations into denser areas.

To do so, in column (2), Table 2, we include a wide array of workers' characteristics (such as education level, age, gender, marital status, household composition), job characteristics (such as the size of the firm where the worker is currently employed, tenure, the total number of days worked in the current job), as well as sector fixed effects. These additional controls almost halve the employment density coefficient yet have limited repercussions for the qualitative results, as the effect of density on the routine intensity of jobs remains strongly negative.[21] In column (3) we introduce locational controls (*i.e.* the share of land that is part of a historic district, or the share of open space in the neighbourhood, the share of young, elderly and foreigners in the same neighbourhood) and travel-to-work-area fixed effects. Hence, we identify the effect of density within labour markets, while controlling for amenities and sorting. We then find that a 100% increase in employment density is associated with a decrease in the RTI of 0.017 standard deviations.

One may argue that current employment density is correlated with locational endowments and sorting so that it does not capture solely the effect of agglomeration economies on the routineness of a job. To mitigate the issue we use a familiar strategy discussed earlier and instrument the log of current employment density with employment density in 1909 and the square of employment density in 1909. Looking at the first-stage Kleibergen-Paap F-statistic, it is unsurprising that the instruments are strong (and have the expected signs).[22] When considering the results in column (4) in Table 2, we now find a substantially stronger effect of employment density on routine task intensity of jobs. This may imply that routine task intensive jobs may be more concentrated in otherwise more attractive locations with higher densities. The coefficient implies that the doubling of employment density leads to a 0.16 standard deviation increase in RTI.

What could be the reason for the effect becoming much stronger when instrumenting for employment density? First, there may be measurement error in employment density, *e.g.* because we use free-flow travel times and because we extrapolate employment observed in the *LFS* to the full population. To the extent this measurement error is random, this means that the coefficient in OLS specifications is biased towards zero. We would argue that this measurement error is unlikely to be correlated to historical employment density. Hence, instrumenting should lead to a stronger (negative) coefficient. We develop this argument more formally in Appendix B.1, where we show that the higher the R^2 of employment density on controls and fixed effects, the more the measurement error will be amplified.

Second, in the past, routine task intensive jobs as a whole were more abundant than non-routine task intensive jobs (Michaels et al., 2016). Most likely, routine jobs were concentrated in otherwise attractive locations, as firms are likely to take up first the most attractive locations. Because of agglomeration economies and sorting, non-routine jobs are now replacing routine jobs in dense areas. However, because replacement is slow, we may still see manufacturing in otherwise expensive locations. Hence, to the extent that routine jobs are still disproportionately located at otherwise attractive locations, we expect to find a (strong) underestimate in the OLS-specifications.

One likely concern when using historical employment density as an instrument for current employment density is that the employment density in 1909 is correlated to current sorting patterns. That is, historically dense places may have attracted high-skilled and high-ability people for centuries and they may be still doing this today. Hence, to address this issue, we calculate the share of people employed in high-skilled, medium-skilled, low-skilled and primary jobs in 1909 as controls. We show in column (5), Table 2, that the effect of employment density on the RTI is only slightly, but not significantly, stronger than the previous

[20] Under regularity conditions, Robinson (1988) shows that the coefficient is a $\sqrt{N}$-consistent and asymptotically normal estimator for the linear parameters.

[21] We note that the worker and job characteristics all have a statistically significant effect on the routineness of jobs. For example, females and foreigners are more likely to perform jobs involving routine tasks. Workers that are employed longer at the firm usually participate in more routine task intensive jobs. The RTI is also positively correlated with firm size while being negatively correlated with the number of workdays in the last year. The exact coefficients are available upon request.

[22] We report first-stage results in Appendix C.1.

Table 2
Routine task intensity and employment density.

Dependent variable: routine task intensity index		*+ Worker, job characteristics*	*+ Location characteristics*	*Instrument for density*	*+ 1909 skill composition*
	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS
Employment density within commuting distance *(log)*	-0.0980***	-0.0428***	-0.0249***	**-0.2072*****	**-0.2312*****
	(0.0058)	(0.0024)	(0.0049)	**(0.0277)**	**(0.0392)**
Share employment in 1909 in low-skilled occupations					0.3422**
					(0.1333)
Share employment in 1909 in medium-skilled occupations					0.0712
					(0.0705)
Share employment in 1909 in high-skilled occupations					0.5239*
					(0.2947)
Worker and job characteristics	No	Yes	Yes	Yes	Yes
Location characteristics	No	No	Yes	Yes	Yes
Industry fixed effects	No	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	No	No	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	473,322	473,322	472,947	472,947	472,947
R^2	0.0057	0.3296	0.3308		
Kleibergen-Paap *F*-statistics				106.7	65.18

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

Table 3
Employment density and wages.

Dependent variable: log of yearly wage		*+ RTI*	*+ Worker, job characteristics*	*+ Location characteristics*	*Instrument for density*	*+ 1909 skill composition*
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS	(6) 2SLS
Employment density within commuting distance *(log)*	0.0903***	0.0684***	0.0574***	0.0255***	**0.0983*****	**0.0941*****
	(0.0040)	(0.0032)	(0.0017)	(0.0028)	**(0.0159)**	**(0.0216)**
Routine task intensity index		-0.2233***	-0.0566***	-0.0560***	-0.0559***	-0.0559***
		(0.0014)	(0.0010)	(0.0010)	(0.0010)	(0.0010)
Share employment in 1909 in low-skilled occupations						0.0046
						(0.0697)
Share employment in 1909 in medium-skilled occupations						-0.0112
						(0.0358)
Share employment in 1909 in high-skilled occupations						0.0719
						(0.1511)
Worker and job characteristics	No	No	Yes	Yes	Yes	Yes
Location characteristics	No	No	No	Yes	Yes	Yes
Occupation 2-digit fixed effects	No	No	Yes	Yes	Yes	Yes
Industry fixed effects	No	No	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	No	No	No	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	473,322	473,322	473,321	472,946	472,946	472,946
R^2	0.0116	0.0913	0.7591	0.7615		
Kleibergen-Paap *F*-statistics					106.5	65.05

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

specification. Overall, agglomeration externalities are complementary to technological improvements so we believe that the strong effects that we find are not unreasonable.

3.2. Wages and employment density

We now move forward to the effect of employment density on wages. We begin by replicating the well-established relationship in the literature that wages increase with the density of areas. In column (1) of Table 3, where we only control for year fixed effects, we find an elasticity of 0.090. In column (2) we control for the routine task intensity of occupations by including the RTI. This should capture whether a task-premium exists for less routine task intensive occupations (Goos et al., 2014; Michaels et al., 2016). In other words, this implies that we control for the sorting of non-routine occupations in dense areas. Indeed, we find that the impact of employment density slightly decreases. The RTI has a negative association with wages, which is in line with the literature (Ozgen, 2020). That is, a standard deviation increase in the RTI is associated with a wage change of $e^{-0.22} - 1 = -19.7\%$ so wage returns to routine tasks seem to be substantially lower.

In column (3) we include a wide array of individual and job characteristics, as well as industry fixed effects. Although this strongly increases the R^2, it does not materially influence the estimated agglomeration elasticity.

However, when we control for locational characteristics, including amenities and the neighbourhood demographic composition, as well as travel-to-work-area fixed effects, the effect of employment density becomes somewhat lower. The coefficient in column (4)

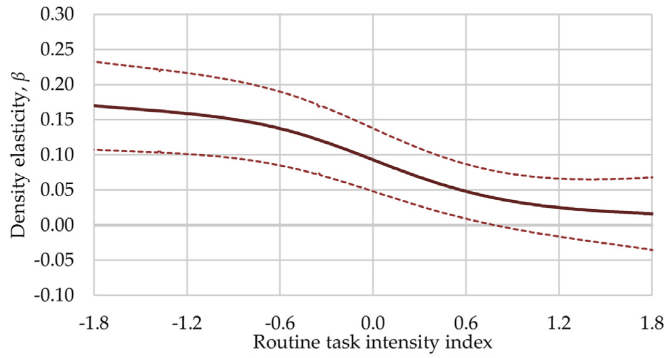


Fig. 6. RTI, density and wages. *Notes*: The regression is based on a control function approach, where the first-stage error is inserted as a control variable in the second stage. We control for the share of skills in 1909, worker, job and location characteristics, as well as occupation 2-digit, industry, travel-to-work area and year fixed effects. The dashed lines constitute 95% confidence intervals, based on 250 cluster-bootstrapped replications.

implies that doubling the employment density increases wages by 2.5%, which is in line with the previous literature on agglomeration economies (see *e.g.* Combes et al., 2008). Also, when adding controls and fixed effects, the effect of RTI becomes somewhat lower: a standard deviation increase in the RTI is associated with a wage decrease of 5.4%.

Columns (5) and (6) present the IV estimates. First, in column (5) we mitigate the issue of unobserved locational endowments being correlated with the instrument historical density in 1909. Column (6) repeats the same regression while including additional controls capturing 1909 skill composition to take sorting on skills into account. The density elasticity becomes stronger. More specifically, the preferred estimate in column (6) implies that doubling employment density leads to a wage increase of 6.5%.[23]

The impact of RTI remains fairly similar across specifications, where we introduce locational controls, suggesting a 5.6% decrease in wages with respect to a standard deviation increase in RTI. These results suggest that denser areas do not only provide higher wages to urban workers but also workers in more complex jobs receive a *task premium* for working in non-routine task intensive occupations.

3.3. Routine task intensity and agglomeration economies

While we have so far established the returns to wages in dense areas by the average level of routinisation, our estimates do not tell us how these returns vary by the *level of the RTI* and hence by the type of tasks workers perform.

To explore the employment density effects on wages across routinisation levels, we make use of the semi-parametric techniques proposed in Section 2.3.2. Semi-parametric estimations enable us to scrutinise this relationship in a much more flexible way. We present these results in Fig. 6.

We show the estimated density elasticity with respect to wages, $\beta_{\mathcal{R}}$, for workers for different levels of the RTI in our sample. We find that the employment density elasticity is only statistically significant and positive for lower levels of RTI. Hence, only workers in *non-routine* task intensive occupations, which include *both* non-routine analytical and non-routine manual task occupations, seem to enjoy a wage premium in dense urban areas. It is plausible that only workers who are at the lower end of the RTI distribution, *i.e.* those who perform highly complex analytical tasks, reap benefits of density. This is because the analytical skills of these workers not only complement the technological advances, but also by using these technologies, those skills may require workers to have more frequent interactions with other workers in the city. This, in turn, spurs faster learning (Davis and Dingel, 2019). Further, cities may foster a better matching between jobs and skills, leading to higher wages, as a thick labour market may improve matching. However, most prior research has predominantly focused on skill levels, proxied by the (formal) degree of education. Consequently, it has been silent about how wage-density elasticities vary over different levels of routine task intensity.[24]

The literature on routine biased technological change provides ample evidence of a large variation in wage returns to tasks, even within the same skill groups (see Michaels et al., 2016). Therefore, one may argue that the finding of agglomeration effects for analytic jobs (*i.e.* for low levels of the RTI) is spurious because most analytic jobs are also high-skilled. To address this concern, in Fig. 7 we revisit the estimated relationship between agglomeration economies and RTI by splitting the sample between high-skilled and medium/low-skilled workers. It is shown that both high and medium/low-skilled workers in non-routine jobs benefit from agglomeration economies, while the effect dissipates quickly for low-skilled workers who perform routine tasks. Also for high-skilled workers, we observe that the agglomeration elasticity becomes lower for higher levels of the RTI. Hence, these results confirm that only workers who are at the lowest levels of RTI gain from the agglomeration economies, which in turn lead to higher wages. Importantly, these findings hold *within* skill groups, suggesting that heterogeneity in returns to agglomeration economies cannot be solely explained by skills, but by tasks.[25]

3.4. Sensitivity checks

We subject these results to a wide range of robustness checks in Appendix C. First, we report first-stage results in Appendix C.1, where we show that employment density in 1909 is positively and strongly related to current employment density. In Appendix C.2 we split the sample by different levels of RTI and estimate our preferred specification (see column (6) in Table C.2). The results confirm the more general semi-parametric findings reported in the previous section: we only detect a density premium for wages for workers employed in non-routine task intensive jobs.

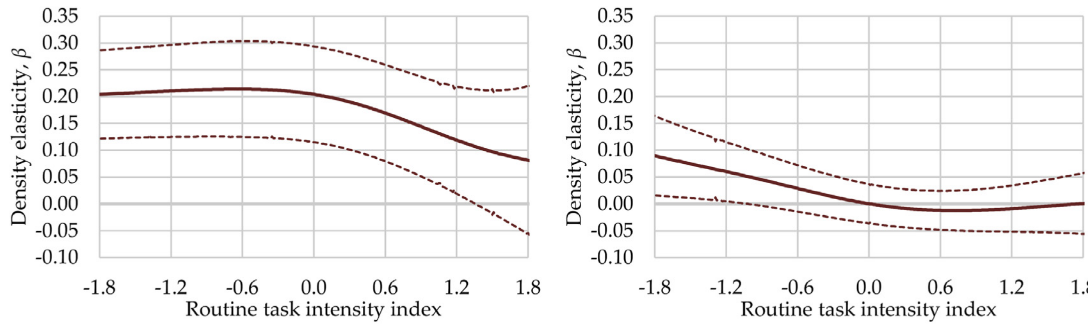
In Appendix C.4, to control for sorting based on unobserved ability, we limit our sample to sibling workers who carry similar genetic ability traits. This siblings sample allows us to create an almost ideal setting to properly identify the effect of the innate ability of workers that potentially correlates with the error terms. Our results presented in Table C.3 exhibit similar point estimates in order of magnitude to those in Table 3 for both OLS and IV estimations. However, due to a substantial decrease in the number of observations the confidence intervals become too large to draw strong conclusions. Still, the point estimates signal that ability bias does not seem to be a major issue.

Appendix C.5 studies whether our results are robust to alternative definitions of the task intensity of jobs. More specifically, in line with other work (such as Grujovic, 2018), we distinguish between analytic, manual and routine intensive task occupation scores. We find that analytic task intensive jobs benefit considerably more from agglomeration economies, while routine and manual task intensive jobs benefit less from density. In other words, our findings do not completely align with

[23] We provide reasons why the IV estimates may be stronger than the OLS estimates in the previous subsection. These also apply here.

[24] Note that the results of the semi-parametric regressions are fully consistent with and suggestive of the same pattern that we report in Appendix C.2, which are standard IV results, where we split the sample for the different levels of RTI.

[25] In Appendix C.3 we also estimate the relationship between the RTI and the density elasticity for low and medium-skilled separately. The results are somewhat imprecise but display the same pattern.



(A) High skilled (B) Medium/low skilled

Fig. 7. RTI, density and wages by skill level. *Notes*: The regressions are based on a control function approach, where the first-stage error is inserted as a control variable in the second stage. We control for the share of skills in 1909, worker, job and location characteristics, as well as occupation 2-digit, industry, travel-to-work area and year fixed effects. The dashed lines constitute 95% confidence intervals, based on 250 cluster-bootstrapped replications.

Table 4
Matching and density.

	Dependent variable: overqualified			*Dependent variable: horizontally mismatched*		
	(1) OLS	(2) 2SLS	(3) 2SLS	(4) OLS	(5) 2SLS	(6) 2SLS
Employment density within commuting distance *(log)*	-0.0172***	**-0.0593******	**-0.0552******	-0.0013	**0.0092**	**0.0160**
	(0.0020)	**(0.0096)**	**(0.0133)**	(0.0018)	**(0.0081)**	**(0.0115)**
Routine task intensity index	0.1816***	0.1815***	0.1815***	-0.0345***	-0.0344***	-0.0344***
	(0.0007)	(0.0007)	(0.0007)	(0.0010)	(0.0010)	(0.0010)
Skill shares in 1909	No	No	Yes	No	No	Yes
Worker and job characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Education field fixed effects	No	No	No	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	445,501	445,501	445,501	352,917	352,917	352,917
R^2	0.3085			0.2508		
Kleibergen-Paap *F*-statistics		108.1	65.76		106.6	65.09

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

Table 5
Employment density, wages and learning.

Dependent variable: log of yearly wage	*Labour market experience*		*Locational experience*		*Both*
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS
Employment density within commuting distance *(log)*	**0.0010**	**0.0640**	**0.0781******	**0.0985******	**0.0082**
	(0.0025)	**(0.0522)**	**(0.0215)**	**(0.0240)**	**(0.0247)**
Employment density within commuting distance *(log)* × Labour market experience *(log)*	**0.0363******	**0.0485******			**0.0385******
	(0.0030)	**(0.0074)**			**(0.0033)**
Employment density within commuting distance *(log)* × Locational experience *(log)*			**0.0110******	**0.0126******	**0.0064******
			(0.0017)	**(0.0020)**	**(0.0021)**
Employment density within commuting distance *(log)* × Age *(log)*		**-0.0246**			
		(0.0191)			
Labour market experience *(log)*	-0.3939***	-0.5088***			-0.4186***
	(0.0396)	(0.0905)			(0.0438)
Locational experience *(log)*			-0.1429***	-0.1661***	-0.0893***
			(0.0220)	(0.0254)	(0.0263)
Routine task intensity index	-0.0565***	-0.0564***	-0.0553***	-0.0553***	-0.0556***
	(0.0010)	(0.0010)	(0.0010)	(0.0012)	(0.0012)
Skill shares in 1909	Yes	Yes	Yes	Yes	Yes
Worker and job characteristics	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	460,383	460,383	468,211	361,248	352,874
Kleibergen-Paap *F*-statistics	33.90	23.09	32.49	32.09	23.20

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

the ranking of the task groups (namely analytic, routine and manual) as in Grujovic (2018) to benefit from agglomeration economies, except for workers performing analytic tasks. Given the strong positive correlations of the RTI with analytic task intensity, and the strong negative correlation of routine and manual task intensity, these results support our approach to reduce dimensionality and collapse these three indices into one index measuring routine task intensity.

In Appendix C.6 we further consider sorting of firms in dense areas. Combes et al. (2012) show that in cities competition is tougher, allowing only the most productive firms to survive. More specifically, multinational firms, in which jobs are more likely to require analytic intensive tasks, may be disproportionally clustered in cities. We take into account the multi-nationality of the firms by splitting the sample between firms with a Dutch headquarters (HQ) and an HQ outside the Netherlands. It is shown that the main result (*i.e.* agglomeration economies only being relevant for non-routine task intensive jobs) is not affected by the location of the HQ of a firm.

We subject the baseline specification (see column (5) in Table 2 and column (6) in Table 3) to additional robustness checks in Appendix C.7. *(i)* We use population in 1900 as an instrument instead of employment density in 1909, *(ii)* we use the share of locally born people living in the area in 1909 – an indicator for mobility – as an alternative instrument for current employment density, *(iii)* we control for amenities in 1909, *i.e.* the share of built-up land and water in 1909, and, finally, *(iv)* we include less-detailed province instead of travel-to-work-area fixed effects. Our findings are not overturned by these additional sensitivity checks.

4. Potential mechanisms

In this section, we aim to explore what could explain the striking pattern depicted in Fig. 6, in which we show that agglomeration economies are only relevant for workers employed in non-routine task-intensive jobs. The literature discussed in the Introduction suggests two main reasons – *matching* and *learning* – why workers particularly in non-routine jobs may benefit from agglomeration economies. We use familiar proxies for matching and learning used in the literature and aim to explore whether matching and learning could indeed explain the variation in the wage-density elasticities across different levels of the RTI. We emphasize that we interpret the results presented here as suggestive, as we are partly constrained by the limited availability of microdata.

4.1. Matching

The literature points out that better matching of firms and workers occurs in large cities due to the potential complementarity of market size and the skill requirements of jobs (Duranton and Puga, 2004). The benefits of better matching may, however, be only relevant for certain jobs. For example, Fallick et al. (2006) show that on the job mobility is highest among computer industry workers, while mobility patterns do not hold for the workers in other industries. The particular reason why we focus on overqualification is that routinisation has significant implications for the reallocation of workers by education levels across occupation groups (see Autor, 2019). Therefore, we explore the effects of density and RTI on skills mismatch. In doing so, we consider two dimensions of skills mismatch. The first is 'vertical' mismatch, defined as the mismatch when a worker's obtained degree is higher than the degree that is required for performing her/his current job (Büchel and van Ham, 2003; Leuven and Oosterbeek, 2011; Kampelmann and Rycx, 2012; Jauhiainen, 2011; Berlingieri, 2019). This type of mismatch is known as 'overqualification' in the economics of education literature (see Duncan and Hoffman, 1981).

Our measure of overqualification is determined by the gap between the actual degree obtained (in 4 education classes) and the required degree for the occupation in the current job (in 4 aggregate occupation classes). The required level of education for a worker's current occupation is based on an inquiry of required degrees from 1996; therefore our measure is not impacted by contemporaneous changes in the labour market. If workers hold a degree of education that is higher than the required one in their current job, they are assigned to be overqualified. However, although our analysis spans between 2006–2016, the occupation level required degree is only available between 2006 and 2011. Based on the 2011 period, we, therefore, calculate the average required degree for each 4-digit occupation for post-2011. This may introduce some measurement error so we offer additional sensitivity checks by solely relying on the spells between 2006 and 2011 where we have the information on required education available for each job and each year.[26] Appendix C.8 provides a more elaborate discussion.

The key advantage of this measure of overqualification is that it is not dependent on the self-assessment of employees, which is common in the literature (see *e.g.* Berlingieri, 2019). One may argue that in some circumstances, the employees may be able to indicate the job requirements thoroughly and this may provide instantaneous information. However, this may not be the case in many occupations. According to Hartog (2000), for example, the concern with the self-assessed measures is that the respondents might have a tendency to upgrade the status of their position, or to overstate the hiring requirements. Therefore, self-reported measures may be subject to bias. The method using a self-reported measure also draws upon self-reflective experience rather than being based on objective indicators. Using a job analyst approach for measuring mismatch in our analysis is also consistent with the underlying methods (*i.e.* O*NET experts' views) used for our measure of routinisation.[27]

Approximately 24% of the people in our sample are overqualified – meaning that their obtained degree is higher than the required degree for their current job. This is comparable to Berlingieri (2019) for Germany, who found that 19% of the people are overqualified. In Fig. 8 we plot the relationship between overqualification and employment density by 4-digit ISCO occupations. We observe a negative relationship: occupations with a higher share of overqualified workers are typically in lower density areas.

The second type of mismatch we explore is called 'horizontal' mismatch, following Boualam (2014). This metric measures the mismatch between the field of education required for the current job and the field of education obtained by the worker. For example, when a worker is trained as an engineer but ends up being a bank manager, this could be considered a mismatch, although the required education level may be comparable.[28]

We then calculate horizontal mismatch $\mathcal{H}_{ijt}$ for worker i in education field f in occupation o:

$$\mathcal{H}_{ijt} = 1 - \frac{N_{fo}/\sum_f N_{fo}}{\max_f(N_{fo}/\sum_f N_{fo})}, \tag{9}$$

[26] Based on the required degree information we obtained from the Dutch LFS, the required education may be different within each occupation. Therefore, we calculate the probability of a worker to be overqualified, given the degree she/he obtained. An example would be as follows: suppose we observe that in the data for 60% of financial controllers a Bachelor's degree is required, while for 40% a Master's degree is necessary. Let's further assume that a worker holds a Master's degree. Our variable measuring overqualification is then equal to 0.6. Furthermore, we run robustness checks where we only keep occupations for which the required education applies to essentially one skill level so that measurement error is negligible.

[27] A further concern is that of frequent job changes, especially among the younger workers. The literature provides evidence that overqualification and tenure is negatively associated, which lends support for on-the-job learning. We observe each worker's firm-specific job history at a given time. We, therefore, address this issue by including tenure with the current employer as a covariate in all of the regressions reported. Further, given that the average tenure in our dataset is around 7 years, this negative association is unlikely to be the result of frequent job changes. We also further include age in all regression to control for age and age-related training effects. To the extent there is still measurement error in the overqualification measure, conditional on the control variables, we expect the remainder to be random and therefore should not affect our estimated coefficients.

[28] We consider 13 education fields, among others including, agrarian, technical, economic, language and culture, management, education. Please note that in the Netherlands in low-skilled education (*e.g.* in high schools) allows for specialising in education fields.

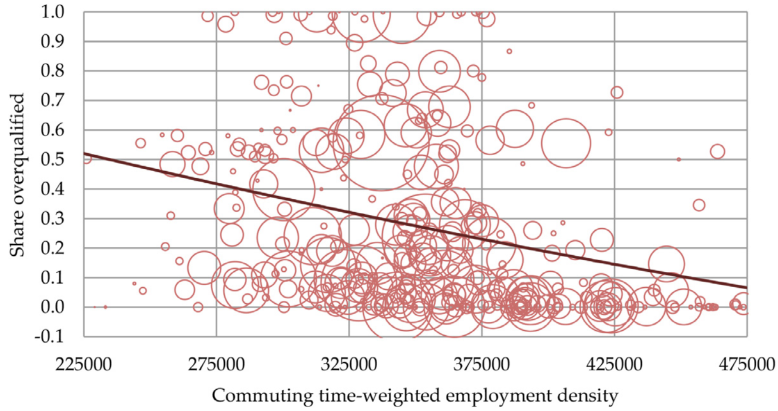


Fig. 8. Overqualification and employment density. *Notes*: Each circle represents the mean of overqualified workers for a given level of employment density in each 4-digit occupation. The size of the dot is proportional to the number of workers in each occupation.

where N_{fo}, which are the number of workers in our sample in education field f and occupation o. Hence, the numerator is the share of people in an education field f in the occupation field o. We further rescale the measure to take into account heterogeneity across occupation fields by dividing this share by the maximum value that is observed over all education fields within this occupation. Hence, when $\mathcal{H}_{ijt}$ is close to one this means that within an occupation, a certain education field is rarely observed, which suggests that there is a mismatch. Note that because field of education is only available until 2013, our dataset will be somewhat smaller.

We first investigate whether our proxies for mismatch are higher in denser places, by replicating Eq. (5), but replacing the dependent variable with our proxies for vertical or horizontal mismatch, respectively. For the overqualification regressions, we omit workers with elementary education because they are at the bottom of the occupational ladder therefore they cannot be overqualified. In the regressions where we focus on horizontal mismatch, we also include the field of education fixed effects. We report the results in Table 4.

As in Berlingieri (2019), we confirm in column (1) that employment density does reduce the probability of being overqualified. In other words, workers in large cities are likely to match to jobs requiring the level of education that they have obtained. The coefficient indicates that a 10% increase in employment density decreases overqualification by 0.17 percentage points. In column (2) we instrument for density and in column (3) we further control for the share of workers with different skill levels in 1909. The coefficient in the preferred specification in column (3) indicates that a 10% increase in employment density decreases overqualification by 0.55 percentage point so this effect is non-negligible.

Furthermore, as expected, the routine task intensity of jobs seems to increase overqualification. One explanation is that technological change makes many routine jobs redundant. Given that routine-intensive jobs are outsourced or offshored, workers in more routine jobs may end up accepting jobs that are below their education level. As a result, routinisation can increase the prevalence of overqualification; as we find in the estimations.

In columns (4)-(6) in Table 4 we focus on horizontal mismatch, as in Boualam (2014). Because we do not observe the field of education after 2013, we have fewer observations. We do not find an effect of density on horizontal mismatch, which is in contrast with Boualam (2014). A likely reason for that is we are better able to control for worker and location characteristics and further focus on horizontal mismatch *within* travel-to-work-areas. If anything, we find that for jobs with a higher degree of routineness, horizontal matching is lower. An explanation might be that more analytic task intensive jobs require less specialised studies, or put differently, the acquired skills are more widely applicable to perform a larger range of occupations.

If matching is a potential explanation for the results we find earlier, mismatch should not only be lower in denser areas but also only be relevant for low levels of routine task intensity. This would imply that only workers in analytic jobs would experience a reduction in overqualification, and hence receive a respective wage premium. To investigate this, Panel A of Fig. 9 indeed shows that for low levels of the RTI, the effect of employment density is negative for overqualified workers, suggesting a decrease in overqualification for workers who perform non-routine task intensive occupations. For those who are in routine occupations, higher density does not yield better matching of jobs and workers. Hence, these results show that a reduction in mismatch only applies to non-routine task intensive jobs, and may therefore be an explanation for the findings shown in Fig. 6. That is, part of the wage-density premium that accrues to analytic workers may be due to better matching in dense areas.

In Panel B of Fig. 9 we show that the effect of density on reducing horizontal mismatch does not vary for different levels of RTI. Hence, horizontal mismatch is unlikely to be a mechanism that explains the wage-density premium found for workers in analytic task intensive jobs.

In Appendix C.8 we further show that also within skill groups we observe that mismatch is mostly reduced for workers in non-routine task intensive occupations. For horizontal mismatch, we consistently do not find a conclusive effect. Appendix C.8 also considers two alternative measures of overqualification. First, we only use data between 2006 and 2011 for which there is exact data on the required level of education for every year. Second, we only keep occupations for which the required degree is for over 90% applying to one skill level, which minimises measurement error issues. We find very similar results, although it seems that the estimates display slightly smaller coefficients. All in all, we find that density reduces skills mismatch only for non-routine task intensive jobs. Hence, better matching in denser areas for non-routine jobs is one useful explanation for why the wage-density premium is decreasing in RTI.

4.2. *Learning*

In line with Davis and Dingel (2019), we expect learning opportunities particularly for analytic task intensive workers to be greater in cities. However, De la Roca and Puga (2017) show that learning externalities may not be instantaneous but may increase with experience

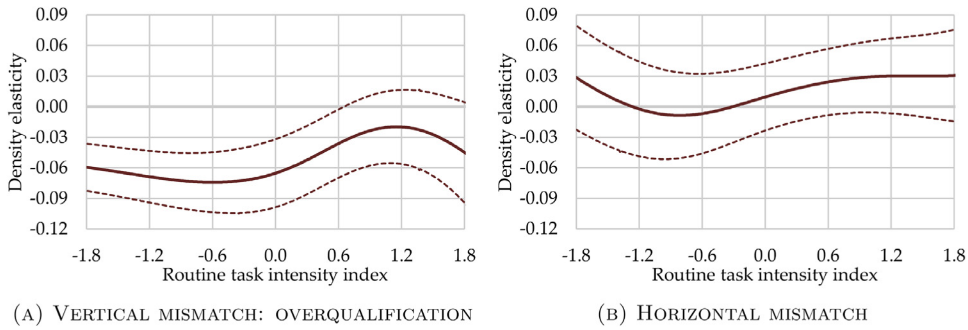


Fig. 9. RTI, density and mismatch. *Notes*: The regressions are based on a control function approach, where the first-stage error is inserted as a control variable in the second stage. We control for the share of skills in 1909, worker, job and location characteristics, as well as occupation 2-digit, industry, travel-to-work area and year fixed effects. The dashed lines constitute 95% confidence intervals, based on 250 cluster-bootstrapped replications.

(*i.e.* proxied by the time spent in large cities).[29] To capture learning effects we use two somewhat crude proxies. Slightly improving on the commonly used potential experience measure in the literature (see *e.g.* Blau and Kahn, 2003), the first one we refer to is (labour market) experience, which we define as age minus the average age of graduation given the degree of education minus 4, that is the age when compulsory education starts in the Netherlands.[30] Of course, this measure of experience is related to age. We do not see this as a problem as experience – and therefore learning – rises with age. However, we also estimate regressions where we allow agglomeration economies to vary by age.

A second measure for learning is in the spirit of De la Roca and Puga (2017). They argue that bigger cities provide workers with opportunities to accumulate valuable experience. They define experience acquired as the years a worker lives in a certain city, which we refer to as 'locational experience'. Because of data restrictions and the fact that our measure is based on an almost continuous geography, we choose a simple measure of experience: the number of years since a person is employed in a certain municipality.[31] The idea is that generally workers who just arrive in a new area does not have a business network and cannot exploit the benefits of learning. If the worker is employed for longer she/he likely has established a business network to reap the benefits from agglomeration economies. Hence, over and above Davis and Dingel's (2019) static model, we assume that learning possibilities are dynamic and increase with the years of experience.

In order to study whether our proxies for learning are relevant, we first test whether our proxies are correlated with the intensity of agglomeration economies. We then re-estimate the wage-density regressions (see Eq. (5)), but now include interactions of density and the variable of interest. The results are reported in Table 5.

In column (1) of Table 5 we investigate whether labour market experience commands higher wage-density premia. We find considerably lower agglomeration economies for inexperienced workers. For example, for someone who just enters the labour market, the elasticity of wages with respect to density is essentially zero, while after 24 years (*i.e.* the median experience), the elasticity is $0.0010 + \ln(24) \times 0.0363 = 0.116$. Hence, it seems that more experienced workers benefit from agglomeration economies. Note that the coefficient on experience is negative. However, because we also control flexibly for age, we do not think the interpretation of this effect is obvious. Although experience comes with age, one may be concerned that we measure just an age effect. In column (2) we, therefore, include an interaction of employment density and the log of age. We find that the interaction effect is statistically insignificant, while the effect of density interacted with experience remains positive and is even somewhat stronger.

Further, in column (3) we show that employment density is also increasing in locational experience. If someone is working in a certain location for 10 years or more, the density elasticity is $(0.0110 \times \ln(10)) = 0.0253$ (*i.e.* 32%) higher. Hence, in line with De la Roca and Puga (2017) we find that there is a complementarity between density and locational experience, for example, because it takes time before workers can exploit local business networks. One may be concerned that locational experience is measured with error because our data goes back only until 1999. We, therefore, exclude censored observations in column (4). While this means that we omit about 25% of the observations, the results are similar and confirm that agglomeration economies still increase in locational experience.

In column (5) we include both proxies for experience. We confirm that agglomeration economies are increasing in labour market experience and locational experience, although the latter effect is partly absorbed by labour market experience.

In Fig. 10 we split our sample between experienced and inexperienced workers and replicate Fig. 6 to see whether learning effects can explain the finding that agglomeration economies decrease in RTI.[32]

In Panel A of Fig. 10, we show that for workers with substantial labour market experience (*i.e.* more than 10 years), the relationship between RTI and the wage-density premium is almost the same as depicted in Fig. 6. Strikingly, in Panel B we show that workers with little labour market experience do not benefit at all from agglomeration economies; neither workers performing analytic task intensive occupations, nor those in routine ones. This strongly suggests that learning effects in cities are not instantaneous (as assumed in Davis and Dingel, 2019), but increase with the years of work experience. These results also suggest that our main result may be explained by learning; because if inexperienced workers in analytic task intensive occupations did not accumulate valuable experience, they do not benefit from agglomeration economies. Workers in routine task-intensive jobs do not benefit from learning, whether they are inexperienced or experienced.

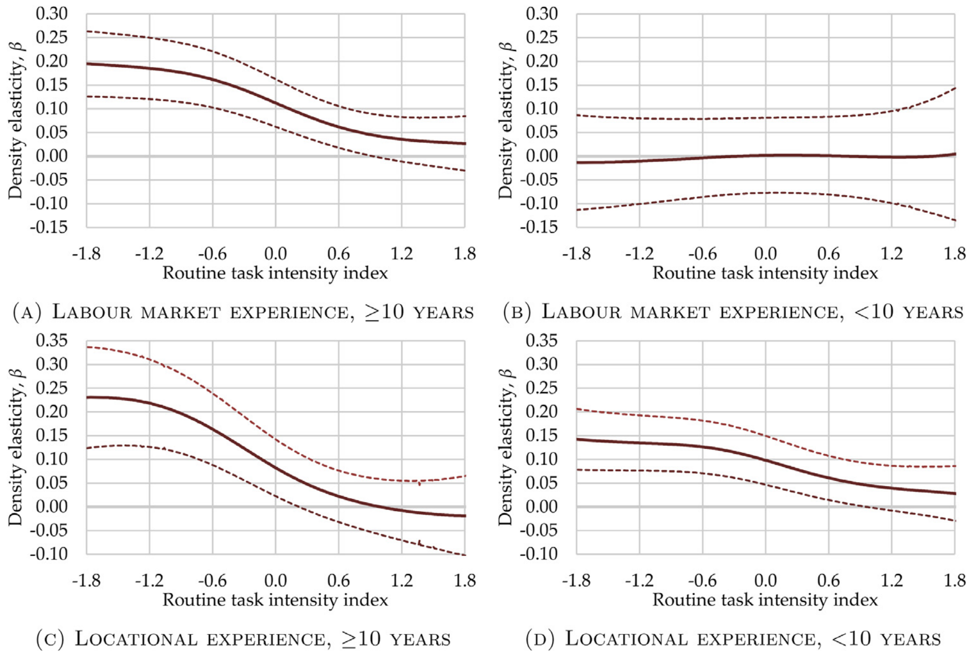
We also test whether locational experience matters for agglomeration economies, by splitting the sample between inexperienced and experienced workers. In Panels C and D of Fig. 10 we observe that the relationship between RTI and the wage-density premium is flatter for inexperienced workers, in line with the previous results. However, given the limitation of our measure, the results are not as pronounced in mag-

[29] A recent work by Peters (2021) also supports this hypothesis in the case of Germany.

[30] We do not observe the actual graduation age, which forces us to use the average age. The ages of graduation for the different degrees are the following: primary=12, secondary=16, vocational=20, bachelor's degree=21, and master's degree=23.

[31] Calculating this locational experience in our data is not straightforward. First, we do know the *firm* at which a worker works, but not the *establishment*. We, therefore, calculate the commuting time between the home location and all establishments of a firm. We then assume that the worker is employed at the *nearest* establishment. Second, although we know at what firm people work before 2006, the *location* of firms where people work is only known from 2006 onwards. Hence, *before* 2006, we assume that firms did not move. We believe that this assumption is pretty innocuous because the share of firms moving each year is small (just about 4%), but most firms (more than 75%) are moving within the municipality (Van Oort et al., 2007). Third, because our administrative data go back only to 1999, for comparability across the sample we have censored observations for people working longer than eight years in certain locations in 2006, nine years in 2007, etc. We show the robustness of our results by excluding censored observations.

[32] We test for different thresholds, but the results presented below do not change considerably. These results are available upon request.



(A) Labour market experience, ≥10 years (B) Labour market experience, <10 years

(C) Locational experience, ≥10 years (D) Locational experience, <10 years

Fig. 10. RTI, density and wages by experience. *Notes*: The regressions are based on a control function approach, where the first-stage error is inserted as a control variable in the second stage. We control for the share of skills in 1909, worker, job and location characteristics, as well as occupation 2-digit, industry, travel-to-work area and year fixed effects. The dashed lines constitute 95% confidence intervals, based on 250 cluster-bootstrapped replications.

nitude as those for labour market experience. Still, we do think that these results are also suggesting that learning effects may be important in explaining why mostly analytic task-intensive workers benefit from agglomeration economies.

5. Conclusions

Wheeler (2001) and Baum-Snow and Pavan (2013) document widening wage inequalities across differently skilled groups particularly between urban and rural regions. Michaels et al. (2016) show that even after controlling for workers' observable characteristics, large wage inequalities remain in urban *versus* rural areas. Canonical models of labour markets that solely distinguish between returns among skilled *versus* unskilled workers cannot fully explain these. Although many studies establish a strong relationship between wages and density, the heterogeneity in the effect through which density augments productivity remains understudied. Moreover, the effect of density on the spatial concentration of tasks and the subsequent implications for wages are not yet known.

To the best of our knowledge, this is the first employee-level study that uses a continuous measure of routinisation to indicate routinisation as a potential source for widening wage inequalities in urban areas. By using rich administrative and survey data from the Netherlands over the perio 2006–2016, we study the effect of agglomeration economies on wages and overqualification when RBTC is explicitly taken into account. We introduce an improved measure of employment density, by using the cumulative distribution of commuting time-weighted distances in the Netherlands. Our measure is based on the actual number of jobs within an area that is a maximum of 2 hours commuting time from a worker's home location. Therefore, it has the advantage of not being confined to administrative borders or an arbitrary spatial unit, in addition to being defined based on each worker's accessibility to jobs. We also construct a measure of routine task intensity by adapting O*NET importance scores for routine tasks into all jobs provided in the *LFS* at the 4-digit ISCO level. By doing so we observe the degree of routineness for every occupation in the Netherlands.

Our findings suggest that employment density significantly increases the number of non-routine task intensive jobs. We find that the magnitude of higher wage returns to density is consistent with the extant literature. However, crucially, we show that density premia only exist for workers in non-routine task intensive occupations. In other words, urban workers enjoy a wage premium due to the complementarity between non-routine tasks and technological change in denser areas and an additional task premium seems to exist for non-routine workers.

To explain these findings, we provide suggestive evidence on the potential channels, such as a better matching of workers and jobs resulting in a lower prevalence of overqualified workers. We show that improved matching is only relevant for workers in non-routine task intensive jobs. Moreover, we show that learning externalities accrue only to analytic workers with more years of experience in the local labour market. We interpret the latter as circumstantial evidence that learning effects are important in explaining why mostly workers in non-routine task intensive occupations receive density premia. All of these findings withstand attempts to mitigate endogeneity issues and a significant number of sensitivity checks.

Urban areas are conducive to the proliferation of complex jobs. Workers in more complex, analytic task intensive, occupations then enjoy higher wage returns, are more resilient to automation effects, and have a higher probability of finding jobs that match with their education. On top of that, they receive a task premium. Given that these forces are absent for routine workers the convergence of wages in urban areas is unlikely, suggesting that wage and skill inequalities are likely to rise.

Policy makers need to recognise *(i)* the strong trend in non-routine jobs clustering in urban areas, and *(ii)* workers performing non-routine analytic jobs are the main beneficiaries of the technological change through routinisation. To manage the transition to better technologies, policy makers may invest in more resources that would enable routine workers to up-skill and adapt faster. Accordingly, better technology adaptation would mean higher earnings for routine workers thereby reducing wage inequalities, and positive externalities in terms of higher quality jobs and better matching for cities as shown by our findings.

Credit Author Statement

Both authors have contributed equally to the paper in terms of methodology, conceptualization and data analysis.

Appendix A. Data

A1. Network distances

We obtain information on travel times by road from the *VUGeoPlaza* which enable us to calculate travel time τ_{jk} between two locations j and k. The dataset provides information on actual driving speeds for every major street in the Netherlands. The actual speeds are usually substantially below the free-flow driving speeds due to traffic lights, roundabouts and intersections. For each neighbourhood, we calculate the total driving time for each location pair. Alternatively, we calculate for each location pair the Euclidian distance and assume an average speed of 10km/h. We then choose the lowest of the network travel time and Euclidean travel time (to avoid any possible inconsistencies in the network).

We report in Fig. A.1 the relation between Euclidian distance and travel time. For short distances (<5km) we observe that it is sometimes faster not to make use of the network so that the Euclidian travel speed is used. Euclidian distance is shorter for the long commutes compared to travel time based on the road network.

The overall correlation between Euclidian distance and travel time is relatively low ($\rho = 0.495$). However, if we focus on trips shorter than 45 minutes (which holds for most commutes), the correlation is much higher ($\rho = 0.633$), even if we exclude trips for which we use the Euclidian travel speed ($\rho = 0.599$).

In Fig. A.2 we display the cumulative distribution of commuting times based on calculations of Gaigné et al. (2017). It indicates that approximately 50% of workers have commutes shorter than 25 minutes, while less than 5% commute for longer than one hour. We use this distribution to calculate employment density measures as in Eqs. (2) and (3).

A2. Historical data

To instrument employment density we use historic instruments. We obtain data on employment in each municipality in 1909 using the Dutch census, which provides us with the employment for 1,571 occupations. For each occupation, we have the number of jobs at the apprenticeship level and the 'master' level. There were 1,121 municipalities in 1909 (as compared to only 418 in 2011).

To determine the skill level (*i.e.* elementary, low, medium, high) for each occupation we use the current ISCO classification of required education for each occupation and determine manually what the required skill level is for each occupation. We always assume that apprentices are one skill level below masters.

Because the historic data is not available at the neighbourhood level, but at the municipality level, to calculate the historic population in each neighbourhood, we calculate the share of each neighbourhood in each municipality and attribute the population proportionally. One may be concerned that municipalities could be quite large, but they were much smaller than they are today and about the size of just four neighbourhoods.

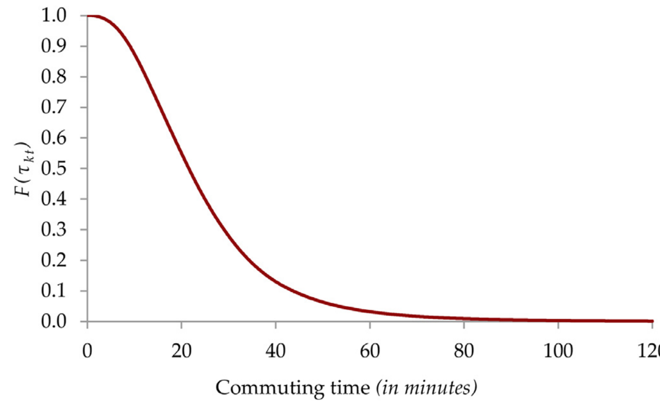


Fig. A.2. Commuting time distribution.

We calculate commuting-time weighted employment density and the share of employment in each skill level in 1909 to determine historic travel times between neighbourhoods. We further use information on railway stations from Koopmans et al. (2012). We enrich these data by adding missing stations from various sources on the internet and create a network with travel times. To approximate the speed, we fit a regression of the length of (current) railway segments between stations on the current observed travel time on the railway network. Based on historic sources, it appears that the average speed is about 50% of what it is currently (about 70km/h). This provides us with the necessary information to predict the travel time between stations in 1909.

In Fig. A.3 we plot the relationship between current travel times and travel times by rail in 1900. The correlation is 0.460, while it is 0.634 for trips that are currently shorter than 45 minutes. The correlation is of course higher if we look at the correlation between current travel times by train and travel times by train in 1900 ($\rho = 0.683$ for all trips and $\rho = 0.885$ for trips shorter than 45 minutes).

A3. Other descriptive statistics

In Fig. A.4 we report trends of the main dependent variables. We observe in Panel A of Fig. A.4 that wages have increased steadily since 2000. In 2016 the median nominal wage in our sample was 42% higher as compared to 2000.

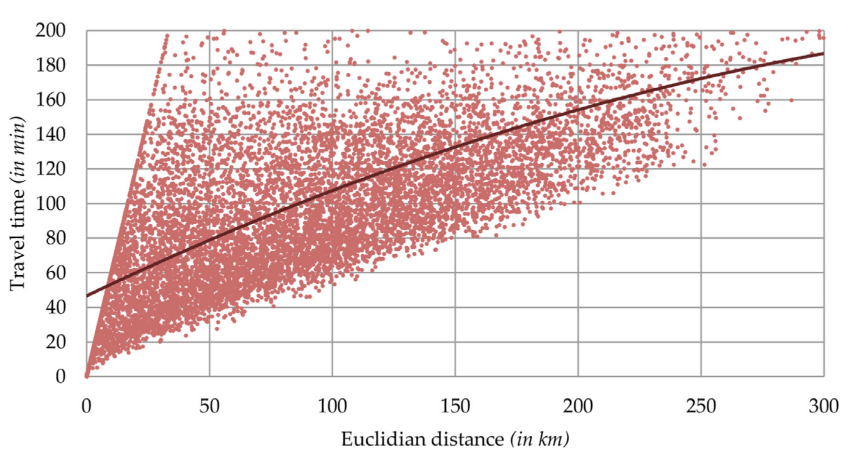


Fig. A.1. Euclidian distance and travel time.

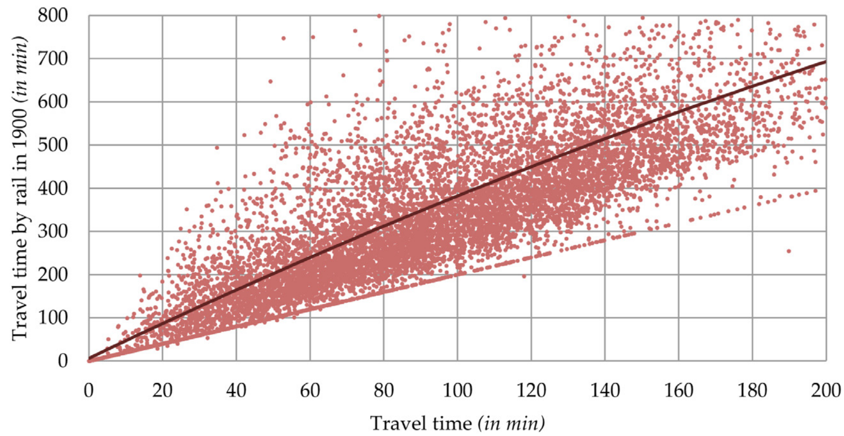


Fig. A.3. Travel time in 1900.

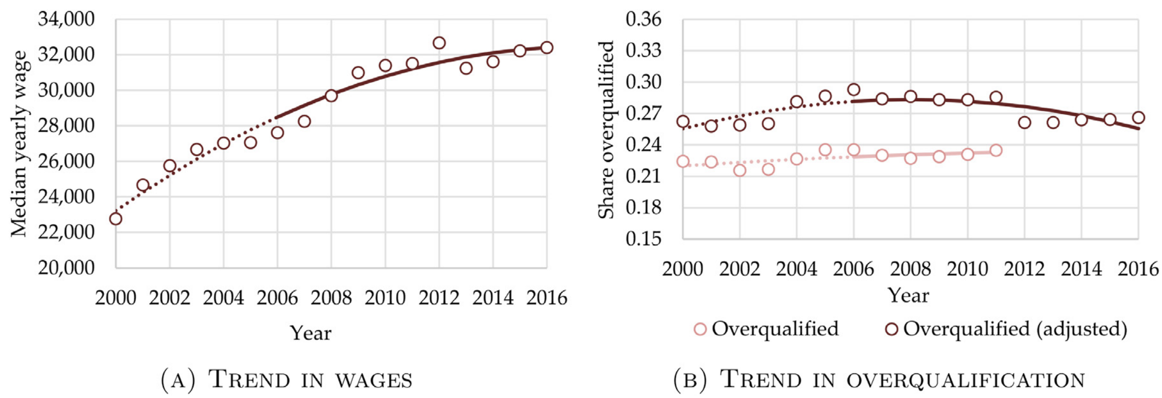


Fig. A.4. Trends in main dependent variables.
Notes: The dotted line is based on data that is not part of the main analyses.

By contrast, in Panel B of Fig. A.4 we do not find strong evidence for any trends in overqualification, neither when using the adjusted measure for the whole time period, nor when using the measure based on the required education for each job. Recall that the latter measure is only available before 2012.

Appendix B. Measurement error

B1. Measurement error

Here we explain why most of the estimates capturing the impact of employment density on respectively the RTI, wages and overqualification are somewhat stronger once we instrument for employment density.

Let us first make the common assumption that the measurement error in employment density is assumed to be uncorrelated with the 'true' unobserved value of employment density. Let us write $\log \mathcal{E}^*_{jt} = \log \mathcal{E}_{jt} + \upsilon$, where υ denotes the measurement error. Cameron and Trivedi, 2005 then show that the estimated parameter using OLS is equal to:

$$\text{plim}\hat{\beta}_{OLS} = \beta \left[1 - \frac{\sigma^2_\upsilon}{\sigma^2_{\log \mathcal{E}^*}(1 - R^2_{\log \mathcal{E}^*|X,\theta_t}) + \sigma^2_\upsilon} \right]. \tag{B.1}$$

where σ^2_υ denotes the variance of the (unobserved) measurement error. Further, $R^2_{\log \mathcal{E}^*|X,\theta_t}$ is the R^2 of an auxiliary regression of $\log \mathcal{E}^*_{jt}$ on controls X_{ijt} and fixed effects, while $\sigma^2_{\log \mathcal{E}^*}$ denotes the variance of the unobserved true value of employment density after removing the linear influence of X_{ijt} and fixed effects.

The more controls we add to the specification the higher the $R^2_{\log \mathcal{E}^*|X,\theta_t}$ is and the lower the conditional variance of $\log \mathcal{E}^*_{jt}$. This implies that the measurement error is magnified when adding more controls.[33] Compared to other agglomeration studies we add many more controls related to household characteristics, as well as fixed effects related to occupation, the labour market area, and time. Hence, $R^2_{\log \mathcal{E}^*|X,\theta_t} \rightarrow 1$ is likely to be higher; and hence potential measurement error in the OLS estimates is likely more severe. This explains why a relatively mild measurement error in employment density may cause a large difference between the OLS and 2SLS estimates, once we control for location characteristics and travel-to-work-area fixed effects.

Appendix C. Additional regression results and sensitivity

C1. First-stage results

We use employment density in 1909 as an instrument for current employment density. In Table C.1 we report first-stage results. Columns (1) and (2) refer to the regressions where the RTI is the dependent variable (see Table 2), and columns (3) and (4) to the regressions where yearly wage is the dependent variable (see Table 3).

The picture that emerges is clear. Employment density in 1909 is a strong instrument for the log of current employment density. Both the linear term and the quadratic term are statistically significant, indicating that there is a non-linear relationship between employment in 1909 and current employment density. Let us consider the estimate in column (4). If we evaluate the elasticity of employment density in 1909 with current employment density, it is 0.139 for the median value of employment in 1909. However, if we consider the elasticity for the 5th percentile value

[33] In case that $R^2_{\log \mathcal{E}^*|X,\theta_t} \rightarrow 1$, $\hat{\beta}_{OLS} \rightarrow 0$

Table C.1
First-stage results.

Dependent variable: log of employment density				
	RTI regressions		*Wage regressions*	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Employment density in 1909 within commuting distance / 50,000	0.2011*** (0.0141)	0.1465*** (0.0132)	0.2006*** (0.0140)	0.1462*** (0.0132)
(Employment density in 1909 within commuting distance / 50,000)^2	-0.0233*** (0.0020)	-0.0164*** (0.0018)	-0.0233*** (0.0020)	-0.0163*** (0.0018)
Share employment in 1909 in low-skilled occupations		2.1846*** (0.2955)		2.1810*** (0.2953)
Share employment in 1909 in medium-skilled occupations		0.9667*** (0.1874)		0.9636*** (0.1872)
Share employment in 1909 in high-skilled occupations		4.9198*** (0.6204)		4.9127*** (0.6198)
Routine task intensity index			-0.0009 (0.0008)	-0.0005 (0.0008)
Worker and job characteristics	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes
Occupation 2-digit fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	472,947	472,947	472,946	472,946
R^2	0.8479	0.8571	0.8480	0.8571

Notes: Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

Table C.2
RTI and agglomeration economies.

	RTI < -1	$-1 \leq$ *RTI* < 0	$0 \leq$ *RTI* < 1	*RTI* ≥ 1
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS
Employment density within commuting distance *(log)*	**0.1679***** **(0.0346)**	**0.1663***** **(0.0364)**	**0.0141** **(0.0213)**	**-0.0009** **(0.0390)**
Routine task intensity index	-0.0269*** (0.0078)	-0.1091*** (0.0057)	-0.0347*** (0.0043)	-0.1058*** (0.0078)
Share of skills in 1909	Yes	Yes	Yes	Yes
Worker and job characteristics	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes
Occupation 2-digit fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	99,434	126,893	179,436	67,181
Kleibergen-Paap *F*-statistics	64.32	59.48	66.43	51.15

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

of employment density in 1909, the implied elasticity is 0.146, while it is only 0.061 for the 95$^{\text{th}}$ percentile value of employment in 1909. This indicates that the places that were the least populated in 1909 grew faster (in percentage terms) than the places that were already populated. Given the observation that current historic amenities prevent a strong population growth in city centres, this seems to make sense.

Columns (2) and (4) in Table C.1 further show that part of the effect of employment density in 1909 on current employment density is due to sorting. Generally speaking, we find that places with concentrations of high-skilled occupations grew faster than places with many elementary or low-skilled occupations.

C2. Parametric RTI regressions

In Table C.2 we analyse the wage-density premium across a range of the RTI categories. As stated earlier in the Introduction, the occupations that have an RTI value less than −1 are almost exclusively those that require non-routine analytic tasks, while the occupations that have an RTI value equal to or larger than 1 specialise in lower-indexed routine tasks. We explore the returns to wages by task levels in Table C.2.

As in previous estimations, the RTI value is negative and highly significant for all RTI sub-categories meaning that a high routinisation level is associated with lower wages. Employment density for analytic workers is positive and significant suggesting an additional wage premium, which is about 70% higher than the average prediction that was 0.094. Employment density is also positive for the workers in column (2), and these workers also receive a task premium. Strikingly, and in line with Fig. 6, a spatial wage premium is absent for all other workers.

C3. Agglomeration economies, skill level and routine task intensity

In Section 3.3 we documented that agglomeration economies are stronger for less routine task intensive occupations. One may, however, argue that there may be a strong correlation between the skill level and the routine task intensity of a job. Hence, what we may measure is that agglomeration economies are stronger for high-skilled jobs rather than less routine task intensive jobs. To test this more explicitly we re-

Table C.3
Addressing the ability bias.

	Dependent variable: RTI		*Dependent variable: log of yearly wage*	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Employment density within commuting distance *(log)*	-0.0227	**-0.2041**	0.0321***	**0.0827**
	(0.0209)	**(0.1456)**	(0.0107)	**(0.0687)**
Routine task intensity index			-0.1095***	-0.0518***
			(0.0035)	(0.0047)
Share employment in 1909 in low-skilled occupations		0.2051		0.0298
		(0.4714)		(0.2175)
Share employment in 1909 in medium-skilled occupations		0.0647		-0.0652
		(0.2662)		(0.1213)
Share employment in 1909 in high-skilled occupations		0.4785		0.1345
		(0.9937)		(0.4549)
Family fixed effects	Yes	Yes	Yes	Yes
Worker and job characteristics	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes
Occupation 2-digit fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	51,184	51,184	51,184	51,184
R^2	0.6962		0.8799	
Kleibergen-Paap *F*-statistics		89.36		88.83

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

Table C.4
Employment density and wages by task groups.

Dependent variable: log of yearly wage	*Analytic*	*Manual*	*Routine*	*All*
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS
Employment within commuting distance *(log)*	**0.0824*****	**0.0954*****	**0.0900*****	**0.0750*****
	(0.0199)	**(0.0216)**	**(0.0206)**	**(0.0191)**
Employment within commuting distance *(log)* × Analytic occupation score	**0.0442*****			**0.0345*****
	(0.0102)			**(0.0111)**
Employment within commuting distance *(log)* × Manual occupation score		**-0.0143***		**-0.0200****
		(0.0076)		**(0.0078)**
Employment within commuting distance *(log)* × Routine occupation score			**-0.0363*****	**-0.0226****
			(0.0094)	**(0.0101)**
Skill shares in 1909	Yes	Yes	Yes	Yes
Worker and job characteristics	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes
Occupation 2-digit fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	472,946	472,946	472,946	472,946
Kleibergen-Paap *F*-statistics	28.22	29.26	27.32	13.54

Notes: We include the indices as well as interact all control variables with the indices capturing manual, routine and analytic task intensity. **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

estimate semi-parametric regressions, but for more detailed education levels. Because we have few observations for workers in elementary occupations we focus on low-skilled and medium-skilled (as the results for high-skilled workers already have been reported in Fig. 7). Fig. C.1 shows the results.

In Panel A of Fig. C.1 we report the results for the elasticity of wages with respect to employment density for workers with maximally a high-school degree. We find that the estimated coefficients are, unsurprisingly, imprecise. Hence, the point estimates are not statistically significantly different from zero. However, the general pattern is that the density elasticity is stronger for non-routine jobs. For the workers with vocational training (shown in Panel B of Fig. C.1), we find that the wage-density premium is (marginally) significant for non-routine task intensive jobs, while it is statistically insignificant and essentially zero for RTI>0.5. Hence, these results confirm the results reported earlier in Fig. 7.

C4. Addressing ability bias

Our estimations would adversely be affected by unobserved ability bias if the ablest and most talented workers sort into denser areas. To correct for this potential bias, we benefit from a novel dataset that permits the linking of workers to their parents. In doing so, we are able to focus on siblings that carry similar genetic ability traits. Because LFS is a survey, this procedure significantly reduces the number of obser-

Table C.5
Employment density, wages and HQ location.

Dependent variable: log of yearly wage			
	All firms	*Dutch firms*	*Non-Dutch firms*
	(1) 2SLS	(2) 2SLS	(3) 2SLS
Employment within commuting distance *(log)*	**0.0766****	**0.0671***	**0.0908**
	(0.0363)	**(0.0389)**	**(0.0614)**
Routine task intensity index	-0.0574***	-0.0500***	-0.0722***
	(0.0021)	(0.0026)	(0.0039)
Skill shares in 1909	Yes	Yes	Yes
Worker and job characteristics	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes
Occupation 2-digit fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Number of observations	99,092	70,355	28,734
Kleibergen-Paap *F*-statistics	56.05	56.56	49.08

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

Table C.6
Sensitivity analysis: routine task intensity.

Dependent variable: routine task intensity index				
	Population in 1900	*Share locals in 1909*	*Built-up and water in 1900*	*Province fixed effects*
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS
Employment density within commuting distance *(log)*	**-0.3545*****	**-0.1525*****	**-0.2059*****	**-0.1560*****
	(0.0516)	**(0.0334)**	**(0.0348)**	**(0.0300)**
Worker and job characteristics	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes
Land use in 1900	No	No	Yes	No
Industry fixed effects	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	472,947	467,401	472,947	472,947
Kleibergen-Paap *F*-statistics	46.43	46.51	74.40	55.39

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments in columns (3) and (4). We use population density in 1900 and population density in 1900 squared as instruments in column (1). In column (2) we use the share of locally born people as instrument. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

Table C.7
Sensitivity analysis: wages.

Dependent variable: log of yearly wage				
	Population in 1900	*Share locals in 1909*	*Built-up and water in 1900*	*Province fixed effects*
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS
Employment density within commuting distance *(log)*	**0.1288*****	**0.1013*****	**0.0813*****	**0.0688*****
	(0.0263)	**(0.0188)**	**(0.0201)**	**(0.0178)**
Routine task intensity index	-0.0558***	-0.0558***	-0.0559***	-0.0559***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
Worker and job characteristics	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes
Land use in 1900	No	No	Yes	No
Industry fixed effects	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	472,946	467,400	472,946	472,946
Kleibergen-Paap *F*-statistics	46.24	46.29	74.26	55.31

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments in columns (3) and (4). We use population density in 1900 and population density in 1900 squared as instruments in column (1). In column (2) we use the share of locally born people as instrument. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

Table C.8
Overqualification and employment density.

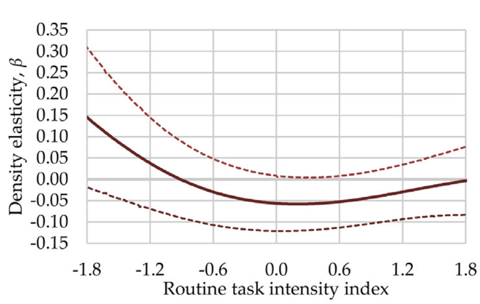
	Dependent variable: overqualified			*Dependent variable: overqualified, adjusted measure*		
	Baseline OLS specification	*Instrument for accessibility*	*+ 1909 skill composition*	*Baseline OLS specification*	*Instrument for accessibility*	*+ 1909 skill composition*
	(1) OLS	(2) 2SLS	(3) 2SLS	(4) OLS	(5) 2SLS	(6) 2SLS
Employment density within commuting distance *(log)*	-0.0173***	**-0.0326*****	**-0.0185**	-0.0180***	**-0.0395*****	**-0.0258**
	(0.0028)	**(0.0124)**	**(0.0173)**	(0.0028)	**(0.0121)**	**(0.0167)**
Routine task intensity index	0.1546***	0.1545***	0.1545***	0.2107***	0.2106***	0.2106***
	(0.0011)	(0.0011)	(0.0011)	(0.0012)	(0.0012)	(0.0012)
Share employment in 1909 in low-skilled occupations			-0.1289**			-0.0461
			(0.0586)			(0.0592)
Share employment in 1909 in medium-skilled occupations			-0.0679**			-0.0691**
			(0.0329)			(0.0350)
Share employment in 1909 in high-skilled occupations			-0.2266			-0.1996
			(0.1386)			(0.1311)
Worker and job characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	247,849	247,849	247,849	221,081	221,081	221,081
R^2	0.2080			0.3060		
Kleibergen-Paap *F*-statistics		109.2	66.03		106.2	64.06

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

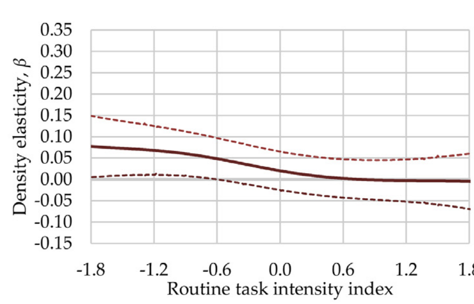
Table C.9
Sensitivity analysis: overqualification.

Dependent variable: worker is overqualified, adjusted measure				
	Population in 1900	*Share locals in 1909*	*Built-up and water in 1900*	*Province fixed effects*
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS
Employment density within commuting distance *(log)*	**-0.0762*****	**-0.0798*****	**-0.0492*****	**-0.0443*****
	(0.0162)	**(0.0132)**	**(0.0122)**	**(0.0108)**
Routine task intensity index	0.1814***	0.1816***	0.1815***	0.1815***
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Worker and job characteristics	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes
Land use in 1900	No	No	Yes	No
Industry fixed effects	Yes	Yes	Yes	Yes
Travel-to-work area fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	445,501	440,295	445,501	445,501
Kleibergen-Paap *F*-statistics	47.13	46.83	74.77	56.16

Notes: **Bold** indicates instrumented. We use employment density in 1909 and employment density in 1909 squared as instruments in columns (3) and (4). We use population density in 1900 and population density in 1900 squared as instruments in column (1). In column (2) we use the share of locally born people as instrument. Standard errors are clustered at the neighbourhood level. *** $p < .01$, ** $p < .05$, * $p < .10$.

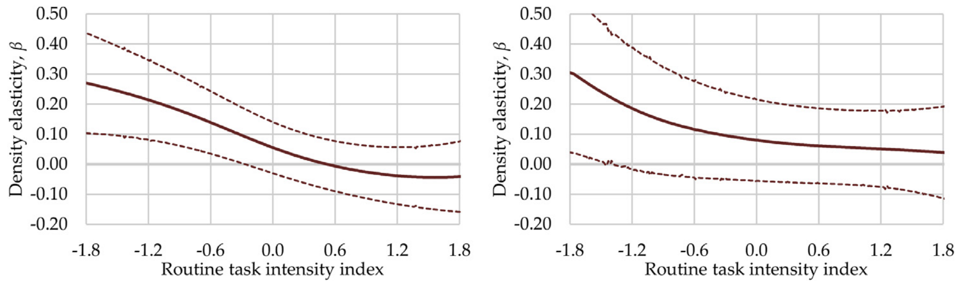


(A) Low-skilled



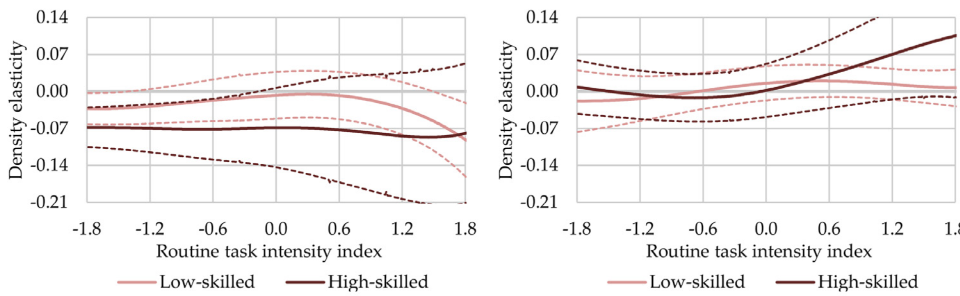
(B) Medium-skilled

Fig. C.1. Agglomeration elasticities for low and medium levels of education.
Notes: The regressions are based on a control function approach, where the first-stage error is inserted as a control variable in the second stage. We control for the share of skills in 1909, worker, job and location characteristics, as well as occupation 2-digit, industry, travel-to-work area and year fixed effects. In Panel A we further include occupation 2-digit fixed effects. The dashed lines constitute 95% confidence intervals, based on 250 cluster-bootstrapped replications.



(A) Firms with a HQ in the Netherlands (B) Firms with a HQ outside the Netherlands

Fig. C.2. RTI, density and wages by HQ location.
Notes: The regressions are based on a control function approach, where the first-stage error is inserted as a control variable in the second stage. We control for the share of skills in 1909, worker, job and location characteristics, as well as occupation 2-digit, industry, travel-to-work area and year fixed effects. The dashed lines constitute 95% confidence intervals, based on 250 cluster-bootstrapped replications.



(A) Vertical mismatch: overqualification (B) Horizontal mismatch

Fig. C.3. RTI, density and mismatch – high and low-skilled.
Notes: The regressions are based on a control function approach, where the first-stage error is inserted as a control variable in the second stage. We control for the share of skills in 1909, worker, job and location characteristics, as well as occupation 2-digit, industry, travel-to-work area and year fixed effects. The dashed lines constitute 95% confidence intervals, based on 250 cluster-bootstrapped replications.

vations towards around 11% of the total sample. Table C.3 reports the estimations.

Columns (1) and (2) present the OLS and IV estimates for the RTI effect of employment density, respectively. In column (2), we find very similar point estimates to the IV regressions in column (5) of Table 2, but larger standard errors due to much fewer observations. Column (3) predicts the agglomeration elasticity with respect to wages and we report a statistically significant relationship between employment density and wages that is of a similar magnitude to column (4) of Table 3. For the IV regressions, the baseline prediction remains remarkably similar to that of the IV estimate in Table 2, yet we lose significance due to larger confidence intervals.

C5. Alternative classifications

Some of the preceding literature that aims to measure whether the returns to tasks are higher in cities use alternative classifications of task intensity, rather than collapsing task intensity into one index, as we do in the current paper.

To reconcile with this literature (see *e.g.* Grujovic, 2018), we consider three different indices of abstract, routine and manual intensive occupations. Please recall that the correlation coefficients between the analytic task intensive occupation scores and the RTI is −0.89, while it is 0.3 for the manual occupation score; and 0.70 for the routine task intensive occupation score. We interact these scores with employment density (and the other controls) to investigate how these different dimensions of tasks are related to agglomeration economies for the preferred specification.

In columns (1)-(3) we add the three different dimensions one-by-one. We first find the expected result that analytic jobs benefit more from agglomeration economies. For a standard deviation increase in the analytic occupation score, the employment density is 0.0442 (54%) higher. In column (2) we find that agglomeration economies are slightly weaker for workers employed in manual task-intensive jobs. The elasticity is 0.0143 (16%) lower. Given the high correlation of the RTI with the routine occupation score, we also find a strong negative effect of the interaction of employment density and routine occupation score. The density elasticity decreases by −0.0363 (30%) for a standard deviation increase in the routine occupation score.

In column (4) we include all four occupation scores. For a standard deviation increase in analytic tasks, the elasticity with respect to employment density is 0.035 (46%) higher. For a standard deviation increase in manual and routine tasks, the employment density elasticity is 0.020 (27%) and 0.023 (30%) lower, respectively. Given these results, and the above-mentioned correlations of the RTI with these categories, this supports our approach to reduce dimensionality and collapse these three indices into one index measuring task intensity.

C6. Multinational corporations

One may be concerned that part of the relationship between the RTI, wages and density as depicted in Fig. 6 is caused by the sorting of firms. More specifically, multinational firms in which jobs are more analytic task-intensive may be disproportionally clustered in denser areas.

To investigate this further we gather data from the *Community Innovation Survey*. For a subsample of firms, we know the location of the *headquarters* (HQ) of the firm. We then distinguish between firms with a Dutch HQ and firms with an HQ outside the Netherlands.

We report results in Table C.5. First, we replicate the baseline estimate for wages reported in column (6), Table 3. We lose about 80% of the observations but still find a comparable, albeit slightly lower, density elasticity (0.0766 *vs.* 0.0941). In column (2) we focus on firms that are

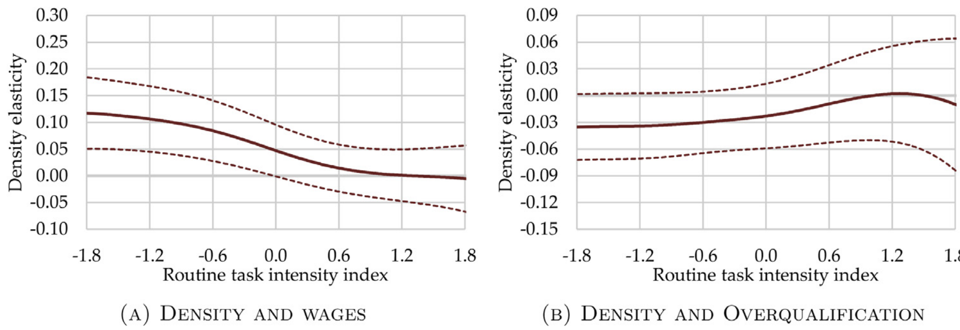


Fig. C.4. Agglomeration elasticities for different RTI levels ≤2011. *Notes*: The regressions are based on a control function approach, where the first-stage error is inserted as a control variable in the second stage. We control for the share of skills in 1909, worker, job and location characteristics, as well as occupation 2-digit, industry, travel-to-work area and year fixed effects. In Panel A we further include occupation 2-digit fixed effects. The dashed lines constitute 95% confidence intervals, based on 250 cluster-bootstrapped replications.

headquartered in the Netherlands. We only find a slightly lower employment density. Given the standard errors, this estimate is not statistically significantly lower.

Column (3) only includes firms with an HQ outside the Netherlands. We find a slightly higher employment density elasticity, but given the large standard errors, the estimate is not significantly different from the estimates reported in columns (1) and (2). Hence, we think it is safe to conclude that the sorting of multinationals is unlikely to explain the result that agglomeration economies are only relevant for workers in non-routine task-intensive jobs.

Moreover, we show in Fig. C.2 that the finding that the wage-density premium only applies to analytic workers (*i.e.* those with a low RTI), also holds for firms with a Dutch HQ; as well as for firms with an HQ outside the Netherlands.

C7. Other sensitivity checks

Here we report the results of a wide array of additional sensitivity checks.

Table C.6 reports the results where the RTI is the dependent variable. We consider the specification reported in column (5) in Table 2 as the preferred estimate. Column (1) in Table C.6 uses alternative instruments, *i.e.* the total population in 1900 that is accessible within commuting distance. Because the data on population are more fine-grained, we expect that measurement error will be reduced. We see that it matters only slightly as the coefficient is somewhat higher. We further use an alternative instrument – the share of locally born people (*i.e.* within the same municipality) in 1909. If the (lack of) mobility of households is correlated over time, the share of locally born people should be correlated positively with current commuting times as immobile households have to commute on average longer to their jobs. The coefficient is then somewhat lower than the baseline estimate. Hence, our baseline estimate of −0.207 seems to be more or less in between the specifications relying on alternative instruments. In column (4) we address the issue that employment density in 1909 may be correlated with land use in 1900. That is, low-density places in 1909 may still have better access to open space. However, we observe that controlling for the share of built-up land and water bodies in 1900 in the neighbourhood hardly changes the results. In column (5) we test whether our results are robust to the inclusion of different geographical fixed effects. That is, we include province fixed effects instead of travel-to-work-area fixed effects. This leads to a slightly lower coefficient for employment density. All in all, these regressions confirm the negative relationship between employment density and the RTI: more routine jobs are less concentrated in cities.

In Table C.7 we repeat the same set of specifications but this time for wages. Recall that our preferred estimate is 0.0941 (see column (6), Table 3). What we observe is that alternative sets of instruments generally lead to similar elasticities for employment density.

C8. Mismatch – alternative measures and sensitivity checks

In this Appendix section, we consider alternative measures and sensitivity checks for the effects of employment density on overqualification. First, we show the relationship between employment density and overqualification for different RTI levels by repeating the baseline specification, yet now distinguishing between low and high-skilled workers. Second, to circumvent measurement error in the overqualification measure, we analyse whether our results are robust to using the ‘original’ overqualification measure where we observe the required education variable for each year between 2006 to 2011. Third, by including all years, but only keeping the occupations for which the required education for more than 90% of the cases applies to only one skill category, we show the robustness with respect to the adjusted overqualification measure. That is, if the required degree is known, there is very little measurement error in the overqualification measure. Fourth, we undertake a couple of sensitivity checks for the baseline results regarding the impact of employment density on overqualification.

The first step is to show in Fig. C.3 that also within skill groups we observe that overqualification is mostly reduced for workers in non-routine task intensive occupations. For horizontal mismatch, we consistently do not find an effect of urban density.

Second, our overqualification measure is subject to measurement error because we calculate the required education by skill level for each occupation in 2006 − 2016 based on data between 2006 and 2011. Hence, recall that overqualification is not a dummy, but a probability. Columns (1)-(3) in Table C.8 show the results for the original overqualification measure. We find largely similar effects, although the effects for the IV-specifications (columns (2) and (3)) seem to be slightly smaller as compared to the baseline results reported in Table 4. Because the standard errors are also somewhat higher, the estimate in column (3) is statistically insignificant.

Using the original overqualification measures from 2006 to 2011, we further replicate the semi-parametric regressions. In Panel A of Fig. 6 we first focus on wages between 2006–2011 to corroborate our findings reported in Fig. 6. We find that the agglomeration elasticity with respect to wages is somewhat lower than the baseline estimate, but the qualitative conclusions remain unchanged: only workers in non-routine task intensive jobs benefit from agglomeration economies. We observe a similar pattern in Panel B of Fig. C.4, where overqualification is only lower for lower levels of the RTI. Hence, the results essentially replicate the results for overqualification reported in Fig. 9a. However, here we see that the estimated effects for non-routine jobs are only statistically significant at the 10% level.

Third, we consider an alternative approach by using the adjusted overqualification measure but only keeping the occupations for which the required degree falls in 90% of the cases in one skill category. This implies that we drop about 50% of the observations.

We find in columns (4)-(6) in Table C.8 that this leads to very similar outcomes as when using the full dataset, but the IV estimates in columns

(5) and (6) are again somewhat smaller. However, the estimates are not significantly different from the baseline results reported in Table 4.

Fourth, we undertake a couple of sensitivity checks for the baseline results in Table C.9. Recall the preferred baseline estimate reported in column (6) of Table 4 is −0.0552. The negative impact of density on overqualification is confirmed by all specifications.[34]

References

Acemoglu, D., Autor, D., 2011. Chapter 12 - skills, tasks and technologies: implications for employment and earnings 4, 1043–1171.
Acemoglu, D., Restrepo, P., 2018. The race between man and machine: implications of technology for growth, factor shares and employment. American Economic Review 108 (6), 1488–1542.
Autor, D., Acemoglu, D., 2011. Skills, Tasks and Technologies: Implications for Emploument and Earnings. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Economics, vol. 4.
Adermon, A., Gustavsson, M., 2015. Job polarization and task-biased technological change: evidence from sweden, 19752005. Scandinavian Journal of Economics 117 (3), 878–917.
Ahlfeldt, G., Redding, S., Sturm, D., Wolf, N., 2016. The economics of density: evidence from the Berlin Wall. Econometrica 83 (6), 1217–2189.
Autor, D.H., 2019. Richard T. Ely lecture: work of the past, work of the future. AEA Papers and Proceedings 109, 1–32.
Autor, D., Dorn, D., 2013. The growth of low-Skill service jobs and the polarization of the U.S. labor market. American Economic Review 103 (5), 1533–1597.
Autor, D.H., Dorn, D., Hanson, G.H., 2015. Untangling trade and technology: evidence from local labour markets. The Economic Journal 125 (May), 621–646. doi:10.1111/ecoj.12245.
Autor, D.H., Levy, F., Murnane, R.J., 2003. The skill content of recent technological change: an empirical exploration*. Q J Econ 118 (4), 1279–1333.
Bacolod, M., Blum, B.S., Strange, W.C., 2009. Skills in the city. J Urban Econ 65 (2), 136–153.
Baum-Snow, N., Pavan, R., 2013. Inequality and city size. Review of Economics and Statistics 95 (5), 1535–1548.
Beaudry, P., Green, D.A., Sand, B.M., 2016. The great reversal in the demand for skill and cognitive tasks. J Labor Econ 34 (S1), S199–S247.
Behrens, K., Duranton, G., Robert-Nicoud, F., 2014. Productive cities: sorting, selection, and agglomeration. Journal of Political Economy 122 (3), 507–553.
Behrens, K., Robert-Nicoud, F., 2015. Agglomeration theory with heterogeneous agents, 5. Elsevier B.V..
Berlingieri, F., 2019. Local labor market size and qualification mismatch. Journal of Economic Geography 19, 1261–1286.
Black, D., Kolesnikova, N., Taylor, L., 2009. Earnings functions when wages and prices vary by location. J Labor Econ 27 (1), 21–47.
Blau, F.D., Kahn, L.M., 2003. Understanding international differences in the gender pay gap. Journal of Labour Economics 21 (1), 106–144.
Böhm, M.J., 2020. The price of polarization: estimating task prices under routinebiased technical change. Quant Econom 11 (2), 761–799.
Boualam, B., 2014. Getting a first job: quality of the labor matching in French cities. Research Papers by the Institute of Economics and Econometrics, Geneva School of Economics and Management, University of Geneva.
Briant, A., Combes, P.P., Lafourcade, M., 2010. Dots to boxes: do the size and shape of spatial units jeopardize economic geography estimations? J Urban Econ 67 (3), 287–302.
Büchel, F., van Ham, M., 2003. Overeducation, regional labor markets, and spatial flexibility. J Urban Econ 53 (3), 482–493. doi:10.1016/S0094-1190(03)00008-1.
Cameron, A., Trivedi, P., 2005. Microeconometrics: Methods and Application. Cambridge University Press, New York.
Ciccone, A., 2002. Agglomeration effects in Europe. Eur Econ Rev 46 (2), 213–227.
Ciccone, A., Hall, R., 1996. Productivity and the density of economic activity. Am Econ Rev 86 (1), 54–70.
Combes, P., Duranton, G., Gobillon, L., 2008. Spatial wage disparities: sorting matters!. J Urban Econ 63 (2), 723–742.
Combes, P.P., Duranton, G., Gobillon, L., Puga, D., Roux, S., 2012. The productivity aadvantages of large cities: distinguishing agglomeration from firm selection. Econometrica 80 (6), 2543–2594.
Davis, D., Dingel, J., 2019. A spatial knowledge economy. American Economic Review 109 (1), 153–170.
Davis, D.R., Dingel, J.I., 2020. The comparative advantage of cities. J Int Econ 103291.
Davis, D.R., Mengus, E., Michalski, T.K., 2020. Labor market polarization and the great divergence: theory and evidence. NBER Working paper #26955 1–57.
D'Costa, S., Overman, H.G., 2014. The urban wage growth premium: sorting or learning? Reg Sci Urban Econ 48, 168–179.

[34] One may argue whether the share of locally born people (see column (2)) may have a direct effect on overqualification. Immobile households are often less skilled and have fewer job market opportunities, possibly leading to more overqualification. If there is a correlation between share of locally born people in 1909 and the share of immobile workers today, the effect of employment density is overstated. Because this instrument may be invalid in the context of overqualification, we do not place put too much trust in the estimate in column (2).

De la Roca, J., Puga, D., 2017. Learning by working in big cities. Review of Economic Studies 84, 106–142.
Duncan, G.J., Hoffman, S.D., 1981. The incidence and wage effects of overeducation. Econ Educ Rev 1 (1), 75–86.
Duranton, G., Kerr, W.R., 2018. The logic of agglomeration.
Duranton, G., Puga, D., 2001. Nursery cities: urban diversity, process innovation, and the life cycle of products. American Economic Review 91 (5), 1454–1477.
Duranton, G., Puga, D., 2004. Micro-foundations of Urban Agglomeration Economies. In: Handbook of Regional and Urban Economics, 4. Elsevier, pp. 2063–2117. doi:10.1016/S1574-0080(04)80005-1.
Ehrl, P., Monteiro Monasterio, L., 2016. Historical trades, skills and agglomeration economies. MPRA 69829.
Fallick, B., Fleischman, C., Rebitzer, J., 2006. Job-Hopping in Silicon Valley: some evidence concerning the microfoundations of a high- Technology cluster. Review of Economics and Statistics 88 (3), 472–481.
Gaigné, C., Koster, H., Moizeau, F., Thisse, J., 2017. Who lives where in the city? amenities, commuting and income sorting. CEPR Discussion Paper 11958.
Glaeser, E., Kallal, H., Scheinkman, J., Shleifer, A., 1992. Growth in cities. Journal of Political Economy 100 (6), 1126.
Glaeser, E., Maré, D., 2001. Cities and skills. J Labor Econ 19 (2), 316–342. doi:10.1086/319563.
Goos, M., Manning, A., 2007. Lousy and lovely jobs: the rising polarization of work in Britain. Rev Econ Stat 89 (1), 118–133.
Goos, M., Manning, A., Salomon, A., 2014. Explaining job polarization: Routine-biased technological change and offshoring. American Economic Review 104 (8), 2509–2526.
Groot, S., De Groot, H., Smit, M., 2014. Regional wage differences in the Netherlands: micro evidence on agglomeration externalities. J Reg Sci 54 (3), 503–524.
Grujovic, A., 2018. Tasks, cities and urban wage premia. CEPREMAP Wp No.1807.
Hartog, J., 2000. Over-education and earnings: where are we, where should we go? Econ Educ Rev 19, 131–147. doi:10.1016/S0272-7757(99)00050-3.
Haworth, C.M.A., Wright, M.J., Martin, N.W., Martin, N.G., Boomsma, D.I., Bartels, M., Posthuma, D., Davis, O.S.P., Brant, A.M., Corley, R.P., Hewitt, J.K., Iacono, W.G., McGue, M., Thompson, L.A., Hart, S.A., Petrill, S.A., Lubinski, D., Plomin, R., 2009. A twin study of the genetics of high cognitive ability selected from 11,000 twin pairs in six studies from four countries. Behav. Genet. 39 (4), 359–370.
Henderson, J., Kuncoro, A., Turner, M., 1995. Industrial development in cities. Journal of Political Economy 103 (5), 1067–1090. doi:10.1086/262013.
Jauhiainen, S., 2011. Overeducation in the finnish regional labour markets. Papers in Regional Science 90, 573–588.
Kampelmann, S., Rycx, F., 2012. The impact of educational mismatch on firm productivity: evidence from linked panel data. Econ Educ Rev 31, 918–931.
Knol, W., Kramer, H., Gijsbertse, H., 2004. Historisch grondgebruik Nederland: Een landelijke reconstructie van het grondgebruik rond 1900. Alterra, Wageningen.
Koopmans, C., Rietveld, P., Huijg, A., 2012. An accessibility approach to railways and municipal population growth, 1840–1930. J Transp Geogr 25, 98–104.
Leuven, E., Oosterbeek, H., 2011. Handbook of the Economics of Education. In: Handbook of the Economics of Education. In: Handbook of the Economics of Education, 4. Elsevier, pp. 283–326.
Li, Q., Racine, J.S., 2007. Nonparametric econometrics: Theory and practice. Princeton University Press, Princeton.
Lin, J., 2011. Technological adaptation, cities, and new work. Rev Econ Stat 93 (2), 554–574.
Melo, P., Graham, D., Noland, R., 2009. A meta-analysis of estimates of urban agglomeration economies. Reg Sci Urban Econ 39 (3), 332–342.
Mendes de Oliveira, M., Santos, M., Kiker, B., 2000. The role of human capital and technological change in overeducation. Econ Educ Rev 19, 199–206.
Michaels, G., Rauch, F., Redding, S., 2016. Tasks and technology in the United States 1880–2000. NBER working paper 18715.
Oesch, D., 2013. Occupational change in Europe. Oxford University Press.
Ozgen, C., 2020. Education surplus, skills mismatch and technological change. University of Birmingham, working paper.
Peters, J.C., 2021. Dynamic agglomeration economies and learning by working in specialised regions. Journal of Economic Geography 20 (3), 629–651.
Roback, J., 1982. Wages, rents, and the quality of life. Journal of Political Economy 90 (6), 1257–1278.
Robinson, P.M., 1988. Root-N-consistent semiparametric regression. Econometrica 56 (4), 931–954.
Rosenthal, S.S., Strange, W.C., 2008. Agglomeration and hours worked. Review of Economics and Statistics 90 (1), 105–118.
Sanchez-Sanchez, N., McGuinness, S., 2015. Decomposing the impacts of overeducation and overskilling on earnings and job satisfaction: an analysis using reflex data. Educ Econ 23 (4), 419–432.
Silverman, B., 1986. Density estimation for statistics and data analysis. Chapman and Hall, New York.
Van Oort, F.G., Ponds, R., Van Vliet, J., Van Amsterdam, H., Declerck, S., Knoben, J., Pellenbarg, P., Weltevreden, J., 2007. Verhuizingen van bedrijven en groei van werkgelegenheid. NAi Uitgevers, Rotterdam.
Van der Velde, L., 2017. Within occupation wage dispersion and the task content of jobs. Oxf Bull Econ Stat 82 (5), 1161–1197.
Violante, G.L., 2002. Technological acceleration, skill transferability, and the rise in residual inequality. Quarterly Journal of Economics 117 (1), 297–338.
Wheeler, C., 2001. Search, sorting, and urban agglomeration. J Labor Econ 19 (4), 879–899.
Wheeler, C.H., 2006. Cities and the growth of wages among young workers: evidence from the NLSY. J Urban Econ 60 (2), 162–184. doi:10.1016/j.jue.2006.02.004.