INSTITUTE OF ECONOMICS
Centre for Economic and Regional Studies

Közgazdaság- és Regionális Tudományi Kutatóközpont
KÖZGAZDASÁG-TUDOMÁNYI INTÉZET

KRTK-KTI WORKING PAPERS | KRTK-KTI MŰHELYTANULMÁNYOK

Return to skills and urban size: Evidence from the skill requirements of Hungarian firms

LÁSZLÓ CZALLER – ZOLTÁN HERMANN

KRTK-KTI WP – 2022/5

February 2022

https://kti.krtk.hu/wp-content/uploads/2022/02/CERSIEWP202205.pdf

INSTITUTE OF ECONOMICS, CENTRE FOR ECONOMIC AND REGIONAL STUDIES
EÖTVÖS LORÁND RESEARCH NETWORK (ELKH)
BUDAPEST, 2022

ABSTRACT

While most empirical studies document that cognitive and social skills are strong predictors of individual earnings, their impact is not homogenous in space. We argue that dense urban settings utilize cognitive and social skills more intensively than rural areas, therefore the labour market return to these skills is higher in cities. Using data from a representative survey recording the skills requirements of Hungarian firms, we show that social skills are rewarded more in dense urban areas. Surprisingly, this pattern is not observed for cognitive skills. We use instrumental variables strategy to correct for measurement errors in skills, and to deal with the endogeneity of agglomeration. Our results are robust to alternative agglomeration measures and a large set of controls, however, returns to skills vary considerably across worker groups and industries.

JEL codes: J24, J31, R12

Keywords: agglomeration, cognitive and social skills, wages, urban labour markets

László Czaller

Agglomeration and Social Networks Lendület Research Group,

Budapest, H-1097, Hungary

and

Institute of Economics, Centre for Economic and Regional Studies,

Budapest, H-1097, Hungary

e-mail: czaller.laszlo@krtk.hu

Zoltán Hermann

Institute of Economics, Centre for Economic and Regional Studies,

Budapest, H-1097, Hungary

and

Institute of Economics, Corvinus University,

Budapest, HU-1093, Hungary

e-mail: Zoltan.hermann@krtk.hu

A készségek kereseti hozama és a városok mérete: a hazai vállalatok készségkövetelményein alapuló elemzés eredményei

CZALLER LÁSZLÓ – HERMANN ZOLTÁN

ÖSSZEFOGLALÓ

Számos tanulmány kimutatta, hogy a kognitív és szociális készségeknek erős hatása van a munkavállalók keresetére, ez a hatás azonban térben nem tekinthető homogénnek. Tanulmányunkban amellett érvelünk, hogy nagyvárosi környezetben a kognitív és szociális készségeket intenzívebben hasznosítják, mint a kisvárosi és falusi térségekben, ezért a nagyvárosokban ezeknek a készségeknek magasabb a kereseti hozama. A hazai cégek készségigényeit rögzítő reprezentatív felmérés adatainak elemzése azt mutatja, hogy a szociális készségek jobban megtérülnek a munkaerőpiacon, ha sűrűn lakott városi térségekben használják őket. Meglepő módon a kognitív készségek esetében ez az összefüggés nem megfigyelhető. A készségek esetleges mérési hibáinak kiküszöbölésére és az agglomerációs változó endogenitásának kezelésére instrumentális változókat használunk. Eredményeink robusztusak a becslési eljárásra, az alternatív agglomerációs mérőszámok megválasztására és a kontrollváltozók széles körére. A készségek kereseti hozama azonban jelentős mértékben eltér a munkavállalói csoportok és iparágak között.

JEL: J24, J31, R12

Kulcsszavak: agglomeráció, kognitív és szociális készségek, bérek, városi munkaerőpiacok

INTRODUCTION

Over the last few decades numerous studies established that cognitive skills and other noncognitive abilities are strong predictors of individual earnings (Borghans, ter Weel, & Weinberg, 2008; Heckman et al., 2006b; Lindqvist & Vestman 2011; Heckman & Kautz, 2012; Deming, 2017) while another line of inquiry in urban economics focused on the general regularity of higher urban wages commonly referred to as the urban wage premium (Glaeser & Maré, 2001; Wheeler, 2006; Combes et al., 2008; Baum-Snow & Pavan, 2012). Much less is known, however, about the extent to which work location contributes to the economic return to different skills. Bacolod et al. (2009) were the first to examine how individual skills are rewarded in cities of different sizes. Relying on the assumption of positive assortative matching between workers and employers, they measured skills by using data on job skill requirements and reported that returns to cognitive and people skills are higher in large cities. Similar conclusions have been drawn by Andersson et al. (2014) for Sweden and Choi (2020) for South Korea.

A common feature of these studies is that they use occupation-level data on job skill requirements to approximate individual skills. One considerable limitation of this approach is that skill requirements are assumed to be homogeneous within occupations, or at least independent of the degree of agglomeration. Recent evidence, however, suggests that the variation of skill contents or stated requirements within occupations is substantial (Autor & Handel, 2013; Deming & Kahn, 2018), therefore, interpreting higher returns to skills in large urban areas as productivity gains from agglomeration is problematic. When skills are measured at the occupational level, the Mincerian approach confounds productivity gains from agglomeration economies with spatial differences in the occupational composition of the labour force.

This paper re-examines the interaction between agglomeration and returns to skills using firm-level data on job skill requirements. We make use of a unique survey recording the skill requirements for more than one thousand Hungarian firms. By matching this dataset to wage survey data, we can measure individual skills directly by employers' requirements and estimate wage regressions where returns to skills are allowed to vary by the degree of agglomeration. Of the many skill types considered in the literature, we focus on cognitive and social skills because presumably these skills contribute most to the microeconomic mechanisms underlying agglomeration externalities. Although previous studies also considered skills that are more related to routine job tasks (e.g. motor or physical skills), these skills can hardly be linked to agglomeration economies (Bacolod et al., 2009; Florida et al., 2012; Barbour & Markusen, 2007; Scott, 2009). Hence, motor and physical skills are omitted from this analysis.

Our contribution to the literature is twofold. First, we provide evidence on the association between agglomeration and skill returns in the context of a Central European small open economy with a low level of agglomeration overall. Second, we approximate individual skills by firm-level job requirements rather than occupation-level measures. Measuring skills using firm-by-occupation level data has several advantages over occupational level alternatives. It allows us to control for both occupations and firm characteristics in wage regressions and thus characterize the variation in returns to skills across locations more accurately. At the same time, the assumption of assortative labour market matching is more likely to hold at the level of the firm. The approach of evaluating worker skills using the requirements of employers builds heavily on the assumption that workers are assigned to firms for which they are well-suited (Abraham & Spletzer, 2009; Ingram & Neuman, 2006). Although skill requirements do not reflect the actual capabilities of workers precisely, there are reasons to believe that employer-employee mismatches are rather exceptions than the regular course of things. Several models on search and recruitment predict assortative matching in the labour market even when search frictions are present (Albrecht & Vroman, 2002; Shimer & Smith, 2000; Chade et al., 2017).

Our main finding is that social skills are rewarded more in agglomerated areas even after controlling for firm-level characteristics and occupations. Returns to social skills increasing with agglomeration are consistent with earlier findings of the literature, however, we find no evidence for agglomeration affecting returns to cognitive skills. We believe that these patterns reflect the effects of extensive changes in the demand for and supply of such skills in Hungary. The massive expansion of enrolment in secondary and tertiary education in the past decades resulted in an increased supply of cognitive skills, especially in cities (Szakálné Kanó et al., 2017). This has recently led to a decrease in the wage returns to higher education as well (Varga, 2020). We suppose that this increased investment in human capital through formal education has driven down the urban premium for cognitive skills. At the same time, the supply of social skills is probably less affected by rational decisions to invest in human capital and therefore has not increased together with cognitive skills.

Separate estimates for different worker subgroups indicate that the urban premium for social skills is higher for educated workers. Besides, in line with Bacolod (2016), social skills seem to be more valuable for men than they are for women, though women in our sample have, on average, better social skills. When we restrict our sample by sectors we find evidence of heterogeneous returns. For the manufacturing sector no sign of significant skill returns can be found, while for service sector workers we find higher returns to social skills in densely populated areas.

The remainder of the paper is structured as follows. The next section presents our empirical model and discusses some of the potential issues regarding its estimation. Section 3 outlines data while Section 4 presents our results. Section 5 concludes.

EMPIRICAL MODEL AND ESTIMATION ISSUES

Our empirical model is simply an analog to a Mincer equation except that it allows skill returns to vary by the degree of agglomeration as in Bacolod *et al.* (2009). Considering only cognitive (c) and social (s) skills the wage equation can be expressed as follows:

$$\log \mathrm{w}_{ir} = \vartheta + \alpha \log a_r + \sum_j \gamma^j \, \mathrm{z}_i^j + \log a_r \sum_j \delta^j \, \mathrm{z}_i^j + \mathbf{X}_i' \boldsymbol{\beta} + \epsilon_{ir} \tag{1}$$

where w_{ir} represents the wage of individual i working in location r, a_r denotes the degree of agglomeration, z_i^j denotes the level of skill j for individual i and X_i is a vector covariates.

Conceptually, the return to skill j for a person working in location r is $\gamma^j + \delta^j \log a_r$, where the focus of our attention is δ^j, a coefficient measuring the association between agglomeration and the labour market return to skill j. For both skill types, $j \in \{c, s\}$, the expected sign of δ^j is positive.

A caution to this approach, however, is that just like the majority of previous studies it does not provide a convincing analysis of the causal effect of skills on earnings (see e.g., Heckman & Kautz, 2012; Hanushek et al., 2015). One of the most extensively discussed concerns comes from unobserved variables such as personality traits and other inherent abilities that might simultaneously affect skill formation and wages. Recognizing these issues, Bacolod *et al.* (2009) and Andersson *et al.* (2014) estimated a series of fixed-effects models to eliminate time-invariant components of ability. However, even in cross-sectional settings where the inclusion of fixed effects is not feasible, there are good reasons to believe that the extent of ability bias can be minimized by including employer characteristics and occupations in X_i. As shown by Heckman et al. (2006b) non-cognitive abilities measured by indices on personality traits, perceived self-worth and locus of control prevail through various channels including education and occupational choice. If workers with better unobserved abilities find positions at productive firms and self-select into prestigious occupations – as the literature suggests –, controlling for these factors will capture the unobserved ability component and thus partly correct for the omitted variable bias. Since it is likely that some of the unobserved abilities considered above are also correlated with the degree of agglomeration, controlling for employers and occupations may also help mitigate the problem of spatial sorting.

Although including worker fixed effects in panel settings became a standard practice lately, in the present context it does not necessarily bring us closer to the identification of causal effects. On the one hand, fixed effects strategies are particularly effective when skill measures act as proxies for other time-invariant abilities, however, if a substantial part of cognitive and social skills comes from (or builds upon) innate abilities and personality traits (Cunha & Heckman, 2008; Hanushek et al., 2015), one should not include fixed effects in a regression that aims to estimate the total effect of skills.[1] Since the exact causal relations between observed skills and unobserved attributes are unknown in most studies that rely on observational data, it is quite difficult to assess whether fixed effects are useful in identifying skill returns. In contrast, using micro data that include a wide range of control variables allows us to infer on the economic conduits through which different skills affect earnings.

Another issue is that measurement error in skills could give rise to standard attenuation bias implying that the estimated coefficients will be biased toward zero. Moreover, if skills are measured using firm-level information on skill demands, it is possible that the actual skills of some workers will differ from their employer's requirements. If unmeasured skill components are correlated with either of our key variables, estimates on returns to skills will be biased.

Apart from these issues there is another limitation to be mentioned. As is generally the case in the literature, our approach does not provide a full characterization of the rate of returns to skills, not only because of the general issues discussed by Heckman et al. (2006a) but also because it does not account for the costs of living and a series of other factors that might affect the actual rate of return (e.g. amenities or the costs of achieving a certain level of skills). Ideally, a structural model on the location choice of workers with different combinations of skills would be required to properly allow for the whole spectrum of equilibrium effects that might affect the actual rate of return to skills. Nevertheless, even in the absence of any structural

[1] Hanushek et al. (2015) argues that the standard ability bias that the empirical literature on returns to schooling aims to address cannot be considered as a bias of returns to skills, because for estimates of skill returns it does not matter where worker skills stem from.

underpinning, wage regressions provide a strong indication of the relative importance of skills in local labour markets and help unravel the mechanisms behind urban wage premium.

DATA AND MEASUREMENT

We draw on multiple data sources. First, we extract data from a detailed survey carried out by the Institution of Economics, Hungarian Academy of Sciences (IE-HAS) in the fall of 2012. A sample of Hungarian firms were asked about the skill requirements of 10 occupations performed within the firm (the 5 most important occupations in terms of employment and another 5 randomly selected occupations),[2] using a standardized list of queries similar to the O*NET Skills Questionnaire. All occupations were graded on a set of skill descriptors, in terms of two separate dimensions: first, on a scale of 1 to 7 according to the 'level' of skill needed to perform the job; and second, on a scale of 1 to 5 corresponding the 'importance' of the skill descriptor to the occupation in terms of frequency of use. The survey accounted for 36 distinct skill descriptors ranging from simple skills such as reading comprehension and writing to more complex ones such as programming or resource management.[3] Scores assigned to the descriptors are provided by the human resource manager, or the executive manager of the firm. Throughout the survey a unique occupational classification consisting of 200 elements was used. Each of these elements are small groups of 4-digit ISCO-88 occupations. Sample firms were selected by stratified sampling, and then, the respondent sample was weighted to the known distribution of the sampling frame (firm size, industry, occupation and location). The final sample consists of 1029 firms and the total number of filled questionnaires is 8568 covering 194 occupations.

We matched this dataset to the 2010 and 2011 waves of the National Employment Office's annual wage survey using firm identifiers and 4-digit ISCO-88 codes. This survey includes the entire public sector, all firms with more than 20 employees and a 20% random sample of firms employing fewer than 20 workers. Firms with less than 50 employees provide data on all workers, while larger firms report only a 10% random sample of their employees. After matching the two datasets and removing observations with missing values we got a dataset containing 14.992 private sector employees. This dataset includes a wide range of individual

[2] In cases when the total number of jobs performed within the firm did not exceed 10, all of the available jobs got into the sample.
[3] The skill descriptors are the same as in the O*NET Skills Questionnaire, except the last one which refers to computer skills.

characteristics such as wages, sex, age, hours of work, educational attainment, occupation, and also detailed information on employers.

Similar to previous studies a subset of the available skill items is used to construct interpretable indices for skills. As a first step the product of the 'level' and the 'importance' scales are calculated for each item to increase variance as in Feser (2003). Cognitive skills are defined as the average of the following six items: reading comprehension, active listening, mathematics, critical thinking, active learning, learning strategies and monitoring. To capture social skills, the following six items are selected and averaged out: social perceptiveness, coordination, persuasion, negotiation, instructing and service orientation.[4] All skill items used in the analysis are further described in Table A1 in the online Appendix. Skill indices are standardized to have a mean of 0 and a standard deviation of 1.

Agglomeration is measured by the distance-weighted average of the number of employees working in the neighbourhood of location r as in Koster *et al.* (2014). We prefer this metric because it is capable of capturing the effects of interactions that reach beyond administrative boundaries. This is particularly important in some parts of Hungary because the delineation of administrative borders does not correspond to the actual extent of local labour markets. Formally, our agglomeration metric is defined as:

$$a_r = \int_s \psi_{rs} n_s \, \mathrm{d}s,$$

where n_s denotes the number of employees in location s, and ψ_{rs} is a spatial weight that gives more importance to closer locations and less to locations that are further away from r. We estimate ψ_{rs} using an Epanechnikov kernel:

$$\psi_{rs} = \left(1 - \left(\frac{d_{rs}}{d_T}\right)^2\right) \mathbf{1}(d_{rs} < d_T),$$

where d_{rs} is the Euclidean distance between locations r and s, d_T is a predefined distance cut-off, and 1(.) is an indicator function that takes value 1 if $d_{rs} < d_T$ condition holds true. Since the wage survey contains information on job sites at the settlement level, a location refers to a settlement except in the case of Budapest where data for all 23 districts are available. Aggregate data on n_s comes from the 2011 Census.

[4] We have experimented with different combinations of the skill items and also tried principal component analysis to create alternative social skill indices but due to the high pairwise correlations between skill items none of these alternative approaches changed any of the results reported below.

The distance cut-off (d_T) is determined by cross-validation.[5] This procedure finds the optimal value for d_T that minimizes the mean squared error (MSE) of the regression model. Cross-validation techniques are frequently used to find the optimal smoothing (bandwidth) parameter for non-parametric approaches and regression discontinuity estimators (see, e.g. Imbens & Lemieux, 2008). The cross-validation criterion can be defined as

$$CV(d_{Tk}) = \frac{1}{N}\sum_{k=1}^{N}\left(Y - \hat{Y}_k(d_{Tk})\right)^2,$$

with the corresponding cross-validation choice for the optimal cut-off:

$$d_T^* = \arg\min_k CV(d_T),$$

where Y is the actual outcome (wage), $\hat{Y}_k(d_{Tk})$ is the predicted outcome, and d_T^* is the optimal cut-off. Accordingly, distance cut-off parameters are determined for each model separately. Since performing this procedure is computationally demanding we assume that the optimal distance cut-off is an integer somewhere between 0 and 40 kilometres in each specification. As will be shown in the next section optimal d_T parameters are much lower than 40 kilometres and vary within a relatively narrow range even when alternative weighting functions are used. Obviously, bandwidth choices might affect both estimates and standard errors (Imbens & Lemieux, 2009), therefore it is important to check whether the key results are dependent on a particular bandwidth choice.

All the other variables used in the analysis come from the Hungarian wage survey (HWS henceforth). The dependent variable is the gross average monthly wage containing the basic wage and other remuneration benefits such as overtime pay, and commissions. Individual control variables include sex, work experience (and its square), educational attainment, occupation and a dummy for part-time work, while firm-level controls include dummies on firm size, ownership, and collective agreements. Work experience is measured as age minus years of schooling minus six and those who work less than 36 hours per week are considered as part-time workers. To control for education, we define four educational dummies such that they correspond to the following ISCED categories: (i) primary education or less (ISCED 0 and 1), (i) lower secondary education (ISCED 2), (iii) upper secondary education (ISCED 3), and (iv) tertiary education or more (ISCED 5 and 6). Firms are classified into five categories by the

[5] A similar procedure has been used by Koster (2013) to find the optimal bandwidth for models explaining commercial property rents.

number of employees and two dummies are introduced to control for ownership; the first one indicates whether at least 50% of the firm is owned by foreigners, and the second one takes value 1 if the majority of the firm is owned by the state or a local government. Some of the regression models include dummies for 2-digit NACE Rev.2. industries. Table A2 in the online Appendix presents some descriptive statistics on the variables used in the analysis.

[FIGURE 1 AROUND HERE]

Figure 1 plots mean skill intensities for different urban size categories (*x* axis). To construct these categories locations are partitioned into five groups using the quintiles of the agglomeration metric (d_T =18kms). Consistent with previous findings, the intensity of skills increases with the degree of agglomeration. Specifically, the mean intensity of cognitive and social skills in the top quintile is about 0.44 points larger than the observed mean in the bottom quintile and significantly larger than the mean of the whole sample (0.18 compared to 0.00). Interestingly, the lowest skill intensities can be observed for the third quantile. The reason for this pattern is that manufacturing activities are often located on the outskirts of middle-sized cities, therefore the share of production-related jobs (e.g. machine operators, assemblers) that require less cognitive and social skills is somewhat higher in these areas.

RESULTS

Baseline estimates

Table 1 presents the baseline results of our regression analysis. Column 1 contains the cognitive skill index, the agglomeration metric and their interaction as explanatory variables. shows that the return to cognitive skills as imputed from the employers' skill requirements grows with the degree of agglomeration. Specifically, a unit increase in log a_r increases the partial effect of cognitive skills by 2.8 percentage points. Column 2 and 3 include individual and firm-level characteristics into the model. While worker characteristics seem to have a smaller impact on the interaction term, it drops considerably after the inclusion of firm-level controls in column 3. These results suggest that cognitive skills operate through schooling and worker selection into firms with different attributes.

[TABLE 1 AROUND HERE]

Results for social skills are presented in the next three columns. The key result is that the return to social skills increases with urban density. However, unlike in the case of cognitive skills, this result persists even if we control for worker and firm characteristics and 2-digit NACE industries. According to estimates of column 6, a unit change in the natural log of the agglomeration metric is associated with a 2.6 percentage point increase in the partial effect of social skills.

When both skill types are included into the model (column 7), the point estimate for the interaction between social skills and agglomeration (δ^j) remain unchanged which suggest that the cognitive skill index do not act as a proxy for social skills. The same results hold even if we look at the skill variation across workers within the same occupation (column 8). Although the joint contribution of occupational dummies to the predictive performance of the model seems to be important (joint F-stat is 82.46 with $p < 0.01$), it does not have a large influence on the key coefficients. The point estimate for δ^j becomes somewhat smaller but it remains significant at the 5% level. These results seem to be robust to the choice of kernel functions and the distance cut-off. Similar estimates are obtained when alternative kernel functions (Gaussian and tricube) are used to construct the agglomeration metric (see Table A3 in the online Appendix) and alternative distance cut-off parameters (d_T) are considered (see Figure A1 in the online appendix).

[FIGURE 2 AROUND HERE]

For a simpler interpretation of the results in Table 1, it is worth estimating the marginal effect of skills at different agglomeration levels. Figure 2 plots the partial effect from one standard deviation increase in skills at different agglomeration levels based on Column 7 in Table 1. When log a_r is observed at the 90th percentile (14.7), a one standard deviation increase in social skills is associated with a 12.1 percent increase in earnings, whereas at the 10th percentile (10.2), the corresponding effect is -3.2 percent but it is not significant. At the median value of log a_r (11.5) the partial effect is virtually zero (-0.4 percent). For cognitive skills the partial effects vary between 5.9 to 6.7 percent but due to large standard errors these effects are never significant.

Endogeneity issues

As noted in Section 2, the Mincerian approach might be exposed to endogeneity issues caused by unobserved worker heterogeneity, measurement error in skills and omitted variables correlated with both agglomeration and wages. In this section we pursue several approaches designed to address these issues.

The first problem we consider is related to unobserved worker heterogeneity. If cognitive and social skill indices capture the effects of personality traits, abilities and other unobserved characteristics, the interaction terms in Eq. (1) will be biased. Moreover, if workers self-select across locations according to these characteristics, the role of agglomeration will be overstated in both interaction terms. To examine the possible extent of omitted variable bias we follow Oster (2019) who proposes a simple procedure to calculate bounding values for unbiased coefficients. This test relies on the assumption that selection on observable covariates from a basic towards a full model is proportional to the selection on unobserved variables (Altonji et al., 2005; Oster, 2019). Assuming that selection in unobservables is equal to selection on observable covariates, and taking columns 4 and 8 of Table 1 as the basic and full models for social skills, Oster's bias-adjusted coefficient for the interaction between social skills and agglomeration is 0.0195. This suggest that after controlling for firm-level variables and occupational dummies any remaining bias related to omitted variables is relatively small, as the bias-adjusted coefficient is close to our prior estimates.[6]

Although this simple heuristic provides some circumstantial evidence that omitted variables bias is not a critical issue, it cannot be ruled out entirely. For example, prior estimates might be biased due to unobserved amenities and location-specific demand shocks. To overcome this issue, we follow the common approach of using historical data on the distribution of population as an instrument for the agglomeration metric. The underlying idea of such historical instruments is that the late 19th century distribution of population is by no means correlated with the recent distribution of amenities and demand shocks but it still predicts the extent of agglomeration well (Combes & Gobillon, 2014). However, because the agglomeration metric is part of both interaction terms in Eq. (1), other instruments should be identified to treat the endogeneity of these interaction terms as well. A natural instrument for the interaction between the agglomeration metric and the cognitive skill index is the log of 1880 population multiplied with the cognitive skill index while for the other interaction term the product of the historical population and the social skill index can be used (Wooldridge, 2002). Since historical data on 1880 population is only available at the LAU-1 level, for the 2SLS models we define a_r as the total employment in LAU-1 units.

[6] For the interaction between cognitive skills and agglomeration Oster's procedure yields -0.0008.

[TABLE 2 AROUND HERE]

Column 1 of Table 2 replicates the last model in Table 1 with the new agglomeration metric and serves as reference point for the following estimates. Column 2 reports 2SLS estimates where the agglomeration metric and both interaction terms are treated as endogenous. As it is shown at the bottom panel the first stage Sanderson-Windmeijer *F*-statistic rejects the null-hypothesis of weak instruments. The second stage estimates are consistent with those obtained using OLS in column 1. As expected, the point estimate of the agglomeration measure is slightly smaller than the one reported in column 1 (0.034 compared to 0.043), but this specification leaves the point estimate of the interaction terms unchanged.

Another concern regarding our prior results is attenuation bias arising from measurement errors in skills. It is possible that the finding of no urban cognitive premium is due to such errors. Probably, the most straightforward way to address this issue is to use an alternative measure of the same skill type as an instrumental variable (Hanushek et al., 2015). This approach essentially extracts the variation that is common to both the skill index and the instrument in the first stage of a 2SLS model, and uses this variation to estimate skill returns in the second stage. Since speaking skills are presumably associated with all of the items that constitute our social skill index, the standardized item of speaking skills can be used as an instrument for social skills. On the same basis, complex problem solving might be a reasonable instrument for cognitive abilities. Again, the interaction between these instruments and the agglomeration metric can be used as instruments for the interaction terms.

Column 3 shows that the instruments based on the standardized skill item of complex problem solving are strong predictors of the endogenous variables (Sanderson-Windmeijer F-stat: 363.5 with $p<0.001$). In the second stage, the point estimate of δ^C is much larger than the reference estimate in the first column (0.021 compared to -0.005) but its remains insignificant. Column 4 repeats the same exercise for social skills. The excluded instruments seem to be quite strong which means that any bias arising from weak identification is not likely in the model. Second stage results are almost exactly similar to previous models reported in the first two columns. The fact that the estimates in column 4 does not change compared to the baseline indicates that attenuation bias from measurement errors is only a minor issue in our setting. The lack of a significant urban cognitive premium cannot be attributed entirely to measurement error.

Of course, none of the models reported in Table 2 is able to deal with the entire spectrum of endogeneity issues but the consistency across different models supports the robustness of our baseline results.

Heterogeneous returns by worker groups

We continue by looking at different worker groups to examine the heterogeneity in returns to skills along multiple dimensions. As often shown in the literature the task content of jobs substantially varies between female and male workers which may result in different skill returns (Borghans *et al.*, 2014; Bacolod, 2016). While women predominate in care work occupations such as nursing and child care as well as in other service jobs involving cognitive tasks and human interactions, the share of male workers is higher in production-related occupations. Table A4 in the online appendix shows that in our sample women have somewhat higher cognitive and social skills than men. Two-sample t-tests based on group means show significant differences for both skills (cognitive skills: t-test is 3.54, with p-value 0.00; social skills: t-test is 5.89, with p-value 0.00).

[TABLE 3 AROUND HERE]

Columns 1 and 2 of Table 3 show significant gender differences in the returns to social skills. While for male workers the interaction between the social skill index and the agglomeration measure is positive and significant at the 5 percent level, for females the same coefficient is almost zero and insignificant. These results suggest that even though women are well-endowed with social skills, they benefit less from their skills in cities compared to men. One reason for this pattern might be that women face barriers that limit their access to professional networks which does not allow them to exploit their social potentials and harness the benefits of agglomeration (Rosenthal & Strange, 2012). For instance, due to the unequal division of labour within the household, women cannot devote as much time to learning and networking as men (Bacolod, 2016). Moreover, having preschool-age children are shown to discourage women from making job changes that would result in higher wages and more productive matches (Looze, 2017).

Another factor along which returns to skills might vary is educational attainment. To analyse heterogeneity between education groups we split the sample into two parts: one for individuals without high-school graduation (corresponding to fewer than 12 years of schooling), and another for those with high-school graduation or more. Columns 3 and 4 report separate estimates for educational subsamples. These results suggest that the urban wage premium tends to be higher when social skills are coupled with a high level of education. For educated workers the interaction between agglomeration and the social skill index is larger than the corresponding full-sample estimate (0.028 as opposed to 0.022 reported in the last column of Table 1) while for the less educated the interaction between social skills and urban density is virtually to zero.

These results are in line with our expectations as more educated workers are more likely to perform non-routine tasks that involve different types of human interactions. Surprisingly, cognitive skills have no effect in either group when we control for firm characteristics and occupations.

Columns 5 and 6 reports that there is a sharp gap in urban social skills premium between non-manual and manual workers.[7] When we restrict our sample to non-manual workers, we find only a weak association between the hedonic price of social skills and agglomeration (0.018 with s.e. 0.010) while for manual workers we find no sign of higher returns in densely populated urban areas. These results highlight that higher return to social skills in large cities stems in part from cities' greater appreciation of non-manual work.

The remaining columns report estimates for manufacturing, services and high-tech industries. For manufacturing workers, the partial correlation between wages and agglomeration is somewhat higher than for the whole sample but neither cognitive, nor social skills are associated with average monthly earnings (column 7). In contrast, for service sector workers a large positive estimate for the interaction between social skills and agglomeration (0.055 with s.e. 0.020) can be found (column 8). A reasonable explanation for this finding is that localized competition in non-tradable services drives urban firms to customize services to the needs of clients, maintain customer relationships and resolve complaints in order to prevent customer attrition or expand their client base (Storper, 2013). Since these activities require skills such as persuasion, service orientation and social perceptiveness (Borghans et al., 2008), when local competition is fierce, it becomes important for firm to hire high-skill employees and incentivise effort with higher wages. In the manufacturing sector, however, this incentive is of lesser importance as firms producing tradable goods can sell to larger markets and thus face more similar product market competition, irrespective of their location.

Overall, there are considerable heterogeneities across worker groups and sectors in the return to social skills, however cognitive skills are not predictive of earnings in any of the worker groups considered.

[7] Workers performing occupations that belong to either of the first five major ISCO groups (managers, professionals, technicians and associate professionals, clerical support workers, service and sales workers) are classified as non-manual workers, while remaining workers are classified as manual workers.

CONCLUSIONS

This paper examines whether returns to cognitive and social skills are greater in dense urban settings. The novelty of the paper is that it draws on a representative survey containing detailed information on the skill requirements of Hungarian firms which allows us to rule out the possibility that spatial differences in skill returns are driven by firm selection or the 'functional specialization' of locations. Although we cannot neutralize every econometric issue that might bias our results, the consistency of estimates across different model specifications provides some support for the underlying importance of social skills in urban labour markets. This result is consistent with the arguments that emphasize the role of knowledge exchange and learning as possible sources of agglomeration economies (e.g., Davis & Dingel, 2019). Moreover, it is also consistent with the idea that fierce competition in cities raises demand for social skills especially in non-traded services and innovative activities (Borghans *et al.*, 2008). Surprisingly, despite its theoretical underpinning we do not find any relationship between returns to cognitive skills and the degree of agglomeration. In fact, cognitive are not predictive of wages which might be the result of the educational expansion that increased the supply of cognitive skills in the last two decades in Hungary. Since cognitive skills can be more easily developed through education and informal human capital investment (Cawley et al., 2001), access to tertiary education plays a major role in driving down the price of cognitive skills.

The results of this paper contribute to a number of other related literatures as well. For example, previous research has shown that early childhood interventions may have long-term benefits for a number of adult outcomes (Cunha & Heckman 2007; Chetty et al., 2011). Although this paper does not examine where skills come from, our results suggest that social skills acquired at an early age may have positive effects on lifetime earnings which makes a case for focusing more on the development of interpersonal skills at every stage of education. As shown in Section 4 even high-school graduates can benefit from their social capabilities, thus building these skills should be started earlier than higher education.

Another strand of the literature investigates the sources of urban wage premium. Social skills might play an important role in explaining higher individual earnings in cities. First of all, workers with better social capabilities may choose to locate in dense urban and self-select into well-paid occupations. Second, since all sorts of agglomeration economies involve some kind of human interaction, social skills might help harness the external benefits of agglomeration and facilitate specialization by reducing coordination costs within firms (Becker & Murphy, 1992; Deming, 2017). An interesting follow-up would be to formalize the mechanisms underlying these ideas in a spatial equilibrium framework. The results of this paper provide a strong empirical rationale for such theoretical investigations.

REFERENCES

Albrecht, J., Vroman, S. (2002). A matching model with endogenous skill requirements. *International Economic Review*, 43 (1), 283–305.

Altonji, J. G., Elder, T. E. & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1), pp. 151–184.

Andersson, M., Klaesson, J. & Larsson, J. P. (2014). The sources of the urban wage premium by worker skills: Spatial sorting or agglomeration economies? *Papers in Regional Science*, 93(4), 727–747.

Autor, D., & Handel, M. (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labour Economics*, 31(2), 559–596.

Bacolod, M. (2016). Skills, the gender gap and cities. *Journal of Regional Science*, 57(2), 290–318.

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

Barbour, E., & Markusen, A. (2007). Regional occupational and industrial structure: does one imply the other? *International Regional Science Review*, 30(1), 72–90.

Baum-Snow, N. & Pavan, R. (2012) Understanding the city size wage gap. *The Review of Economic Studies*, 79 (1), 88–127.

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

Borghans, L., ter Weel, B., & Weinberg, B. A. (2008). Interpersonal styles and labour market outcomes. *Journal of Human Resources*, 43(4), 815–858.

Borghans, L., ter Weel, B., & Weinberg, B. A. (2014). People skills and the labour-market outcomes of underrepresented groups. ILR Review, 67(2), 287-334.

Cawley, J., Heckman, J. J., & Vyclatil, E. (2001). Three observations on wages and measured cognitive ability. *Journal of Labour Economics*, 8 (4), 419–442.

Chade, H., Eeckhout, J., & Smith, L. (2017). Sorting through search and matching models in economics. *Journal of Economic Literature*, 55(2), 493–544.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from project Star. *Quarterly Journal of Economics*, 126 (4), 1593–1660.

Choi, T. (2020). Agglomeration effect of skill-based local labour pooling: Evidence of South Korea. *Sustainability*, 12(8), 3198.

Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63 (2), 723–442.

Cunha, F., & Heckman, J. J. (2007). The technology of skill formation. *American Economic Review*, 97 (2), 31–47.

Davis, D. R., & Dingel J. I. (2019). A spatial knowledge economy. *American Economic Review*, 109 (1), 153–170.

Deming, D. J. (2017). The growing importance of social skills in the labour market. *The Quarterly Journal of Economics*, 132(4), 1593–1640.

Deming, D. J., & Kahn, L. B. (2018). Skill requirements across firms and labour markets: Evidence from job postings for professionals. *Journal of Labour Economics*, 36(S1), S337–S369.

Feser, E. (2003). What regions do rather than make: A proposed set of knowledge-based occupation clusters. *Urban Studies*, 40 (10), 1937–1958.

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

Glaeser, E. L., & Maré, D. C. (2001). Cities and skills. *Journal of Labour Economics*, 19(2), 316–342.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, 73 (1), 103–130.

Heckman, J. J., Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4), 451–464.

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006a). Earnings functions, rates of return and treatment effects: the Mincer equation and beyond. In E. A. Hanushek, & F. Welch (Eds.), *Handbook of the Economics of Education vol. 1*, pp. 307–458. Amsterdam, North-Holland.

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

Imbens, G., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 42 (2), 615–635.

Ingram, B. F., & Neumann, G. R. (2006). The returns to skill. *Labour Economics*, 13(1), 35–59.

Koster, H. R. A. (2013). Rocketing rents the magnitude and attenuation of agglomeration economies in the commercial property market. *SERC Discussion Papers 0125*, Spatial Economics Research Centre, LSE, London.

Koster, H. R. A., van Ommeren, J., Rietveld, P. (2014). Agglomeration economies and productivity: A structural estimation approach using commercial rents. *Economica*, 81(321), 63–85.

Lindqvist, E., & Vestman, R. (2011). The labour market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics*, 3 (1), 101–128.

Looze, J. (2017). Why do(n't) they leave?: Motherhood and women's job mobility. *Social Science Research*, 65, 47–59.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37 (2), 187–204.

Rosenthal, S., & Strange, W. (2012). Female entrepreneurship, agglomeration, and a new spatial mismatch. *The Review of Economics and Statistics*, 94(3), 764–788.

Scott, A. (2009). Human capital resources and requirements across the metropolitan hierarchy of the USA. *Journal of Economic Geography*, 9(2), 207–226.

Shimer, R., & Smith, L. (2000). Assortative Matching and Search. *Econometrica*, 68(2), 343–369.

Storper, M. (2013). *Keys to the City: How Economics, institutions social interactions, and politics shape development*. Princeton University Press.

Szakálné Kanó, I., Kazemi-Sánta, É. & Lengyel, I. (2017). Territorial distribution of highly educated individuals in Hungary after 1990. *Regional Statistics*, 7 (2), 171–189.

Varga, J. (2020). The labour market situation of young graduates, overqualification and the value of higher education degrees. In: Fazekas, K., Csillag, M., Hermann, Z. & Scharle, Á. (eds.) *The Hungarian Labour Market 2019*. Institute of Economics, Centre for Economic and Regional Studies, Budapest, 137–143.

Wheeler, C. H. (2006). Cities and the growth of wages among young workers: Evidence from the NLSY. *Journal of Urban Economics*, 60(2), 162–184.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

TABLE 1 Estimates on returns to social skills

Dep. var.: log of gross monthly wage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Agglomeration	0.078***	0.052***	0.046***	0.079***	0.052***	0.050***	0.051***	0.043***
	(0.014)	(0.012)	(0.009)	(0.014)	(0.012)	(0.009)	(0.009)	(0.009)
Cognitive	-0.038	-0.161	-0.042				0.086	-0.047
	(0.157)	(0.114)	(0.084)				(0.138)	(0.130)
Cognitive × Agglomeration	0.028**	0.020**	0.012				-0.008	0.006
	(0.013)	(0.010)	(0.008)				(0.012)	(0.011)
Social				-0.220	-0.330***	-0.219**	-0.288**	-0.289**
				(0.192)	(0.120)	(0.098)	(0.137)	(0.131)
Social × Agglomeration				0.033**	0.030***	0.026***	0.028**	0.022**
				(0.016)	(0.010)	(0.008)	(0.012)	(0.011)
R-squared	0.199	0.425	0.602	0.150	0.422	0.605	0.607	0.671
Individual controls		Yes	Yes		Yes	Yes	Yes	Yes
Firm controls			Yes			Yes	Yes	Yes
Occupation dummies								Yes
Distance cut-off	15 km	17 km	18 km	16 km	17 km	18 km	18 km	19 km

Note: Standard errors clustered by firms in parentheses. All regressions include a constant term and a dummy for the year 2011. Individual controls include sex, age, age squared, part-time work and educational attainment. Firm-level controls include foreign and state ownership dummies, collective agreement, firm size and industry dummies (2-digit NACE Rev. 2. classes). The agglomeration measure is calculated using Epanechnikov kernel. Number of observations: 14,990. ***, ** and * indicate significance at the 1, 5 and 10% levels.

TABLE 2 Endogeneity issues

dep. var.: log of gross monthly wage	LAU-1	Endogenous agglomeration	Measurement error I.	Measurement error II.
	(1)	(2)	(3)	(4)
Agglomeration	0.043***	0.034***	0.041***	0.044***
	(0.010)	(0.011)	(0.010)	(0.010)
Cognitive	0.092	-0.056	-0.276**	
	(0.145)	(0.160)	(0.127)	
Cognitive × Agglomeration	-0.005	0.007	0.021	
	(0.012)	(0.013)	(0.012)	
Social	-0.368**	-0.355**		-0.282**
	(0.143)	(0.158)		(0.139)
Social × Agglomeration	0.029**	0.028**		0.027**
	(0.012)	(0.013)		(0.011)
R-squared	0.667	0.666	0.665	0.665
Estimation method	OLS	2SLS	2SLS	2SLS
Endogenous variables		Aggl, Cogn × Aggl, Soc × Aggl,	Cogn, Cogn × Aggl,	Soc, Soc × Aggl
Excluded instrument		Log pop in 1880, + interactions with skills	Complex skills, + interaction with aggl skills	Speaking skills + + interaction with aggl skills
Sanderson-Windmeijer F-test		1861.7 ($p < 0.001$)	363.5 ($p < 0.001$)	335.6 ($p < 0.001$)
Individual controls	Yes	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes

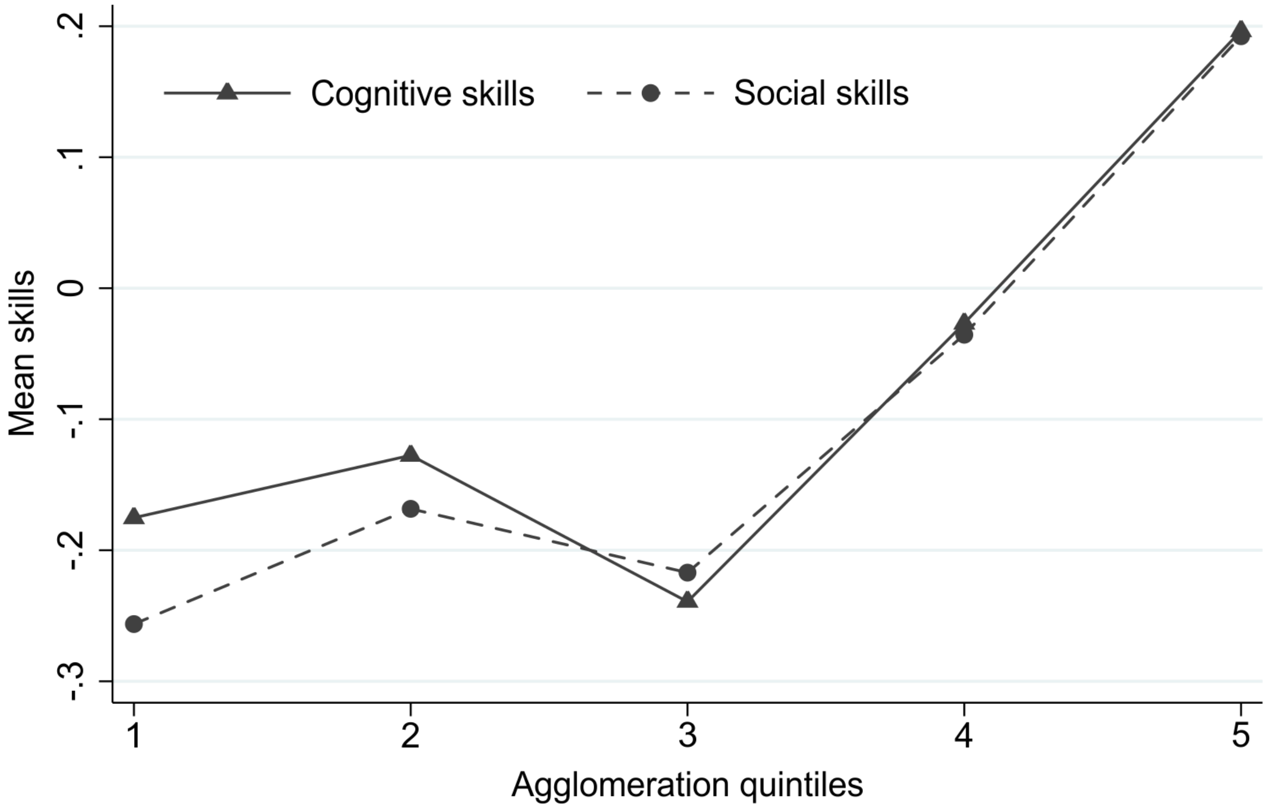
Note: Standard errors clustered by firms in parentheses. All regressions include a constant term and a dummy for the year 2011. Individual controls include sex, age, age squared, part-time work and educational attainment. Firm-level controls include foreign and state ownership dummies, collective agreement, firm size and industry dummies (2-digit NACE Rev. 2. classes). Number of observations: 14,990. ***, ** and * indicate significance at the 1, 5 and 10% levels.

TABLE 3 Estimates on labour market returns to skills for different sub-samples of workers

	Men	Women	High school or more	Less than high-school	Non-manual	Manual	Manufacturing	Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Agglomeration	0.039***	0.053***	0.064***	0.017*	0.052***	0.015	0.052***	0.036**
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.018)	(0.010)	(0.016)
Cognitive	0.079	-0.206	0.036	0.095	-0.059	0.130	-0.086	0.405
	(0.149)	(0.147)	(0.147)	(0.154)	(0.154)	(0.172)	(0.162)	(0.287)
Cognitive × Agglomeration	-0.004	0.017	-0.001	-0.006	0.007	-0.011	0.010	-0.026
	(0.013)	(0.013)	(0.012)	(0.014)	(0.012)	(0.016)	(0.014)	(0.022)
Social	-0.365**	-0.081	-0.369***	-0.135	-0.189	-0.032	-0.214	-0.795***
	(0.156)	(0.143)	(0.143)	(0.158)	(0.143)	(0.281)	(0.168)	(0.257)
Social × Agglomeration	0.028**	0.006	0.028***	0.010	0.018*	0.004	0.018	0.055***
	(0.013)	(0.012)	(0.011)	(0.015)	(0.010)	(0.026)	(0.014)	(0.020)
R-squared	0.676	0.691	0.653	0.574	0.654	0.604	0.685	0.708
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance cut-off	17 km	18 km	20 km	9 km	20 km	9 km	18 km	20 km

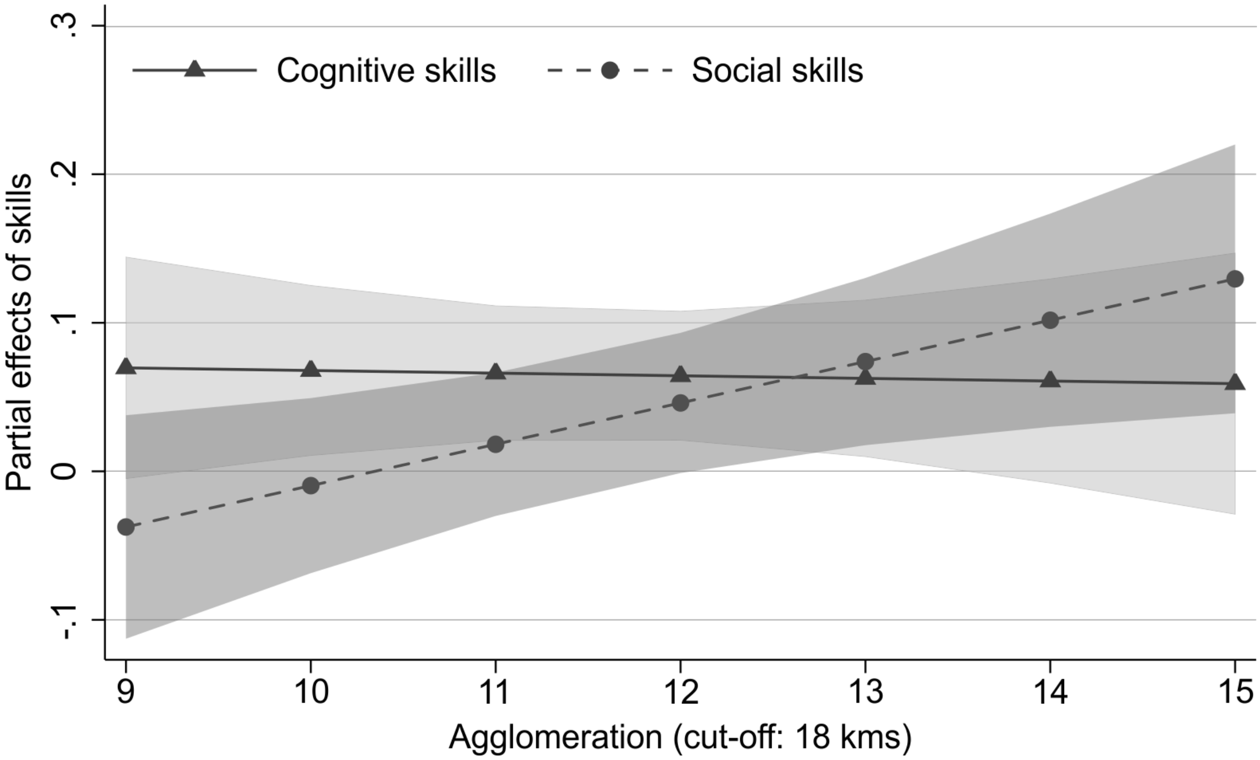
Note: Standard errors clustered by firms in parentheses. All regressions include a constant term and a dummy for the year 2011. Individual controls include sex, age, age squared, part-time work and educational attainment. Firm-level controls include foreign and state ownership dummies, collective agreement, firm size and industry dummies (2-digit NACE Rev. 2. classes). The agglomeration measure is calculated using Epanechnikov kernel. Number of observations: 14,990. ***, ** and * indicate significance at the 1, 5 and 10% levels.

FIGURE 1 Spatial distribution of skills



Notes: Authors' own calculations based on the IE-HAS’s skill survey and data from the 2010 and 2011 annual wage surveys. The agglomeration measure is calculated using Epanechnikov kernel. Distance cut-off: 18 kms.

FIGURE 2 Partial effects of skills at different agglomeration levels



Notes: Estimates are based on column 7 in Table 1.

TABLE A1 Description of skill items

Skill items	Description
Cognitive skills	
Active learning	Understanding the implications of new information for both current and future problem-solving and decision-making.
Active listening	Giving full attention to what other people are saying, taking time to understand the points being made, asking questions as appropriate, and not interrupting at inappropriate times.
Critical Thinking	Using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems.
Learning Strategies	Selecting and using training/instructional methods and procedures appropriate for the situation when learning or teaching new things.
Mathematics	Using mathematics to solve problems.
Monitoring	Monitoring/Assessing performance of yourself, other individuals, or organizations to make improvements or take corrective action.
Reading Comprehension	Understanding written sentences and paragraphs in work related documents.
Social skills	
Coordination	Adjusting actions in relation to others' actions.
Instructing	Teaching others how to do something.
Negotiation	Bringing others together and trying to reconcile differences.
Persuasion	Persuading others to change their minds or behaviour.
Service orientation	Actively looking for ways to help people.
Social Perceptiveness	Being aware of others' reactions and understanding why they react as they do.

Notes: Descriptions are based on O*NET.

TABLE A2 Descriptive statistics

Variable	Mean	Std.dev.	Min	Max
Main variables				
Log gross montly wage	11.98	0.64	10.99	15.76
Cognitive skills	0	1		
Social skills	0	1	-1.10	2,98
Agglomeration (distance cutoff: 18 km)	12,01	1,79	8,93	14,72
Controls				
Sex (male = 1)	0.61	0.49	0	1
Experience	24.19	11.57	1	65
Cognitive skills	0	1	-1.54	2,82
Education (primary)	0.15	0.36	0	1
Education (lower secondary)	0.37	0.48	0	1
Education (upper secondary)	0.31	0.46	0	1
Education (tertiary)	0.17	0.38	0	1
Part-time work	0.17	0.38	0	1
Collective agreement	0.26	0.44	0	1
Foreign ownership (>50% = 1)	0.16	0.37	0	1
State ownership (>50% = 1)	0.17	0.37	0	1
Firm size (5-21)	0.14	0.32	0	1
Firm size (21-50)	0.30	0.46	0	1
Firm size (51-300)	0.28	0.45	0	1
Firm size (301-1000)	0.15	0.36	0	1
Firm size (1000+)	0.13	0.33	0	1
Instruments				
Speaking skills	0	1	-1.41	2,11
Complex problem solving skills	0	1	-1.06	2.31
Log population in 1880	11.13	1.21	8.63	12.91

Notes: Authors' own calculations based on the IE-HAS's skill survey and data from the 2010 and 2011 annual wage surveys. Number of observations: 14,990.

TABLE A3 Robustness on the choice of the kernel function

Dep. var.: log of gross monthly wage	Epanechnikov (from Table 1) (1)	Gaussian (2)	Tricube (3)
Agglomeration	0.043***	0.044***	0.042***
	(0.009)	(0.009)	(0.009)
Cognitive	-0.031	-0.063	-0.018
	(0.128)	(0.126)	(0.127)
Cognitive × Agglomeration	0.005	0.008	0.004
	(0.011)	(0.011)	(0.011)
Social	-0.293*	-0.267**	-0.298**
	(0.149)	(0.127)	(0.128)
Social × Agglomeration	0.022**	0.021**	0.022**
	(0.011)	(0.010)	(0.011)
R-squared	0.671	0.671	0.670
Individual controls	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes
Distance cut-off	17 km	18 km	20 km

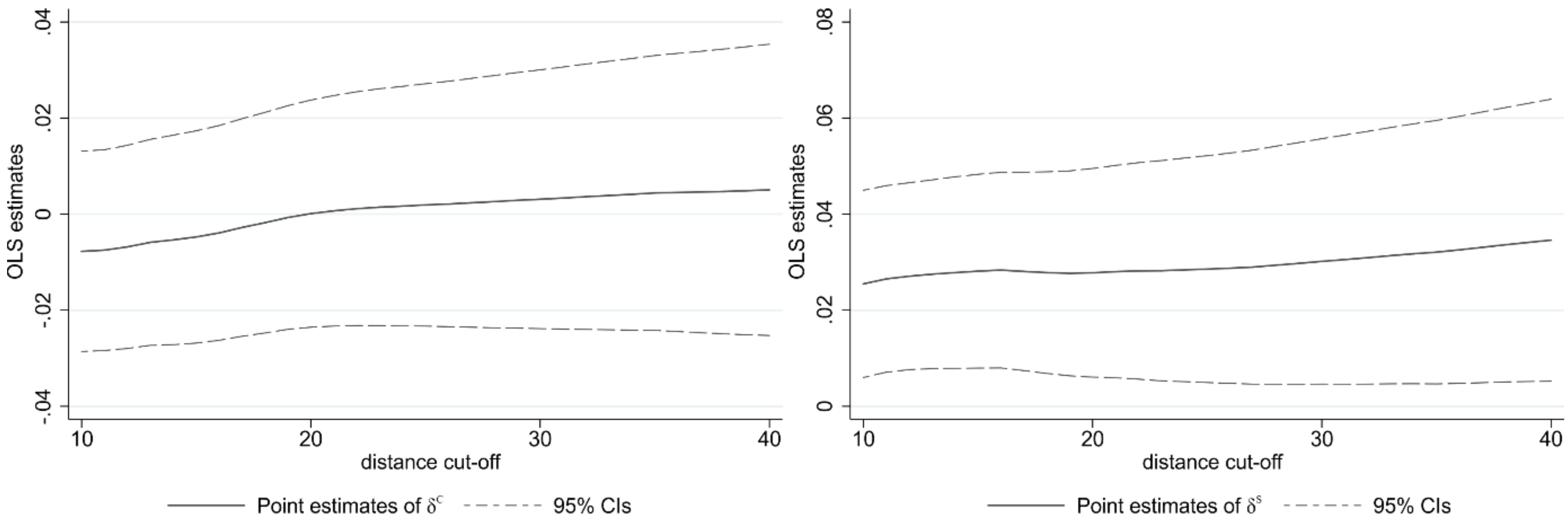
Note: Standard errors clustered by firms in parentheses. All regressions include a constant term and a dummy for the year 2011. Individual controls include sex, age, age squared, part-time work and educational attainment. Firm-level controls include foreign and state ownership dummies, collective agreement, firm size and industry dummies (2-digit NACE Rev. 2. classes). Column 1 replicates the last model in Table 1 for easier comparison. Number of observations: 14,990. ***, ** and * indicate significance at the 1, 5 and 10% levels.

TABLE A4 Cognitive and social skills by worker groups and industries

	Cognitive skills		Social skills		N
	Mean	SD	Mean	SD	
Women	-0.03	1.03	0.06	1.03	5,794
Men	0.01	0.98	-0.04	0.98	9,196
Age (<40 years)	0.01	0.97	0.02	0.98	6,642
Age (≥40 years)	-0.01	1.02	-0.02	1.02	8,348
Less than high school	-0.71	0.62	-0.61	0.51	7,757
High school or more	-0.27	0.75	-0.25	0.79	7,233
Non-manual workers	0.59	1.02	0.65	1.03	5,046
Manual workers	-0.52	0.62	-0.58	0.50	9,944
Manufacturing	0.00	0.98	-0.01	1.01	8,288
Services	0.02	1.02	0.05	0.99	5,870
Overall sample	0.00	1.00	0.00	1.00	14,990

Note: Statistics are calculated using sampling weights. High-tech industries are defined using the classification of Eurostat (see Table A3). Other groups are as defined in the main text.

FIGURE A1 Robustness to the choice of distance cutoffs.



Notes: The agglomeration measure is calculated using Epanechnikov kernel. The estimated model is similar to the one presented in column 7 of Table 1. Confidence intervals are calculated using clustered standard errors and the delta method.