ISSN 1503-299X

WORKING PAPER SERIES

No. 14/2013

Education, experience and dynamic urban wage premium

Fredrik Carlsen, Jørn Rattsø, and Hildegunn E. Stokke,
Department of Economics,
Norwegian University of Science and Technology

Department of Economics

Norwegian University of Science and Technology
N-7491 Trondheim, Norway
www.svt.ntnu.no/iso/wp/wp.htm

Education, experience and dynamic urban wage premium[*]

Fredrik Carlsen, Jørn Rattsø and Hildegunn E. Stokke

Department of Economics, Norwegian University of Science and Technology

fredrik.carlsen@svt.ntnu.no; jorn.rattso@svt.ntnu.no; hildegunnes@svt.ntnu.no

Abstract

We analyze static and dynamic agglomeration effects across education groups. The data are based on administrative registers covering all full time workers in the private sector of Norway during 2001-2010, about 6.5 million worker-year observations, including place and sector of work experience since 1993. Accounting for unobservable abilities with identification based on movers, the static urban wage premium is similar across education groups. When the history of work experience in different regions and sectors is included, we show that the dynamic wage premium increases in education level and that highly educated in high wage sectors have the largest learning advantage.

JEL codes: J24, J31, J61, R12, R23

Key words: Agglomeration economies, sorting, education, worker experience

Date: November 1, 2013

[*] We appreciate discussions at the 2012 European Meeting of the Urban Economics Association, the 2012 North American Meeting of the Regional Science Association International, the 2013 Meeting of the Western Regional Science Association, the 2013 Congress of the European Economic Association, the Urban and Regional Economics seminar at LSE, and at the Labour Economics/Empirical Microeconomics workshop at the Norwegian School of Economics, and comments from Ragnhild Balsvik, Nate Baum-Snow, Paul Cheshire, Jorge De la Roca, Steve Gibbons, Christian Hilber, Ana I. Moreno-Monroy, Jarle Møen, Henry Overman, Kjell Gunnar Salvanes and Olmo Silva. We are grateful for research assistance from Paivi Lujala and the cooperation of Statistics Norway. An earlier version of this paper was titled "Urban wage premium and the role of education: Identification of agglomeration effects for Norway".

1. Introduction

Highly educated individuals have higher productivity and tend to live in cities. The observation has led to a literature dealing with 'skilled cities' or 'smart cities' (Glaeser and Saiz, 2004, Shapiro, 2006, and Winters, 2011). The urban concentration of highly educated explains a large part of the observed urban – rural wage gap. We investigate the importance of education and address both unobservables and the geographic and sectoral history of experience. The approach allows for separation between static and dynamic agglomeration effects and links the dynamic effect to both location and employment sector of experience. The bottom line is that the learning advantage of living in large cities is increasing in the level of education and is largest for highly educated in high wage sectors.

Recent studies of agglomeration effects separating between education groups apply individual level census data. Hedonic regressions clarify how the wages reflect characteristics of the workers and allow for the estimation of a regional fixed wage effect that controls for heterogeneity. Most authors conclude that static agglomeration effects are higher for those with the highest education level. Wheeler (2001) shows how the effect is increasing with the level of education. Rosenthal and Strange (2008) find that the urban wage premium for workers with college degree is higher than the rest. Bacolod et al. (2009) concludes that the effect of population size increases with education level, but that the effect is equal for workers with college and high school degrees. Some contrarian evidence is published by Lee (2010) involving health care workers. He concludes that the urban wage premium decreases as the skill level rises. The interpretation is that high skilled prefer to live in large cities and smaller cities must pay more to recruit them.

Glaeser and Mare (2001) innovated the handling of heterogeneity and sorting using individual panel data and individual fixed effects for the US. They conclude that larger local markets have higher wages and productivity even when heterogeneity is taken into account. Combes et al. (2008) argue that sorting is more important, but they do not have observation of education level and their individual fixed effect consequently represents a mixed bag of education, skill, experience and ability. Recent contributions have added an investigation of experience and where experience is made. Andersson et al. (2013), Baum-Snow and Pavan (2012) and Gould (2007) find that wage-level effects and return to experience are the most important factors explaining the wage premium and that moving to cities promote human capital accumulation.

Baum-Snow and Pavan and the analysis of Gould work with structural models and smaller datasets. De la Roca and Puga (2012) have rich register data of Spain and estimate a model with identification based on movers and including the individual history of experience. They find that working in a larger city gives an immediate wage premium that is expanded over time when living in a large city. De la Roca (2011) expands the analysis by looking at both initial and return migration. D'Costa and Overman (2013) show that workers with experience in cities experience higher wage growth. In recent analyses addressing the spatial distribution of college graduates, Wang (2013) confirms that unobserved abilities matter for the sorting and Ahlin et al. (2013) elaborate on the individual characteristics of importance. Our main contribution is the analysis of experience across education groups and with analysis of the interaction between geography and sector of experience.

We use administrative register data for the whole working population of Norway during 2001 to 2010. We exclude part time workers and workers in the public and primary sectors and apply a dataset with about 6.5 million worker-year observations. The rural-urban dimension is included by separating between 89 labor market regions characterized by population size. The role of industrial composition is represented by 54 industrial sectors. The key worker observables are education level and work experience from 1993 onwards. In the analysis we separate between primary, secondary and tertiary education. The experience is distinguished between large cities versus smaller regions, and high wage sectors versus the rest. We start out with a raw agglomeration elasticity of hourly wages with respect to population size of 0.066. When considering the wage premium for observationally equivalent workers, the elasticity drops to 0.045. A doubling of the population size approximately increases the wage level by 4.5%. This is consistent with the literature. Introducing individual fixed effects and identification based on movers (13% of the workers move at least once during 2001-2010) reduces the elasticity to 0.033. Controlling for observable worker characteristics and unobserved ability, the elasticity is down to about half of the raw elasticity.

Existing studies accounting for observed individual characteristics show that the static agglomeration effect increases with the level of education. The Norwegian data are consistent with this and comparable to Wheeler (2001) we find elasticities of 0.03 for primary educated, 0.044 for secondary educated and 0.062 for tertiary educated. The differences across education groups disappear when we correct for unobserved abilities. All three education groups have an agglomeration elasticity of about 0.03. It follows that the ability selection is more important for the higher education groups.

Glaeser and Resseger (2010) introduce worker experience in the understanding of the wage premium, and their results suggest that learning effects are stronger in more skilled metropolitan areas. De la Roca and Puga (2012) relate work experience and residence to investigate how urbanization matters for learning and to calculate the dynamic wage premium. Following their approach we analyze learning across education groups by separating between experience made in large cities (more than 150 000 inhabitants, basically covering the 4 large cities in Norway) and the rest. Also the experience with respect to industrial sector is taken into account, distinguishing between the top 10 high wage sectors and the rest. De la Roca and Puga find that workers in large cities accumulate more valuable experience. We show that this learning effect depends on both education level and sector of employment. The initial wage premium is not affected by the inclusion of worker experience history, but the experience effect adds to the medium-term wage premium because of the large city advantage in experience. The calculated wage effect given the average 8 years of local experience increases the agglomeration elasticity to 0.047 in high wage sectors and 0.043 in other sectors. The medium-term premium is 3.3% for a low educated worker outside the top wage sectors, while a highly educated with experience in a high wage sector has a premium of 4.7%. The experience effect is portable and is hardly reduced when the individual moves to a smaller region.

The methodological challenges of endogeneity and omitted variables have been addressed in some recent studies, but generally the OLS bias is seen as unimportant. Our analysis suggests that the bias may be of particular importance for highly educated. The studies addressing endogeneity usually apply historical population densities as instruments. To avoid the persistence effect and include first nature geography we use annual change of population from 1875 to 1920 and share of regional area with mountains and lakes as instruments for population size. The IV estimates of static agglomeration effects confirm that the OLS upward bias is small. However, for the medium-term premium there is downward OLS bias for highly educated, and the bias is quite large. Our understanding is that highly educated have broader motivation to move to large cities and that the consequent crowding has negative effects for productivity. The dynamic wage premium for a person with tertiary education working in a high wage sector is about the double compared to a person with primary education outside the high wage sectors. The results are robust to the inclusion of four types of amenities shown to be important in the literature: School quality, crime, cultural services, and winter temperature.

Section 2 discusses our econometric strategy and data. The estimates of static agglomeration effects across education groups are presented in section 3. Section 4 moves on to dynamic agglomeration effects based on experience effects dependent on education and sector. The robustness of the results is investigated in section 5. Section 6 summarizes our conclusions and indicates future research.

2. Econometric strategy and data

The analysis deals with the economic importance of urbanization. The measurement of urban scale requires a definition of the agglomeration relevant for the variation in income and productivity. Black and Henderson (2003) discuss the basic understanding of urban evolution. Empirical studies have used measures of size and density of employment and population. We concentrate on population size here. Population density is most relevant if the agglomeration effects work over short distances within city areas. Norwegian cities are small by international comparison, and most regions have large unpopulated areas. In this setting population size is the best measure of the degree of urbanization. Based on information about commuting flows between municipalities, Statistics Norway has divided Norway into 89 travel-to-work areas, denoted economic regions. The economic regions conform to NUTS-4 regions, as defined by the European Union standard of regional levels. This level of aggregation captures functional regions understood as common labor markets.

Our measures of the regional wage levels are computed from three administrative registers: the employment, tax and education registers. The employment register links workers and firms, and gives information on work contracts for all employees during 1993-2010. It includes the number of days worked, the type of contract[1], and from 2001 onwards it also has information on the exact number of hours worked per week. Based on this we calculate the number of hours worked per year, which is combined with data on annual wage income from

[1] The employment register separates between three contract types: Full time contracts with at least 30 hours work per week, part time contracts with 20 – 29 hours work per week, and part time contracts with less than 20 hours work per week.

the tax register to give a measure of hourly wages for all employees during 2001-2010.[2] Information about the contracts back to 1993 is used to calculate a measure of work experience for each worker. We concentrate on workers with full time contracts (at least 30 hours per week). Workers with more than two contracts during a year, as well as workers with one full time and one part time contract are excluded. Workers with two full time contracts are excluded if the number of days worked that year exceeds 455 days. This means that we allow for a maximum of 3 months overlap between the two contracts. To avoid extreme observations, we exclude individuals working less than 50 hours or more than 3500 hours per year. Similar, workers with hourly wage below 70 NOK or above 1250 NOK are also excluded. Finally, we focus on workers between 25 and 65 years old.

The workers are allocated to 60 employment sectors. Since the productivity of resource based sectors are unrelated to urbanization, we exclude the primary sectors agriculture, fishing and forestry. In the public sector wages are determined by national regulation and public sector workers are excluded (sectors public administration, education and health care). We are left with 54 sectors and the largest are construction, domestic trade, retail sales and business services. The education register covers the whole adult population and gives information about the highest completed education level in the beginning of October each year. We also have information on the age, gender, immigration status and home region of all individuals.

The final dataset includes about 650 000 workers every year during 2001-2010, giving a total of about 6.5 million worker-year observations. Workers can enter and leave the dataset during the 10-year period, and in total 1.1 million different workers are included in the dataset. We compute regional wage measures for the whole sample of workers, as well as for three subgroups of workers according to the level of education: tertiary (workers that have completed at least one year at college/university), secondary (workers that have completed at least one year of secondary education) and primary (workers with not more than compulsory schooling). Workers with secondary education account for the largest share with about 3.4 million observations. The subgroups of workers with primary and tertiary education contain 1.3 million and 1.8 million observations, respectively.

[2] Self-employed workers are not included.

As noticed in the introduction, the heterogeneity of the population represents an important challenge in the estimation of agglomeration effects. Geographical sorting of workers may create correlations between urban scale/density and observable and unobservable workers characteristics, such as education, experience and ability. Sorting may therefore introduce measurement errors in the regional wages. Our first measure of the regional wage levels controls for observable worker characteristics, as well as sector and year fixed effects, and is based on estimation of the following hedonic equation for the period 2001 – 2010:

$$\ln w_{irst} = \sigma_r + \mu_s + \gamma_t + X_{it}\beta + \varepsilon_{irst} \tag{1}$$

where w_{irst} is the hourly wage income for worker *i* in region *r*, sector *s* and year *t*, σ_r is the set of regional indicators, while sector and year fixed effects are represented by μ_s and γ_t, respectively. β is a vector of parameters and ε_{irst} is an error term. The vector of observable worker characteristics (X_{it}) includes dummies for age (5-year intervals), gender, immigration status and education level (primary, secondary, tertiary), as well as a measure of work experience calculated in days (and expressed in years) from 1993 onwards. When estimates are done for subgroups according to the level of education, the education dummies are not included in X_{it}.

Since we have panel data of workers, movements between regions can be used to control for unobservable worker characteristics. In our dataset, 13% of the workers move at least once during 2001 – 2010. To estimate our second measure of the regional wage levels worker fixed effects (η_i) are added to the hedonic equation above:

$$\ln w_{irst} = \sigma_r + \mu_s + \gamma_t + \eta_i + X_{it}\beta + \varepsilon_{irst} \tag{2}$$

The estimated regional indicators (σ_r) give a measure of the regional wage levels that controls for sorting of workers based on both observable characteristics like education level and experience and unobservable characteristics like abilities.[3] Sector and year fixed effects are still taken into account.

[3] Observable worker characteristics that are time-invariant (gender and immigration status) are not included in this regression.

As argued by De la Roca and Puga (2012), the regional wage measures estimated in equations (1) and (2) only capture static urban wage premiums, while working in urban areas are likely to affect wages over time through learning effects. They capture the dynamic effect on regional wages by allowing the value of experience to vary with the type of region it is both accumulated and currently used in. We follow this approach and separate out the large city regions from the rest. The large city group is defined by labor market regions with more than 150 000 inhabitants in 2010. The group consists of 7 regions that basically cover the 4 large cities in Norway: Oslo, Stavanger, Bergen and Trondheim. As an extension of previous research, we also separate out the top 10 high wage sectors to check whether the learning effect varies with sectoral experience. We follow the suggestion of Liu et al. (2012) and rank industries with respect to the sectoral wage premium estimated. Our third, and preferable, measure of regional wages is therefore based on estimation of the following hedonic equation:

$$\ln w_{irst} = \sigma_r + \mu_s + \gamma_t + \eta_i + X_{it}\beta + \sum_{j=1}^{2}\sum_{k=1}^{2}\delta_{jkr} e_{ijkt} + \varepsilon_{irst} \tag{3}$$

where e_{ijkt} is the work experience acquired by worker *i* in region type *j* and sector type *k* up until time *t*, and δ_{jkr} are parameters. With this specification we are able to separate between the static and the dynamic urban wage premium. The static premium is still given by the estimated regional indicators (σ_r). To calculate the medium-term premium we use the estimated coefficients in equation (3) to add the wage effect of the average worker experience in a region to the static wage premium. The learning effect is allowed to differ between the large cities and the rest, and dependent on sector experience.

The identification of the agglomeration effect is based on the two step approach whereby the regional wage measures are regressed on regional population size in step two. In our setting the two step method is necessary to calculate the dynamic effects since we add the effect of learning to the regional fixed effects. When the determination of regional wages is understood in a migration equilibrium, population variables are potentially endogenous due to reverse causality and omitted production and consumer amenities. Migration of workers to regions with high wages will cause a spurious correlation between population size and the regional wage level and upward bias in OLS estimates. Omitted consumer amenities positively correlated with urban scale may bias estimates of agglomeration effects downward. In our understanding, these opposite bias factors probably explain why instrumentation has not produced very different estimates of the wage premium compared to OLS in the literature.

Ciccone and Hall (1996) innovated the handling of endogeneity by using lagged population variables as instruments. The instruments should be independent of present wage level and productivity. Long lags of population work well as instruments when they are important for the early urbanization, the urbanization process is persistent, and the background factors initiating the first urbanization are unimportant now. Glaeser et al. (2013) summarize the persistence for US counties and show that recent population sizes are closely correlated with numbers back to the mid 19^{th} century. Eaton and Eckstein (1997) confirm the same pattern for France and Japan. It is argued that large changes in production structure and techniques over such long periods of time imply that old population densities are less relevant today. Combes et al. (2010) extend the menu of instruments to include geographical and geological variables. We use historical census data to compute instruments for contemporary regional size and include first geography measured as share of mountains and lakes in the regional area. In a robustness extension we control for amenities representing school quality, crime, cultural services and climate.

We report our preferred first stage model in the first column of Appendix Table A1. Annual change of population size from 1875 to 1920 and share of mountains and lakes in total regional area are used as instruments for present population size (average for the period 2001-2010). The *F*-statistic is around 10 (weak instrument test suggested by Stock and Yogo, 2005) indicating that the instruments are jointly significant. The *LM* test clearly rejects the null hypothesis of underidentified model and confirms that our instruments are relevant. The endogeneity of the instruments are checked through the Hansen-J test, and the null hypothesis that instruments are uncorrelated with the error term is not rejected in any of the specifications. The effect of the change in population size from 1875 to 1920 is statistically significant at the 1% level and the first geography variable at 5% level. We have experimented with instrumentation based on historical population levels. These instruments have much better prediction power, but we want to avoid the linkage between historical and present population size levels. The results presented below hold with this alternative instrumentation.

Table 1 presents descriptive statistics for the main variables used in the individual and regional level regressions. The average worker in our dataset has an hourly wage of 238 NOK, is about 43 years old, and has 8 years of work experience. On average the experience made in a large city region is 3.6 years, in a high wage sector 0.7 years, and in high wage sector in large city about 0.4 years. About 12% of the workers are immigrants, and 72% are male. The average region has about 50 000 inhabitants, but there are large variations in population size across regions.

Table 1 about here

The top and bottom regions with respect to relative wages per hour are shown in Table 2. The regions with highest wage level include the capital Oslo and nearby regions, as well as the ‘oil capital’ Stavanger/ Sandnes and the second most populous region Bergen. The regions with the lowest wage level are all much smaller in population size and density and then relating to smaller cities and regional centers. The table illustrates our argument that population size is a better measure of urbanization than population density. The Stavanger/Sandnes and Bergen regions have more population than Bærum/Asker, but their density is much lower because they include a large sparsely populated territory outside the cities. The bottom regions in terms of wages and population size have extremely low population densities because they cover large unpopulated areas. The top regions have higher education levels, in particular Oslo and neighbor Bærum/Asker where around 50% of the population has tertiary education. The bottom regions have higher shares of the population with only primary education.

Table 2 about here

3. Static agglomeration effects across education groups

We start out presenting the raw agglomeration elasticity of hourly wages with respect to population size of 0.066 in column 1 of Table 3. A doubling of population size is associated with 6.6% higher wage level. To check out observable background factors we run an individual level regression over the whole sample including all individual characteristics, regional indicators and year and sector fixed effects, as described by equation (1) in section 2. The results are given in column 2. The education wage gap is 8% from primary to secondary education and 29.4% from primary to tertiary education. The male wage advantage is 17%.

Non-western immigrants have 9.2% lower wages on average. Experience matters and one year of experience adds 1.3% to wages. The effect is approximately linear. The agglomeration elasticity is reduced to 0.045 when the observables are accounted for. The size of this elasticitiy is consistent with the literature (see overviews of Puga, 2010 and Combes et al., 2011) and confirms that the urbanization effects in our dataset are consistent with previous results. The importance of unobserved characteristics has been a source of concern and only a few studies have been able to follow movers between regions to identify the ability factor. The individual regression in column 4 includes worker fixed effects (as described by equation (2) in section 2), and interestingly a large part of the education gap is explained by unobservables. The gap between primary and secondary educated is reduced to 1.8% and the gap between primary and tertiary education to 11.7%. The effect of a year of experience increases to 8.7% indicating a negative correlation between ability and experience. The agglomeration elasticity is 0.033 when observable and unobservable individual characteristics are included. The elasticity is down to about half of the raw elasticity.

Table 3 about here

Instrumentation of the population size effect increases the coefficients a bit. As shown in the second row of Table 3, the IV-estimate of the raw elasticity is up to 0.072 (from 0.066). When individual characteristics are taken into account the IV-elasticity is up to 0.046 (from 0.045) and with identification based on movers the IV-estimate is reduced to 0.031 (from 0.033). The OLS bias for this set of coefficients is negligible.

Our main interest is the separate effects of each education group. The static wage premiums with and without worker fixed effects are reported in Table 4. The first row presents OLS-estimates, while the second shows the IV models. The background individual regressions are shown in Appendix Tables A2 and A3. Taking away the ability factor adds to the experience effect for all education groups. The individual regressions show that the experience effect is higher for the highly educated. One extra year of experience increases wages of primary and secondary educated by about 7.6%, while the highly educated get a wage increase of 10.5%.

We are able to reproduce the differences of agglomeration elasticities across education groups obtained by Wheeler (2001) when controlling only for observables. The results are shown in the first three columns of the first row in Table 4. A doubling of the population size increases wages of primary educated workers by 3%, while individuals with higher education get about 6.2% wage increase. The difference between the estimated elasticity for the lower and higher education group is significantly different from zero at the 1% level. Wheeler did not account for endogeneity/ omitted variables or worker fixed effects. In the second row of columns 1-3 we show that the Wheeler result survives instrumentation. The IV-estimates confirm that the OLS bias is negligible, and the difference between the top and bottom education groups still is statistically significant. But Wheeler's conclusion does not hold when worker fixed effects are included, as shown in columns 4-6. In these models the agglomeration effect is the same for all three education groups and implies that a doubling of the population size increases wages by about 3% for all. The additional agglomeration effect for individuals with secondary and higher education shown above is related to unobserved abilities. When we take away the ability effect, the urban wage premium is the same and independent of education. Again the IV-estimates are close to the OLS-estimates, and the differences across education groups are not statistically significant.

Table 4 about here

Previous studies conclude that high education groups have higher static agglomeration effect. We have shown that the differences between education groups disappear when we control for unobservable individual characteristics. The urban wage premium does not increase with the education level in our dataset. We find that the size of the static agglomeration effect is similar across education groups, both individuals with lower and higher education benefit from agglomeration economies.

4. Dynamic agglomeration effects – experience by type of region and sector

The dynamics of agglomeration are related to the accumulation of experience. We separate between experience in large cities versus smaller regions and high wage sectors versus other sectors. Table 5 presents individual regression with worker fixed effects and with experience disaggregated based on where it is accumulated and currently used. The main result that stands out is the higher wage effect by having experience from large cities. One year experience in large cities adds an extra 1.1% to the wage level. This effect explains the dynamic wage premium below. In addition, the learning benefits from working in large cities are highly portable. Even if workers relocate to smaller regions, they still benefit from the large region experience. The interaction term between experience in large cities and currently working in a smaller region is insignificant. The education gap is similar to above, about 2% from primary to secondary and about 12% from primary to tertiary. The initial agglomeration effect, the elasticity of wages with respect to population size is 0.033 using OLS and 0.031 using IV, the same as before disaggregating experience (see column 5, Table 3).

The dynamic wage premium reflects the accumulation of experience over time, and in particular accounts for where the experience is accumulated. The calculation assumes a medium-term effect consisting of the initial effect and the average 8 years of work experience. The medium-term effect is higher than the initial effect, and about the same using OLS and IV (0.043 versus 0.048). The OLS estimate has some downward bias. The result is similar to De la Roca and Puga (2012, Table 2) and only with a slightly smaller additional effect of experience. In our dataset about 2/3 of the gain from working in a large city follows from the static effect and about 1/3 is dynamic.

Table 5 about here

We have shown in section 3 that the static agglomeration effect does not differ between education groups when worker fixed effects are included. The result is reproduced when experience is disaggregated by type of region and the estimated elasticities are reported in columns 1-3 of Table 6, with OLS and IV estimates in the first and second row, respectively. The initial urban wage premium controlling for observables and unobservables following a doubling of the population size is about 3% and independent of education. However, the dynamic wage premium differs according to education, as can be seen from columns 4-6 of Table 6. In the OLS models, the medium-term effect of doubling the population size is 3.3% higher wages for primary educated, 3.7% for secondary educated and 4.1% for highly educated. The IV estimates in the second row indicates that the small downward bias in the OLS estimates in Table 5 is entirely driven by tertiary educated workers. While the OLS coefficients are upward biased for primary and secondary educated, there is downward bias for highly educated. We will discuss this below when we take into account industrial structure. The effect in our context is to increase the difference between primary and secondary educated on the one hand and the tertiary educated on the other. The medium-term premium is 4.7% for highly educated compared to 3% for primary educated, and the difference is statistically significant at the 5% level.

The point estimates and the background individual regressions indicate that experience in the large cities is more valuable for the highly educated. As seen from Appendix Table A4, one year of experience in the large cities increases wages of highly educated by 1.1%, while wages of primary and secondary educated only increase by 0.3% and 0.5%, respectively. The combination of the experience effects discussed above and these estimates lead us to conclude that the highest educated gain more from living in cities because they have more return to experience. The learning effect from working in large cities documented in Table 5 tends to increase with education.

Table 6 about here

Regions differ with respect to industrial structure. So far we have estimated the agglomeration effect independent of how education groups allocate to sectors. The importance of industrial sector is investigated by separating out the top 10 high wage sectors based on the fixed sectoral effects of the individual models reported in Appendix Table A4.[4] In Table 7 the experience effect is distinguished between large cities and other regions, high wage sectors and the rest, and high wage sectors in large cities (as described by equation (3) in section 2). In addition portability is analyzed for a separate group with experience in large cities, but now working in a smaller region. Comparing the first columns of Tables 5 and 7 we see that the size of the effect of experience, the additional effect of experience in large cities, and the effects of secondary and tertiary education are unaffected. The new result is that industrial sector matters. A year of experience in a high wage sector adds 0.5% to the wage, and experience in a high wage sector in a large city adds another 0.3%. A worker gains 0.8% by having a year of experience in a high wage sector in a large city compared to experience in another sector outside the large cities. The initial premium is the same as in Table 5. We can now calculate the medium-term premium dependent of sector of experience. The medium-term gain of doubling the population size is 5.3% for workers in high wage sectors and 4.6% for workers in low wage sectors. The aggregate agglomeration effect depends on regional industrial structure.

Table 7 about here

[4] The top 10 high wage sectors are specific for each education group, but differences across groups are minor. High wage sectors are typically oil related sectors, business services and transport sectors.

The introduction of the industrial sector effect of experience modifies the dynamic wage premium across education groups. Table 8 displays the medium-term premium for each education group, when the experience is made in the 10 high wage sectors (columns 4-6) and when experience is made in the other sectors (columns 1-3). The first row shows OLS-estimates and the second row is based on IV. The individual regressions underlying these results are shown in Appendix Table A5. The average effect of a year of experience is still higher for the highly educated, 9.7% compared to 7.5% for those with only primary education and 7.4% for those with secondary education. The new results (comparing Appendix Tables A4 and A5) are the additional effects of experience in high wage sectors and in particular for high wage sectors in large cities. The extra wage effect of experience in high wage sectors in large cities is only statistically significant for primary and tertiary educated, 0.4% and 0.6% respectively. The estimates show that secondary educated take benefit of experience in high wage sectors independent of what type of region they work in. Primary educated only benefit from working in high wage sectors in large cities. It is of interest to pursue these differences in future research. At this stage we only note that the additional dynamic agglomeration effect in high wage sectors is higher for primary educated and tertiary educated.

The different estimates of the individual regressions for each education group explain the elasticities shown in columns 4-6 of Table 8 – primary educated have higher dynamic premium in high wage sectors than secondary educated, but tertiary educated still have the highest wage premium. Our main finding is that the learning effect depends on both the level of education and sector of employment. The differences between highly educated and primary educated in the same type of sector are statistically significant at the 10% level. The dynamic wage premium for a person with high education working in a high wage sector is about double compared to a person with low education outside the high wage sectors, and the difference is significant at the 1% level. Interestingly, the downward OLS bias is quite large for highly educated. Our understanding is that highly educated have broader motivation to move to large cities and that the consequent crowding has negative effects for productivity. It is of interest to pursue this issue.

Table 8 about here

5. Robustness check

Given the volume of data involved here, there are many ways of specifying the models estimated. The analysis concentrates on the particular importance of large cities and high wage sectors. We have studied alternative definitions of the most urban regions, and the urbanization effects hold when we extend beyond the large cities although the size of the effects is somewhat smaller. In the analysis of industrial sector effects we have also looked at an alternative where the most advanced sectors are defined by share of workers with tertiary education. Both the wage premium of the advanced sectors and the interaction effect with large cities remain. We have also studied the heterogeneity of the population with respect to gender and race. The results are robust when we study men only and when we exclude foreign immigrants.

A major concern in the estimation of population size effects is the role of amenities motivating migration. Four types of amenities have been shown to be important in the literature – school quality, cultural services, crime, and climate. We have checked the robustness of the results with various combinations of these control variables in Table 9 using the IV method. The measure of school quality is based on Borge and Naper (2006). They have estimated municipal fixed effects based on individual data of student achievement in mathematics and with other relevant controls. The weighted municipal effects are aggregated to regional school quality. Cultural amenities are measured as net per capita regional spending on museums in the year 2010. Public safety is measured by number of drug related crimes per 1000 inhabitant and as an average over the period 1994-2001. Finally, climate is represented by the average winter temperature during 1971-2000. Comparing the estimates including amenity controls of Table 9 with the results of Table 8, we see that the dynamic agglomeration effects are somewhat reduced across the board, but that the differences between education groups and between high wage sectors and the rest remain. The highly educated have the highest dynamic agglomeration effect, and the effect is more than three times larger for highly educated in high wage sectors versus primary educated outside high wage sectors. The amenity variables have the expected signs. In particular, crime has a positive and significant effect on regional wages indicating that workers in areas with high extent of crime are compensated with higher wages. Since crime is positively correlated with urban scale this is the main reason behind the drop in agglomeration elasticities. We have looked at alternative measures of the four types of amenities and always reproduce our main conclusions.

Table 9 about here

The regions in our analysis are defined based on commuting patterns, but input-output linkages might extend beyond the regional boundaries. To take into account that some agglomeration effects can take place at a larger scale than the labor market region, we control for a region's market potential. Market potential is defined as the sum of the population size of a region's neighboring regions, weighted by the inverse distance between the regions.[5] Table 10 shows the estimates of the dynamic wage premium when both regional population size and market potential are included as independent variables. The first stage regressions for population size and market potential for the IV-estimates are documented in Appendix Table A1. The addition of the market potential variable takes away some effect from population size, but the main findings are still the same. The medium-term premium increases with the level of education, and tertiary educated workers in high wage sectors gain most from experience in large cities. The static wage premium is similar across education groups (not reported in the table).

Table 10 about here

6. Concluding remarks

We have used register data for all full time workers in the private sector in Norway (about 6.5 million worker-year observations) to study the agglomeration wage premium. The individual panel data include observations of education levels as well as personal, labor market and employment sector characteristics, and also allow for identification of unobserved individual effects based on migration between regions. The main focus is the analysis of differences in population size effects for the income level across education groups. The endogeneity of the population measures and omitted variable bias is addressed using instrument variables based on historical population structure and first geography. We do not know other studies of the agglomeration effect in education groups with this instrumentation of population variables and taking into account individual unobserved effects and experience by region and sector.

[5] Regions are defined as neighbors if the distance between regional centers is less than 200 km.

The data allows for a separation between static and dynamic agglomeration effects taking into account the location of the work experience of the individuals. The experience is distinguished between large cities (more than 150 000 inhabitants) and smaller regions. Also the experience with respect to industrial sector is taken into account, distinguishing between the top 10 high wage sectors and the rest. The initial wage premium is not affected by the inclusion of worker experience history, but the experience effect adds to the medium-term wage effect since experience in large regions is found to be more valuable. The initial premium is the same across education groups, but the experience effect differs with respect to education, in particular for the highly educated in high wage sectors.

The estimation has raised two issues worth pursuing in future research; how different education groups gain from experience and the source of possible downward bias of OLS estimates of dynamic effects for highly educated.

References

Ahlin, L., M. Andersson and P. Thulin (2013), Market thickness and the early labor market career of university graduates: An urban advantage? CIRCLE Working Paper No. 2013/2, Lund University.

Andersson, M., J. Klaesson and J.P. Larsson (2013), The sources of urban wage premium by worker skills: Spatial sorting or agglomeration economies? *Papers in Regional Science*, forthcoming.

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

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

Black, D. and J.V. Henderson (2003), Urban evolution in the USA, *Journal of Economic Geography* 3, 343-372.

Borge, L. and L. Naper (2006), Efficiency potential and efficiency variation in Norwegian lower secondary schools, *FinanzArchiv* 62, 221-249.

Ciccone A. and R. Hall (1996), Productivity and the density of economic activity, *American Economic Review* 86, 1, 54-70.

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

Combes P-P., G. Duranton, L. Gobillon, and S. Roux (2010), Estimating agglomeration economies with history, geology and worker effects, in E. Glaeser (ed), *Agglomeration Economics*, Chicago: Chicago University Press.

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

D'Costa, S. and H. Overman (2013), The urban wage growth premium: Sorting or learning? SERC Discussion Paper 135, London School of Economics.

De la Roca, J. (2011), Selection in initial and return migration: Evidence from moves across Spanish cities, IMDEA Working Paper No. 2011-21.

De la Roca, J. and D. Puga (2012), Learning by working in big cities, CEPR Discussion Paper No. DP9243.

Eaton, J. and Z. Eckstein (1997), Cities and growth: Theory and evidence from France and Japan, *Regional Science and Urban Economics* 27, 443-474.

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

Glaeser, E., G. Ponzetto and K. Tobio (2013), Cities, skills and regional change, *Regional Studies*, forthcoming.

Glaeser, E. and M. Resseger (2010), The complementarity between cities and skills, *Journal of Regional Science* 50, 1, 221-244.

Glaeser E.L. and A, Saiz (2004), The rise of the skilled city, *Brookings-Wharton Papers on Urban Affairs* 5, 47-94.

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

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

Liu, K., K.G Salvanes and E. Sørensen (2012), Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession, NHH Discussion Paper SAM 16 2012.

Puga, D. (2010), The magnitude and causes of agglomeration economies, *Journal of Regional Science* 50, 1, 203-219.

Rosenthal, S. and W. Strange (2008), The attenuation of human capital spillovers, *Journal of Urban Economics* 64, 373-389.

Shapiro, J. (2006), Smart cities – quality of life, productivity, and the growth effects of human capital, *Review of Economics and Statistics* 88, 324-335.

Stock J. H. and M. Yogo (2005), Testing for weak instruments in linear IV regression, in D. W. K. Andrews and J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, 80-108, Cambridge: Cambridge University Press.

Wang, Z. (2013), Smart city: Learning effects and labor force entry, mimeo, Department of Economics, Brown University.

Wheeler, C. (2001), Search, sorting and urban agglomeration, *Journal of Labor Economics* 19, 4, 879-899.

Winters J. (2011), Why are smart cities growing? Who moves and who stays, *Journal of Regional Science* 51, 2, 253-270.

Table 1: Descriptive statistics

Panel a

	Mean	St dev	Min	Max
Individual level data				
Hourly wage (in NOK)	238	130.8	70	1250
Total work experience (in years)	7.9	4.2	0	17.6
Experience in large cities (in years)	3.6	4.7	0	17.6
Experience in high wage sectors (in years)	0.7	2.3	0	17.3
Experience in high wage sectors in large cities (in years)	0.4	1.7	0	17.3
Age	42.5	10.5	25	65
Regional level data				
Regional indicators eq. (3)	5.27	0.05	5.15	5.37
Regional indicators eq. (3), primary	5.29	0.06	5.15	5.37
Regional indicators eq. (3), secondary	5.30	0.06	5.15	5.39
Regional indicators eq. (3), tertiary	5.40	0.05	5.28	5.50
Regional population size	52 220	77 805	5 442	540 030
ΔPop size$_{1875\text{-}1920}$	208.8	531.8	-57.5	4538.4
Share mountain and lake	0.08	0.05	0.01	0.26
Average winter temperature (Celsius)	-3.2	3.7	-12.8	3.0
School quality (math grade)	3.5	0.1	3.2	3.8
Drug crime (per 1000 inhabitants)	6.3	3.1	2.5	13.6
Public expenditure museums	71.0	59.3	-19.0	344.9

Panel b

	Share of obs.
Primary education	0.193
Secondary education	0.529
Tertiary education	0.278
Male	0.717
Immigrant	0.117
Immigrant, western	0.087
Immigrant, non-western	0.029

Notes: Work experience is calculated in days from 1993 onwards, and expressed in years. We separate between large city regions and the rest. The large city group is defined as regions with more than 150 000 inhabitants in 2010, which includes 7 regions. We also separate out the top 10 high wage sectors based on the fixed sectoral effects from individual level regressions. Secondary education corresponds to workers that have completed at least one year of secondary education, while tertiary education includes workers with at least one year at university/college. Western immigrants are defined as immigrants from Europe, Japan, North America, Australia or New Zealand. Regional population size is given as an average during 2001-2010. ΔPop size$_{1875\text{-}1920}$ represents the annual change in regional population size from 1875 to 1920, while 'Share mountain and lake' is the share of the regional area covered by mountains or lakes. The average winter temperature is given in Celsius degrees and is the average during 1971-2000. The measure of school quality is based on student performance in mathematics adjusted for student and family characteristics (estimated by Borge and Naper, 2006), and is given on a scale from 1 to 6 with 6 as the best. The number of drug related crimes per 1000 inhabitants is measured as an average during 1994-2001, while net per capita public expenditures on museums is from the year 2010.

Table 2: Summary statistics, top and bottom regions ranked by the average hourly wage level during 2001-2010.

	Relative hourly wage	Population size	Population density	Primary education	Secondary education	Tertiary education
High wage regions						
Bærum/Asker	1.45	157 173	536.4	9.4	36.4	54.2
Stavanger/Sandnes	1.28	234 595	70.0	16.6	51.9	31.5
Oslo	1.24	540 030	1189.5	13.8	36.8	49.4
Follo	1.22	108 688	185.2	16.6	48.3	35.0
Bergen	1.14	368 789	72.2	18.0	52.4	29.5
Low wage regions						
Nord-Troms	0.85	11 363	1.6	29.1	60.6	10.3
Røros	0.85	7 736	2.4	15.8	67.7	16.5
Tynset	0.84	15 511	1.6	18.2	64.2	17.5
Nord-Gudbrandsdalen	0.83	19 699	2.0	22.9	65.9	11.2
Vadsø	0.82	15 748	1.6	31.3	52.2	16.5

Notes: The first column gives the regional level of hourly wages before adjusting for observable and unobservable individual characteristics and before adding learning effects, measured relative to the average wage level across all 89 regions. Population density is measured as inhabitants per square kilometer. The last three columns give the share of workers with primary, secondary and tertiary education, respectively.

Table 3: Estimation of the static urban wage premium

Dependent variable	(1) Average regional hourly wage (log)	(2) Log hourly wage	(3) Regional indicator coefficients column (2)	(4) Log hourly wage	(5) Regional indicator coefficients column (4)
OLS estimation:					
Regional pop size (log)	0.066***		0.045***		0.033***
	(0.01)		(0.006)		(0.004)
IV estimation:					
Regional pop size (log)	0.072***		0.046***		0.031***
	(0.008)		(0.006)		(0.004)
Regional indicators		Yes		Yes	
Worker fixed effects		No		Yes	
Year fixed effects		Yes		Yes	
Sector fixed effects		Yes		Yes	
Age controls		Yes		Yes	
Experience		0.013***		0.087***	
		(0.0001)		(0.0003)	
$(\text{Experience})^2$		-0.000***		-0.001***	
		(0.0000)		(0.0000)	
Secondary education		0.08***		0.018***	
		(0.0004)		(0.0019)	
Tertiary education		0.294***		0.117***	
		(0.0004)		(0.0029)	
Immigrant, western		-0.011***			
		(0.0005)			
Immigrant, non-western		-0.092***			
		(0.0008)			
Male		0.17***			
		(0.0003)			
Observations	89	6.512.359	89	6.512.359	89

Notes: The regressions in columns (2) and (4) are based on yearly data for all full time workers in the private sector during 2001-2010. Sector fixed effects are at the 2-digit level and include 54 sectors. Regional indicators are at the NUTS-4 level, and correspond to 89 labor market regions. The age controls are given as 5-year intervals. Work experience is calculated in days from 1993 onwards, and expressed in years. Secondary education corresponds to workers that have completed at least one year of secondary education, while tertiary education includes workers with at least one year at university/college. Western immigrants are defined as immigrants from Europe, Japan, North America, Australia or New Zealand. The regressions in columns (1), (3) and (5) are at the regional level, and are done with both OLS and IV estimation, as illustrated in the first and second row, respectively. The population size during 2001-2010 is instrumented by the annual change in population size during 1875-1920 and the share of regional area consisting of mountains and lakes. The first stage estimation is documented in Appendix Table A1. Standard errors are given in parenthesis. *** indicates significance at the 1 percent level. All regressions include a constant term.

Table 4: Static urban wage premium by education groups

Dependent variables	Regional indicator coefficients from individual regressions by education groups, with or without worker fixed effects					
	(1)	(2)	(3)	(4)	(5)	(6)
Education group	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary
Worker fixed effects	No	No	No	Yes	Yes	Yes
OLS estimation:						
Regional pop size	0.03***	0.044***	0.062***	0.03***	0.033***	0.031***
(log)	(0.005)	(0.006)	(0.007)	(0.005)	(0.004)	(0.004)
IV estimation:						
Regional pop size	0.029***	0.043***	0.065***	0.025***	0.029***	0.032***
(log)	(0.005)	(0.006)	(0.007)	(0.004)	(0.003)	(0.004)
Observations	89	89	89	89	89	89

Notes: In columns (1) – (3) the dependent variable is regional indicator coefficients by education groups from individual level regressions controlling for observable individual characteristics, as well as sector and year indicators, as documented in columns (1) – (3) of Appendix Table A2. In columns (4) – (6) the dependent variable is regional indicator coefficients by education groups from individual level regressions controlling for both observable and unobservable individual characteristics, as well as sector and year indicators, as documented in columns (1) – (3) of Appendix Table A3. The regressions are done with both OLS and IV estimation, as illustrated in the first and second row, respectively. The population size during 2001-2010 is instrumented by the annual change in population size during 1875-1920 and the share of regional area consisting of mountains and lakes. The first stage estimation is documented in Appendix Table A1. Clustered standard errors (at the regional level) are given in parentheses. *** indicates significance at the 1 percent level. All regressions include a constant term.

Table 5: Estimation of the dynamic urban wage premium – controlling for experience by type of region

Dependent variable	(1) Log hourly wage	(2) Initial premium (regional indicator coefficients column (1))	(3) Medium-term premium (initial + 8 years local experience)
OLS estimation:			
Regional pop size (log)		0.033***	0.043***
		(0.004)	(0.005)
IV estimation:			
Regional pop size (log)		0.031***	0.048***
		(0.004)	(0.004)
Regional indicators	Yes		
Worker fixed effects	Yes		
Year fixed effects	Yes		
Sector fixed effects	Yes		
Age controls	Yes		
Experience	0.081***		
	(0.0003)		
$(\text{Experience})^2$	-0.001***		
	(0.0000)		
Experience large cities	0.011***		
	(0.0002)		
$(\text{Experience large cities})^2$	-0.000***		
	(0.0000)		
Experience large cities x now in smaller	-0.000		
	(0.0002)		
Secondary education	0.02***		
	(0.0019)		
Tertiary education	0.119***		
	(0.0029)		
Observations	6.512.359	89	89

Notes: The regression in column (1) is based on yearly data for all full time workers in the private sector during 2001-2010. Sector fixed effects are at the 2-digit level and include 54 sectors. Regional indicators are at the NUTS-4 level, and correspond to 89 labor market regions. The age controls are given as 5-year intervals. Work experience is calculated in days from 1993 onwards, and expressed in years. We separate between large city regions and the rest. The large city group is defined as regions with more than 150 000 inhabitants in 2010, which includes 7 regions. Secondary education corresponds to workers that have completed at least one year of secondary education, while tertiary education includes workers with at least one year at university/college. The dependent variable in column (2) is the regional indicator coefficients from the regression in column (1). In column (3) the dependent variable also includes the medium-term urban wage premium. This is calculated by adding the wage effect of the average experience in a region (about 8 years) to the initial premium, based on the estimated coefficients in column (1). The regressions in columns (2) and (3) are done with both OLS and IV estimation, as illustrated in the first and second row, respectively. The population size during 2001-2010 is instrumented by the annual change in population size during 1875-1920 and the share of regional area consisting of mountains and lakes. The first stage estimation is documented in Appendix Table A1. Standard errors are given in parenthesis. *** indicates significance at the 1 percent level. All regressions include a constant term.

Table 6: Static and dynamic urban wage premium by education groups – controlling for experience by type of region

Dependent variable	(1) Initial premium	(2) Initial premium	(3) Initial premium	(4) Medium-term premium	(5) Medium-term premium	(6) Medium-term premium
Education group	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary
OLS estimation:						
Regional pop size (log)	0.03***	0.032***	0.031***	0.033***	0.037***	0.041***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)
IV estimation:						
Regional pop size (log)	0.025***	0.027***	0.032***	0.03***	0.035***	0.047***
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	89	89	89	89	89	89

Notes: In columns (1) – (3) the dependent variable is regional indicator coefficients by education groups from individual level regressions controlling for both observable and unobservable individual characteristics, as well as sector and year indicators, as documented in Appendix Table A4. In columns (4) – (6) the dependent variable also includes the medium-term urban wage premium. This is calculated by adding the wage effect of the average experience in a region (about 8 years) to the initial premium, based on the estimated coefficients in Appendix Table A4. The regressions are done with both OLS and IV estimation, as illustrated in the first and second row, respectively. The population size during 2001-2010 is instrumented by the annual change in population size during 1875-1920 and the share of regional area consisting of mountains and lakes. The first stage estimation is documented in Appendix Table A1. Clustered standard errors (at the regional level) are given in parentheses. *** indicates significance at the 1 percent level. All regressions include a constant term.

Table 7: Estimation of the dynamic urban wage premium – controlling for experience by type of region and sector

	(1)	(2)	(3)	(4)
Dependent variable	Log hourly wage	Initial premium (regional indicator coefficients column (1))	Medium-term premium (initial + 8 years local experience)	Medium-term premium (initial + 8 years local experience)
High wage sector			No	Yes
OLS estimation:				
Regional pop size (log)		0.033***	0.043***	0.047***
		(0.004)	(0.005)	(0.005)
IV estimation:				
Regional pop size (log)		0.031***	0.046***	0.053***
		(0.004)	(0.004)	(0.005)
Regional indicators	Yes			
Worker fixed effects	Yes			
Year fixed effects	Yes			
Sector fixed effects	Yes			
High wage sector x Year fixed effects	Yes			
Age controls	Yes			
Experience	0.08***			
	(0.0003)			
$(\text{Experience})^2$	-0.001***			
	(0.0000)			
Experience large cities	0.011***			
	(0.0002)			
$(\text{Experience large cities})^2$	-0.000***			
	(0.0000)			
Experience large cities x now in smaller	-0.000			
	(0.0002)			
Experience high wage sector	0.005***			
	(0.0004)			
$(\text{Experience high wage sector})^2$	-0.000***			
	(0.0000)			
Experience high wage sector in large cities	0.003***			
	(0.0003)			
Secondary education	0.021***			
	(0.0019)			
Tertiary education	0.119***			
	(0.0029)			
Observations	6.512.359	89	89	89

Notes: The regression in column (1) is based on yearly data for all full time workers in the private sector during 2001-2010. Sector fixed effects are at the 2-digit level and include 54 sectors. Regional indicators are at the NUTS-4 level, and correspond to 89 labor market regions. The age controls are given as 5-year intervals. Work experience is calculated in days from 1993 onwards, and expressed in years. We separate between large city regions and the rest. The large city group is defined as regions with more than 150 000 inhabitants in 2010, which includes 7 regions. We also separate out the top 10 high wage sectors based on the fixed sectoral effects of the individual model reported in Table 5. The dependent variable in column (2) is the regional indicator coefficients from the regression in column (1). In columns (3) and (4) the dependent variable also includes the medium-term urban wage premium, dependent on type of sector. This is calculated by adding the wage effect of the average experience in a region (about 8 years) to the initial premium, based on the estimated coefficients in column (1). The regressions in columns (2) – (4) are done with both OLS and IV estimation, as illustrated in the first and second row, respectively. The population size during 2001-2010 is instrumented by the annual change in population size during 1875-1920 and the share of regional area consisting of mountains and lakes. The first stage estimation is documented in Appendix Table A1. Standard errors are given in parenthesis. *** indicates significance at the 1 percent level. All regressions include a constant term.

Table 8: Dynamic urban wage premium by education groups and type of sector

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Medium-term premium	Medium-term premium	Medium-term premium	Medium-term premium	Medium-term premium	Medium-term premium
Education group	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary
High wage sector	No	No	No	Yes	Yes	Yes
OLS estimation:						
Regional pop size	0.033***	0.036***	0.039***	0.039***	0.036***	0.047***
(log)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.006)
IV estimation:						
Regional pop size	0.03***	0.035***	0.045***	0.039***	0.035***	0.057***
(log)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Observations	89	89	89	89	89	89

Notes: The dependent variables capture the medium-term urban wage premium by education groups and type of sector. It is calculated by adding the wage effect of the average experience in a region (about 8 years) to the initial premium given by the regional indicator coefficients in the individual level regressions in Appendix Table A5. The regressions are done with both OLS and IV estimation, as illustrated in the first and second row, respectively. The population size during 2001-2010 is instrumented by the annual change in population size during 1875-1920 and the share of regional area consisting of mountains and lakes. The first stage estimation is documented in Appendix Table A1. Clustered standard errors (at the regional level) are given in parentheses. *** indicates significance at the 1 percent level. All regressions include a constant term.

Table 9: Dynamic urban wage premium by education group and type of sector – with control variables (IV estimation only)

Dependent variable	(1) Medium-term premium	(2) Medium-term premium	(3) Medium-term premium	(4) Medium-term premium	(5) Medium-term premium
Effect of regional pop size (log)by education group and type of sector:					
Primary, low wage sector	0.016**	0.019***	0.015**	0.027***	0.016**
	(0.008)	(0.006)	(0.007)	(0.007)	(0.007)
Secondary, low wage sector	0.024***	0.026***	0.024***	0.032***	0.025***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Tertiary, low wage sector	0.037***	0.039***	0.037***	0.042***	0.038***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)
Primary, high wage sector	0.027***	0.029***	0.026***	0.037***	0.027***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Secondary, high wage sector	0.024***	0.027***	0.024***	0.032***	0.025***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Tertiary, high wage sector	0.051***	0.053***	0.051***	0.055***	0.052***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
Control variables					
Average winter temperature	Yes	No	Yes	Yes	Yes
School quality (math grade)	Yes	Yes	No	Yes	Yes
Drug crime (per 1000 inhab.)	Yes	Yes	Yes	No	Yes
Public expenditure museums	Yes	Yes	Yes	Yes	No
Observations	89	89	89	89	89
IV estimation tests					
F-statistic weak identification (H_0: instruments jointly insign.)	7.8	9.0	7.2	10.0	7.3
p-value LM test (H_0: model underidentified)	0.05	0.02	0.04	0.06	0.03
p-value J-test (H_0: instruments uncorrelated with error term)	0.34 – 0.99	0.17 – 0.66	0.35 – 0.99	0.27 – 0.91	0.26 – 0.86

Notes: The dependent variables are the same as in Table 8. Each column shows the estimated effect of regional population size on the medium-term urban wage premium for six groups dependent on level of education and type of sector. The columns differ with respect to control variables included. The regressions are based on IV estimation. The population size during 2001-2010 is instrumented by the annual change in population size during 1875-1920 and the share of regional area consisting of mountains and lakes. The first stage estimation for the specification with all four control variables is documented in Appendix Table A1. Clustered standard errors (at the regional level) are given in parentheses. *** and ** indicate significance at the 1 and 5 percent levels, respectively. All regressions include a constant term.

Table 10: Dynamic urban wage premium by education groups and type of sector – including market potential

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Medium-term premium	Medium-term premium	Medium-term premium	Medium-term premium	Medium-term premium	Medium-term premium
Education group	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary
High wage sector	No	No	No	Yes	Yes	Yes
OLS estimation:						
Regional pop size (log)	0.019***	0.023***	0.028***	0.025***	0.023***	0.036***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.006)
Market potential (log)	0.027***	0.026***	0.021***	0.027***	0.026***	0.021***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
IV estimation:						
Regional pop size (log)	0.018***	0.023***	0.034***	0.028***	0.023***	0.047***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)
Market potential (log)	0.023***	0.022***	0.02***	0.022***	0.022***	0.019***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)
Observations	89	89	89	89	89	89

Notes: The dependent variables are the same as in Table 8. Market potential is measured as the sum of the population size of a region's neighboring regions, weighted by the inverse distance between the regions. The regressions are done with both OLS and IV estimation, as illustrated in the upper and lower half of the table, respectively. The population size and market potential during 2001-2010 are instrumented by the annual change in population size during 1875-1920, the change in market potential during 1875-1920, and the share of regional area consisting of mountains and lakes. The first stage estimations are documented in Appendix Table A1. Clustered standard errors (at the regional level) are given in parentheses. *** indicates significance at the 1 percent level. All regressions include a constant term.

Appendix Table A1: IV estimation, first stage regressions

	(1) Regional pop size (log)	(2) Regional pop size (log)	(3) Regional pop size (log)	(4) Regional market potential (log)
ΔPop size$_{1875\text{-}1920}$	0.001***	0.001***	0.001***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
ΔMarket potential$_{1875\text{-}1920}$			0.092***	0.326***
			(0.019)	(0.06)
Share mountain and lake	-3.097**	-1.39	-1.811	-2.27
	(1.197)	(1.133)	(1.103)	(1.674)
Average winter temperature		0.035*		
		(0.019)		
School quality (math grade)		-1.408**		
		(0.677)		
Drug crime (per 1000 inhab.)		0.054*		
		(0.03)		
Public expenditure museums		-0.004***		
		(0.001)		
Observations	89	89	89	89
R^2	0.40	0.55	0.49	0.52
F-statistic weak identification	10.6	7.8	17.0	12.5
(H_0: instruments jointly insign.)				
p-value LM test	0.01	0.05	0.01	0.01
(H_0: model underidentified)				

Notes: Clustered standard errors (at the regional level) are given in parentheses. ***, ** and * indicate significance at the 1, 5 and 10 percent levels. All regressions include a constant term.

Appendix Table A2: Individual level regressions by education groups – without worker fixed effects and past/current experience by type of region/sector

	(1)	(2)	(3)
Dependent variable	Log hourly wage	Log hourly wage	Log hourly wage
Education group	Primary	Secondary	Tertiary
Regional indicators	Yes	Yes	Yes
Worker fixed effects	No	No	No
Year fixed effects	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes
Age controls	Yes	Yes	Yes
Experience	0.004***	0.005***	0.022***
	(0.0003)	(0.0002)	(0.0003)
$(\text{Experience})^2$	0.000***	0.000***	-0.000***
	(0.0000)	(0.0000)	(0.0000)
Immigrant, western	-0.005***	-0.012***	-0.019***
	(0.0011)	(0.0007)	(0.0009)
Immigrant, non-western	-0.062***	-0.088***	-0.135***
	(0.0015)	(0.0013)	(0.0016)
Male	0.141***	0.178***	0.158***
	(0.0008)	(0.0005)	(0.0006)
Observations	1.255.194	3.448.571	1.808.594
R^2	0.20	0.27	0.30

Notes: The regressions are based on yearly data for all full time workers by education groups during 2001-2010. Workers in primary and public sectors are excluded. Secondary education corresponds to workers that have completed at least one year of secondary education, while tertiary education includes workers with at least one year at university/college. Sector fixed effects are at the 2-digit level and include 54 sectors. Regional indicators are at the NUTS-4 level, and correspond to 89 labor market regions. The age controls are given as 5-year intervals. Work experience is calculated in days from 1993 onwards, and expressed in years. Western immigrants are defined as immigrants from Europe, Japan, North America, Australia or New Zealand. Standard errors are given in parenthesis. *** indicates significance at the 1 percent level. All regressions include a constant term.

Appendix Table A3: Individual level regressions by education groups – with worker fixed effects, but without past/current experience by type of region/sector

	(1)	(2)	(3)
Dependent variable	Log hourly wage	Log hourly wage	Log hourly wage
Education group	Primary	Secondary	Tertiary
Regional indicators	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes
Age controls	Yes	Yes	Yes
Experience	0.076***	0.077***	0.105***
	(0.0006)	(0.0004)	(0.0005)
$(Experience)^2$	-0.001***	-0.000***	-0.001***
	(0.0000)	(0.0000)	(0.0000)
Observations	1.255.194	3.448.571	1.808.594
R^2	0.65	0.72	0.77

Notes: The regressions are based on yearly data for all full time workers by education groups during 2001-2010. Workers in primary and public sectors are excluded. Secondary education corresponds to workers that have completed at least one year of secondary education, while tertiary education includes workers with at least one year at university/college. Sector fixed effects are at the 2-digit level and include 54 sectors. Regional indicators are at the NUTS-4 level, and correspond to 89 labor market regions. The age controls are given as 5-year intervals. Work experience is calculated in days from 1993 onwards, and expressed in years. Standard errors are given in parenthesis. *** indicates significance at the 1 percent level. All regressions include a constant term.

Appendix Table A4: Individual level regressions by education groups – with worker fixed effects and past/current experience by type of region

	(1)	(2)	(3)
Dependent variable	Log hourly wage	Log hourly wage	Log hourly wage
Education group	Primary	Secondary	Tertiary
Regional indicators	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes
Age controls	Yes	Yes	Yes
Experience	0.075***	0.075***	0.098***
	(0.0007)	(0.0004)	(0.0006)
$(\text{Experience})^2$	-0.001***	-0.000***	-0.001***
	(0.0000)	(0.0000)	(0.0000)
Experience large cities	0.003***	0.005***	0.011***
	(0.0006)	(0.0003)	(0.0004)
$(\text{Experience large cities})^2$	-0.000**	-0.000***	-0.000***
	(0.0000)	(0.0000)	(0.0000)
Experience large cities x now	-0.000	-0.002***	-0.001***
in smaller	(0.0005)	(0.0003)	(0.0003)
Observations	1.255.194	3.448.571	1.808.594
R^2	0.65	0.72	0.77

Notes: The regressions are based on yearly data for all full time workers by education groups during 2001-2010. Workers in primary and public sectors are excluded. Secondary education corresponds to workers that have completed at least one year of secondary education, while tertiary education includes workers with at least one year at university/college. Sector fixed effects are at the 2-digit level and include 54 sectors. Regional indicators are at the NUTS-4 level, and correspond to 89 labor market regions. The age controls are given as 5-year intervals. Work experience is calculated in days from 1993 onwards, and expressed in years. We separate between large city regions and the rest. The large city group is defined as regions with more than 150 000 inhabitants in 2010, which includes 7 regions. Standard errors are given in parenthesis. ***, ** indicate significance at the 1 and 5 levels. All regressions include a constant term.

Appendix Table A5: Individual level regressions by education groups – with worker fixed effects and past/current experience by type of region and sector

	(1)	(2)	(3)
Dependent variable	Log hourly wage	Log hourly wage	Log hourly wage
Education group	Primary	Secondary	Tertiary
Regional indicators	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes
High wage sector x Year fixed effects	Yes	Yes	Yes
Age controls	Yes	Yes	Yes
Experience	0.075***	0.074***	0.097***
	(0.0007)	(0.0004)	(0.0006)
$(\text{Experience})^2$	-0.001***	-0.000***	-0.001***
	(0.0000)	(0.0000)	(0.0000)
Experience large cities	0.003***	0.005***	0.01***
	(0.0006)	(0.0003)	(0.0004)
$(\text{Experience large cities})^2$	-0.000***	-0.000***	-0.000***
	(0.0000)	(0.0000)	(0.0000)
Experience large cities x now in smaller	-0.000	-0.002***	-0.001***
	(0.0005)	(0.0003)	(0.0003)
Experience high wage sector	0.001	0.009***	0.003***
	(0.0016)	(0.0006)	(0.0008)
$(\text{Experience high wage sector})^2$	0.000**	-0.000***	-0.000***
	(0.0001)	(0.0000)	(0.0000)
Experience high wage sector in large cities	0.004***	-0.000	0.006***
	(0.001)	(0.0004)	(0.0006)
Observations	1.255.194	3.448.571	1.808.594
R^2	0.65	0.72	0.77

Notes: The regressions are based on yearly data for all full time workers by education groups during 2001-2010. Workers in primary and public sectors are excluded. Secondary education corresponds to workers that have completed at least one year of secondary education, while tertiary education includes workers with at least one year at university/college. Sector fixed effects are at the 2-digit level and include 54 sectors. Regional indicators are at the NUTS-4 level, and correspond to 89 labor market regions. The age controls are given as 5-year intervals. Work experience is calculated in days from 1993 onwards, and expressed in years. We separate between large city regions and the rest. The large city group is defined as regions with more than 150 000 inhabitants in 2010, which includes 7 regions. We also separate out the top 10 high wage sectors based on the fixed sectoral effects of the individual model reported in Appendix Table A4. Standard errors are given in parenthesis. *** and ** indicate significance at the 1 and 5 percent levels, respectively. All regressions include a constant term.