The urban wage growth premium: Sorting or learning?

Sabine D'Costa[1]
Henry G. Overman[2]

[1] Westminster Business School, University of Westminster and Spatial Economics Research Centre, United Kingdom
[2] LSE, Spatial Economics Research Centre and What Works Centre for Local Economic Growth, United Kingdom

This is a copy of the final published version of an article published in Regional Science and Urban Economics,volume 48, pp. 168-179, September 2014.

The published version is available at:

http://dx.doi.org/10.1016/j.regsciurbeco.2014.06.006

Regional Science and Urban Economics 48 (2014) 168–179

Contents lists available at ScienceDirect

Regional Science and Urban Economics

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

The urban wage growth premium: Sorting or learning?☆

Sabine D'Costa [a,*], Henry G. Overman [b,1]

[a] University of Westminster and Spatial Economics Research Centre, United Kingdom
[b] LSE, Spatial Economics Research Centre and What Works Centre for Local Economic Growth, United Kingdom

ARTICLE INFO

Article history:
Received 28 March 2013
Received in revised form 24 June 2014
Accepted 27 June 2014
Available online 5 July 2014

JEL classification:
R23
J31

Keywords:
Urban wage premium
Agglomeration
Cities
Wage growth
Worker mobility

ABSTRACT

This paper is concerned with the urban wage premium and addresses two central issues about which the field has not yet reached a consensus: first, the extent to which sorting of high ability individuals into urban areas explains the urban wage premium and second, whether workers receive this wage premium immediately, or through faster wage growth over time. Using a large panel of worker-level data from Britain, we first demonstrate the existence of an urban premium for wage levels, which increases in city size. We next provide evidence of a city size premium on wage growth, but show that this effect is driven purely by the increase in wage that occurs in the first year that a worker moves to a larger location. Controlling for sorting on the basis of unobservables we find no evidence of an urban wage growth premium. Experience in cities does have some impact on wage growth, however. Specifically, we show that workers who have at some point worked in a city experience faster wage growth than those who have never worked in a city.

1. Introduction

The urban economics literature provides ample evidence for the existence of an urban wage premium: wages are higher in large urban areas, by between 1% and 11% depending on the sample considered. See, for example, Carlsen at al. (2013), Combes et al. (2008), Di Addario and Patacchini (2008), Fu and Ross (2010), Glaeser and Maré (2001), Melo and Graham (2009), Mion and Naticchioni (2009) and Yankow (2006). Rosenthal and Strange (2004) and Puga (2010) provide reviews. Despite this research, the field has still not reached a consensus on three central issues: first, the extent to which sorting of high ability workers into urban areas can explain observed wage premiums, second, whether urban workers receive this wage premium immediately, or through faster wage growth over time, and third, which of the different agglomeration economies might generate this wage premium. This paper is primarily concerned with the first two of these questions.

To consider these issues we use individual-level data for a large panel of British workers for the period 1998 to 2008. We begin by documenting the existence of an urban wage premium which persists when we control for both observed and unobserved time invariant characteristics of workers (using the panel dimension of our data). We also provide evidence of an urban premium on wage growth, but show that this is driven purely by the increase in wage that occurs in the year that a worker moves from a rural to an urban area. When we exclude move years, we find no evidence of an urban premium for wage growth. If, as Glaeser and Maré (2001) and De la Roca and Puga (2014) argue, an urban wage *growth* premium is evidence of (or at least consistent with) faster human capital accumulation in cities, then for Britain either this mechanism is not at work or faster accumulation is for some reason *not* reflected in faster wage growth for current urban workers. Wheeler (2006) suggests that human capital accumulation as an explanation of an urban wage growth premium might be particularly important for younger workers. In the British context we find some evidence to support this hypothesis. When we restrict our sample to male workers who were 'young' (between 16 and 21) at the beginning of our time period we find some evidence of an urban wage growth premium, over and above that coming from the one-time effect of moving across locations of different sizes.

☆ This work contains statistical data from ONS which is Crown copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. Copyright of the statistical results may not be assigned, and publishers of this data must have or obtain a licence from HMSO. The ONS data in these results are covered by the terms of the standard HMSO "click-use" licence. This work was supported by the Economic and Social Research Council [grant number ES/J021342/1].

* Corresponding author at: Westminster Business School, University of Westminster, 35 Marylebone Road, London NW1 5LS, United Kingdom. Tel.: +44 207 911 5000.

E-mail addresses: s.dcosta@westminster.ac.uk (S. D'Costa), h.g.overman@lse.ac.uk (H.G. Overman).

[1] London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom.

http://dx.doi.org/10.1016/j.regsciurbeco.2014.06.006

We next turn to the issue of whether working in an urban area affects the extent to which wage growth occurs on the job ('within-job') or as a result of moving jobs ('between-job'). It is possible that the absence of an effect overall might hide opposing effects on these two different components (which some have argued might be useful in distinguishing between learning and matching explanations of the urban wage premium). Once again, however, when we control for unobserved characteristics of workers we find no evidence that working in a larger urban area has an effect on either of these two components of wage growth. Again, this contrasts with some of the existing literature for the US, although in this instance the problem appears more to be one of the interpretation of available estimates.[2]

Finally, we consider whether past city 'experience' (i.e. having worked in a city at some point) affects longer-term wage growth. In order to do this, we change our comparison group to those rural workers with no prior experience in cities. We find that in comparison to this group, all workers – those currently working in cities as well as rural workers with past experience in cities – enjoy a wage growth premium. This finding helps reconcile our results with papers emphasising the importance of learning in cities, although in contrast to De la Roca and Puga (2014) we find that both learning *and* sorting matter for understanding the effect of cities on wage growth.

The rest of the paper is structured as follows. The next section reviews related literature. Section 3 outlines our data and provides basic summary statistics. Section 4 provides evidence on the urban wage premium in Britain, while Section 5 considers wage growth. Section 6 then turns to the issue of between versus within-job moves, while Section 7 considers the long-term effects of urban experience. Section 8 presents robustness checks and Section 9 concludes.

2. Existing literature on the urban wage growth premium

As discussed in the Introduction a growing number of papers provide evidence of an urban wage premium (see references above). A number of explanations have been offered for the existence of this premium. According to the productivity hypothesis, market size may facilitate sharing, learning or matching (Duranton and Puga, 2004), increasing productivity in larger locations. Alternatively, according to the selection hypothesis, the direction of causality may be reversed: workers move to productive areas (for reasons that are nothing to do with size) so that productivity increases density (and not vice-versa). If wages are higher in larger cities because of better learning (Glaeser, 1999) or better matching (Zenou, 2009), this implies that not only wage levels, but also wage growth, may be higher in larger locations. Empirically identifying these effects (either static or dynamic) is difficult because once we allow for heterogeneous workers, it may be that higher ability workers self-select into larger locations driving a link between size and wages, assuming that higher ability workers are better paid (Combes et al., 2008) or see faster wage growth.

This paper is specifically concerned with wage growth, that is, with the dynamic aspects of the productivity and selection hypotheses which have received much less consideration in the literature. Wheeler (2006) estimates the impact of density on annual wage growth and on the within-job and the between-job components of annual wage growth. Using a sample of young male workers in the US he finds that wage growth is positively associated with labour market size, and that this is due to between-job wage growth rather than within-job growth. Of course, if more productive individuals select into larger labour markets, as indicated in Combes et al. (2008) and in De la Roca (2011), and these individuals have inherently faster wage growth than average then this, rather than any urban wage growth premium, could explain the higher wage growth in larger cities. If selection or spatial sorting explains the relationship between city size and wage growth, then including worker fixed effects in a panel data specification should make the effect of city size on wage growth disappear. Indeed, when Wheeler (2006) includes fixed effects he finds no significant effect of labour market size on either between-job or within-job wage growth. Our results when including fixed effects are consistent with this finding. Controlling for selection we find no evidence of an urban wage growth premium.

This finding stands in marked contrast to that of a recent paper by De la Roca and Puga (2014) who try to disentangle the static urban wage premium (from working in a city in a given year) from a dynamic urban wage premium (due to higher returns to experience in bigger cities). In contrast to much of the recent literature De la Roca and Puga (2014) find a central role for learning and little evidence of sorting on unobserved ability. We show how our results can be reconciled with theirs once we recognise that unobservable characteristics mean that some workers experience faster wage growth than others independent of location. Controlling for this re-establishes the central role of sorting on unobservables in explaining the urban wage growth premium for current urban workers.[3]

Conceptually, the key to reconciling the two sets of results is to distinguish between three possible sources of faster wage growth for workers who move to and work in cities. We refer to the first source as a 'mobility effect' which is the wage growth that arises because of the increase in wages that occurs at the moment a worker moves from a smaller to a bigger city. In static models, as pointed out by Glaeser and Maré (2001), this jump occurs because of the standard urban wage premium. In the full dynamic specification outlined by de la Roca and Puga (2014) workers experience an additional 'mobility effect' if past experience (learning) is better rewarded in urban locations. A second potential source of faster wage growth in bigger cities is a 'pure' wage growth effect which occurs if otherwise identical workers see faster wage growth in larger cities. Estimates of the size of both the mobility effect and the pure growth effect may be biased upwards by a selection effect. This occurs if more able workers self-select into cities on the basis of characteristics that are unobservable to the econometrician. The full dynamic specification estimated by De la Roca and Puga (2014) controls for the selection effect in terms of wage levels, but needs to impose additional assumptions to control for the selection effect in terms of wage growth (specifically that the effect of unobservables on wage growth is proportional to the effect of unobservables on wage levels). We show that the simplest way to deal with this second selection effect, which does not require us to impose this assumption, is to use panel data to estimate a fixed effects specification for wage growth, rather than wage levels. To control for the mobility effect, we simply drop data corresponding to the move year. We provide more details below.

Yankow (2006) adopts a different approach which allows him to separate the mobility effect from a growth effect, but that does not allow for sorting on unobservables. Using a sample of young US workers from the NLSY, he finds that workers moving into cities experience wage growth in the first year after the move that is 6 percentage points higher than workers remaining in non-urban areas. He also finds a symmetric effect for out of city migrants, who experience wage growth that is 6 percentage points lower than those staying in non-urban areas. In the medium-term out-city migrants have no significant difference in wage growth from non-urban workers. In contrast to these findings, when we consider the role for past experience controlling for selection on unobservables, we find that there are some long-run growth benefits to city experience. This helps reconcile our substantive findings with

[2] See Sections 2 and 6 for details.

[3] This distinction has some parallels with that made in the 'escalator region' literature associated with the work of Fielding (1989, 1992).This literature, which focuses on occupation or social classes, argues that more successful regions (the South East in the UK) attract a disproportionate share of young and qualified workers and act as 'escalators', providing upward social mobility for some of those attracted. Empirical work provides some descriptive evidence, but fails to deal with the question of selection on unobservables (in terms of either wage levels or growth).

De la Roca and Puga (2014): both learning *and* sorting matter for understanding the effect of cities on wage growth.

To summarise, relative to the existing literature, we develop a methodology for studying the urban wage growth premium that allows for the possibility of a mobility effect while controlling for the sorting of wages on the basis of unobservable characteristics that might affect both wage levels and growth.

3. Data[4]

Our analysis is based on the Annual Survey of Hours and Earnings (ASHE) and its predecessor the New Earnings Survey (NES) and covers 1998–2008. NES/ASHE[5] is constructed by the Office of National Statistics (ONS) based on a 1% sample of employees on the Inland Revenue Pay As You Earn (PAYE) register for February and April.[6] ASHE provides information on individuals including their home and work postcodes, while the NES provides similar data but only reports work postcodes. The sample is of employees whose National Insurance numbers end with two specific digits (these have been the same since 1975), meaning NES/ASHE provides an individual-level panel, in which workers are observed for multiple years (up to 11 years in our sample). The sample is replenished as workers leave the PAYE system (e.g. to self-employment) and new workers enter it (e.g. from school).

We allocate workers to locations according to their work postcode allowing us to use the whole sample. The National Statistics Postcode Directory (NSPD) provides a mapping from every postcode to higher-level geographic units. We assign individuals to Travel to Work Areas (TTWAs) using each individual's work postcode. Given the way TTWAs are constructed (so that, in general, at least 75% of the resident population also work within the same area) the work TTWA will also be the home TTWA for the majority of workers. We define cities as TTWAs with more than 100,000 workers in 1999. Sometimes, we further distinguish between small cities, big cities and the London TTWA. We define small cities as TTWAs with 100,000 to 250,000 inhabitants in 1999 and big cities as TTWAs with 250,000 to 1 million inhabitants. Full lists of cities and their size, by size category, are provided in Table A1 of the Appendix.

NES/ASHE includes information on occupation, industry, whether the job is private or public sector, the workers' age and gender and detailed information on earnings including basic pay, overtime pay, basic and overtime hours worked. We use basic hourly earnings as our measure of wages. NES/ASHE does not provide data on education but information on occupation works as a fairly good proxy for our purposes.[7] NES/ASHE provides national sample weights but as we are focused on sub-national (TTWA) data we do not use them in the results we report below.

We follow the existing literature and use a sample of male workers, in order to avoid concerns about the drivers of labour force participation and mobility of female workers.[8] 118,420 male workers are observed, on average, over 8.37 years. Our main outcome of interest is annual wage growth, defined so that wage growth in year t is the growth experienced between $t-1$ and t. Since workers can leave and re-enter the NES/ASHE sample, there are some gaps of more than one year in the data. We calculate annual wage growth when we have wage data for consecutive years, which leaves us with 519,889 observations of annual wage growth. 41% of the workers move across TTWAs at least once in the period (we refer to these workers as 'movers').

Table 1 provides statistics on key explanatory variables for the overall sample, non-movers (i.e. those who never move) and movers (i.e. those who move at least once). In terms of age, the largest number of observations are for workers over 45 (41% of observations overall), while observations for young workers under the age of 24 represent only 7% of our data. 6% of observations are for part-time work, 21% for a public sector job and 56% for a job subject to a collective agreement. In terms of mobility, movers are slightly oversampled (43% of observations are for the 41% of individuals who move at some point during the study period).[9] Overall, 9% of the wage growth observations are for a period when a worker moved across TTWAs. On average, 76% of the observations are for workers working in a city in year t. Of these, the largest category consists of observations for individuals working in small cities, representing 34% of the overall sample. Table 1 also provides summary statistics for various measures of wage growth defined and discussed in more detail below. Average wage growth is 7.1% per annum, wage growth coming from within-job growth is lower, at 6.5% per annum, while wage growth coming from job changes is on average 9.2% per annum.[10]

Movers differ from non-movers mainly in terms of their age and their wage growth. Movers are on average younger and less experienced than non-movers. They also have higher wage levels (basic hourly earnings are £12.5 compared to £11.7 for non-movers) and higher rates of wage growth (7.7% compared to 6.6%).

Five rows near the bottom of the table report wage growth statistics by type of location. These statistics are calculated for non-move years (for reasons discussed below). The first two rows show that wage growth is on average higher in urban areas (6.9%) than in rural areas (6.4%). The next three rows show that it is highest in London and higher for big cities than for small cities, although the difference between big and small cities is very small. Wage growth on average is higher for movers than for non-movers, in any type of location. The fact that this is the case for all three city size categories mitigates concerns about sample selection issues between movers and non-movers, at least across cities of different sizes.

4. The urban wage premium

In this section we estimate the size of the urban wage premium for British cities, ignoring any dynamics (including returns to experience as emphasised by De la Roca and Puga, 2014). That is, we consider the effect of working in a city on wage levels. We have panel data on (log) wages w_{it} for individual i at time t. We follow Glaeser and Maré (2001) and Combes et al. (2008) and use the panel dimension of our

[4] The basic description of the NES/ASHE data is taken from Gibbons et al. (2010).

[5] See Office for National Statistics (2012).

[6] We drop data for Northern Ireland.

[7] Table A2 lists the one-digit SOC occupation categories. We used the LFS to check the median of years of education for each occupation category and we obtain similar results using the median years of education for an individual's occupation rather than our preferred proxy of occupation dummies.

[8] Results which are available upon request using female workers indicate that although the female sample differs from the male sample in terms of several observable characteristics, geographical mobility and mean wages, the urban wage premium is qualitatively similar but more pronounced for women than for men, and more so when we compare young women to young men. Results analysing wage growth reveal very similar patterns for women and for men.

[9] The proportion of movers in our dataset, 41%, is higher than that reported in other work: De la Roca and Puga (2014) have 14% movers across 73 *urban areas* in Spain and Carlsen et al. (2013) have 13% movers across *all* 89 Norwegian economic regions. This is partly due to higher interregional mobility rates in the UK than in other European countries and partly due to different definitions of "movers" across studies of the urban wage premium. We define movers as individuals having changed TTWAs *at some point* during the period (including moves between rural TTWAs and moves between urban TTWAs). Using US data, Wheeler (2006), defining movers as workers moving between labour markets which include both metropolitan areas and non-metropolitan counties, has a proportion of movers in the NLSY dataset of 33%. Our estimate of the proportion of movers in Glaeser and Maré (2001), where they define movers as individuals having changed urban–rural status at least once, is around 24% in the NLSY sample and 28% in the PSID sample. So our figures do not look exceptionally high relative to these more comparable US figures.

[10] Unfortunately neither Yankow (2006) nor Wheeler (2006) provides comparable wage growth statistics for their samples. United States Bureau of Labour Statistics data report that wages grew by 3% per annum for the decade 1998–2008. As discussed in the text, our sample shows figures for Britain that are almost twice as high over the same period. Some of this difference may be real, some may reflect the fact that our wage growth figures are inflated because they reflect the growth for workers in continuous employment (assuming unemployed workers are more likely to experience lower wage growth).

Table 1
Summary statistics (for non-movers, movers and overall sample).

	Overall	Non-movers	Movers
Age (years)	41.5	42.7	40
16–24 years (%)	7	7	8
25–34 years (%)	23	20	27
35–44 years (%)	29	27	31
45+ years (%)	41	45	35
Occupation class 1 (%)	18	17	21
Occupation class 2 (%)	13	12	13
Occupation class 3 (%)	13	12	14
Occupation class 4 (%)	8	8	9
Occupation class 5 (%)	14	16	12
Occupation class 6 (%)	4	4	3
Occupation class 7 (%)	5	4	6
Occupation class 8 (%)	13	15	11
Occupation class 9 (%)	12	12	11
Part time (%)	6	6	5
Public sector (%)	21	21	20
Collective agreement (%)	56	57	55
Basic hourly earnings	12.1	11.7	12.5
Move at least once (%)	43		
Change jobs (%)	21	13	31
Change TTWA (%)	9		21
Work in city (%)	76	77	74
Work in small city (%)	34	32	36
Work in big city (%)	28	29	27
Work in London (%)	14	16	11
Rural with past city experience (%)	5		11
Rural with past small city experience (%)	3		7
Rural with past big city experience (%)	2		5
Rural with past London experience (%)	1		1
Wage growth (%)	7.1	6.6	7.7
Within wage growth (%)	6.5	6.4	6.8
Between wage growth (%)	9.2	8.4	9.6
Annual wage growth in rural areas (%)	6.4	6.2	6.8
Annual wage growth in urban areas (%)	6.9	6.8	7.2
Annual wage growth in small cities (%)	6.8	6.5	7.1
Annual wage growth in big cities (%)	6.8	6.6	7.2
Annual wage growth in London (%)	7.6	7.5	7.9
Number of workers	118,420	69,245	49,175

Notes: Authors' own calculations based on ASHE/NES and LFS data using 519,889 observations for 118,420 workers. One-digit occupation classes are as defined in the Standard Occupation Classification of the Census (see Table A2). Wage growth variables described in Section 5. Wage growth by city size category is for non-move years. Other variables as described in the text.

data to control for selection on unobservables by including fixed effects for each individual i and estimating:

$$w_{it} = \alpha_i + x_{it}^{'}\beta + d_{it}\gamma + \lambda_t + \varepsilon_{it} \tag{1}$$

where α_i is the fixed effect for worker i, x_{it} is a vector of individual and job-specific variables measuring gender, age and other characteristics, d_{it} is a dummy variable that takes value 1 if the individual works in a city at time t, λ_t are a set of time dummies and ε_{it} is the error. β is a vector of coefficients that capture the "returns" to different individual characteristics, while γ is the coefficient which captures the urban wage premium.

As is well known, estimating Eq. (1) without fixed effects correctly identifies the urban wage premium only if we have data on all individual characteristics that affect both sorting and wages.[11] Also, we cannot rule out the possibility that something unobserved changed for the individual that both affected their wage and their place of work. Finally, as highlighted by Combes et al. (2008) identification comes from movers who may not be representative of the population as a whole. These caveats notwithstanding, in the absence of random allocation (or something that as good as randomly assigns people) tracking individuals and observing the change in wages experienced when they move between areas is the best we can do to identify the urban wage premium.

[11] More precisely we need data on all individual characteristics that affect wage and that are correlated with the city dummy.

Table 2
Urban wage premium.

	(1)	(2)	(3)
	OLS	OLS	FE
City	0.141***	0.084***	0.023***
	(0.003)	(0.002)	(0.002)
Age		0.032***	0.028***
		(0.000)	(0.002)
Age^2		−0.001***	−0.001***
		(0.000)	(0.000)
Part time		−0.093***	−0.023***
		(0.004)	(0.004)
Collective agreement		−0.002	0.005***
		(0.002)	(0.001)
Public sector		0.054***	0.034***
		(0.004)	(0.005)
Year dummies	Yes	Yes	Yes
Occupation dummies	No	Yes	Yes
Industry dummies	No	Yes	Yes
Worker fixed effects	No	No	Yes
N	519,889	519,889	519,889
R^2	0.051	0.569	0.498
Number of workers			118,420

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variable is log annual basic hourly earnings.

Results reported in column 1 of Table 2 show that, in the absence of any individual controls, the city premium is quite large at 14.1%. Introducing worker and job characteristics (age, experience, part-time status, collective agreement, public sector job as well as occupation and industry dummies) reduces the city wage premium to 8.4%. Results are reported in column 2. Results in column 3 control for the possibility of sorting across locations on the basis of unobservable worker characteristics by introducing worker fixed effects. Controlling for the sorting of workers further reduces the city wage premium to 2.3%.[12]

Given that the results in Table 2 suggest that an urban wage premium persists even after controlling for individual and job characteristics it is of interest to know whether the effects differ according to labour market size. The theories we rely on relate to the role of large and dense agglomerations rather than small settlements, and predict that the agglomeration effects should be strongest in the largest cities. In line with Glaeser and Maré (2001),Yankow (2006), Gould (2007) and Baum-Snow and Pavan (2012), we therefore focus on estimating the effects of working in cities of different sizes rather than the effect of size or density per se. In Table 3 we report results when we replicate the previous analysis, separating cities into the three size categories described in Section 3.

Results in column 1 (from a specification including only year dummies) show that working in London is associated with a 35.5% higher wage than working in a rural area. The comparable figures are 10.6% for big cities and 8.3% for small cities. The city size premium drops considerably as we introduce explanatory variables. The London premium drops to 23.5% once we control for individual and job characteristics (column 2), and then to 7.1% once we control for unobservable

[12] All our results (with the exception of the within/between job results for the sample of young workers) are robust to the introduction of two city characteristics, diversity and skill share, that may influence wages. We define industrial diversity as the inverse of the Herfindahl index of industry shares of total employment in a TTWA. TTWA skill share is the proportion of the TTWA labour force that has a level of education equal to or higher than NVQ level 4 [National Vocational Qualification level 4 is a qualification in the UK obtained through assessment and training which is informally equivalent to a Higher National Certificate, Higher National Diploma or a first degree.] Data on aggregate employment and the industrial structure of a TTWA comes from the Business Structure Database (BSD). Data on educational attainment is from the Labour Force Survey (LFS). Once we include these two city characteristics, the city wage premium drops further to 1.9% (full results are available upon request).

Table 3
Urban wage premium by city size category.

	(1)	(2)	(3)
	OLS	OLS	FE
Small city	0.083***	0.048***	0.014***
	(0.004)	(0.002)	(0.003)
Big city	0.106***	0.062***	0.025***
	(0.004)	(0.003)	(0.003)
London	0.355***	0.235***	0.071***
	(0.005)	(0.003)	(0.004)
Age		0.032***	0.028***
		(0.000)	(0.002)
Age^2		−0.001***	−0.001***
		(0.000)	(0.000)
Part time		−0.095***	−0.023***
		(0.004)	(0.004)
Collective agreement		0.001	0.005***
		(0.002)	(0.001)
Public sector		0.052***	0.034***
		(0.004)	(0.005)
Year dummies	Yes	Yes	Yes
Occupation dummies	No	Yes	Yes
Industry dummies	No	Yes	Yes
Worker fixed effects	No	No	Yes
N	519,889	519,889	519,889
R^2	0.080	0.582	0.500
Number of workers			118,420

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variable is log annual basic hourly earnings.

time-invariant worker characteristics (column 3).[13] The ranking of the wage premium (largest in London, smaller in big cities, smallest in small cities) is unchanged across specifications. Finally, the reduction in the estimated premium for big and small cities changes in the same way as that of London as we move across specifications. Results which are available on request show that this pattern for small, big cities and London is consistent with that obtained where the city dummies are replaced by the size of the city and its square. The estimated coefficients predict that the wage premium is maximised at a TTWA size of 2.86 m, 600,000 below the size of the London TTWA. Using a continuous measure of city size in a fixed effects model, we find an elasticity of wages with respect to city size of 1.6%.

Our results in Table 3 are comparable to the urban wage premia estimated in Glaeser and Maré (2001) and in Yankow (2006) from US data. With worker fixed effects, Glaeser and Maré (2001) find a premium of 2.6% in non-dense metropolitan areas and 4.5% in dense metropolitan areas using the PSID, and 7% in non-dense metropolitan areas and 10.9% in dense metropolitan areas using the NLSY. Yankow (2006) finds a wage premium of 5% in large cities and 4% in small cities.

5. The urban wage *growth* premium

We turn now from the issue of an urban premium for wage levels to the question of whether such a premium is also observed for the growth of individual wages. De la Roca and Puga (2014) assume that wages are determined by individual characteristics (both observable and unobservable), by the current city of residence and by experience accumulated by the worker to date. The value of experience is allowed to vary depending on both where the experience is accumulated and where that experience is currently being employed. When distinguishing only between urban and rural locations, this model implies that (log) wages of worker i at time t are given by:

$$w_{ict} = \alpha_i + x_{it}^{'}\beta + d_{it}\gamma + \sum_{j=C,R}\delta_{jc}e_{ijt} + \lambda_t + \varepsilon_{ict} \quad (2)$$

where, as before α_i is the fixed effect for worker i, x_{it} is a vector of individual and job-specific variables, d_{it} is a dummy variable that takes value 1 if the individual works in a city at time t, λ_t are a set of time dummies and ε_{it} is the error. Comparing to Eq. (1) the additional components involve e_{ijt}, the total experience accumulated to date by worker i in either cities or rural areas, and δ_{jc} the coefficients capturing the returns to this experience — with the returns indexed by both where the experience was accumulated (j, either a city or a rural area) and where it is currently being used (c, also either a city or a rural area).

De la Roca and Puga (2014) estimate Eq. (2) directly using all observations from a panel of Spanish workers. We adopt an alternative solution of simply first differencing Eq. (2) to give us the following equation for wage growth:

$$\Delta w_{ict} = \left(d_{it} - d_{i(t-1)}\right)\gamma + \sum_{j=C,R}\delta_{jc}e_{ijt} - \sum_{j=C,R}\delta_{jc(t-1)}e_{ij(t-1)} + \Delta\varepsilon_{ict} \quad (3)$$

where, for simplicity we have assumed $x_{it} = x_{it-1}$ and $\lambda_t = \lambda_{t-1}$ to allow us to focus on the items of interest. In years where the worker does not move this equation simplifies to:

$$\Delta w_{ict} = \delta_{jj} + \Delta\varepsilon_{ict} \quad (4)$$

where δ_{jj} is the value of an additional year's urban experience for urban workers or the value of an additional year's rural experience for rural workers. The expression is more complicated in periods where the worker moves. For example, for a worker moving from rural to urban, the expression becomes:

$$\Delta w_{ict} = \left(d_{it} - d_{i(t-1)}\right)\gamma + \delta_{CC} + (\delta_{CC} - \delta_{CR})e_{iC(t-1)} - (\delta_{RR} - \delta_{RC})e_{iR(t-1)} + \Delta\varepsilon_{ict} \quad (5)$$

where the first term captures the static urban premium for moving from rural to urban; the second term captures the dynamic benefits of a year of urban experience (assuming moves occur at the beginning of the period); the third term captures the urban premium for *previous* urban experience; and the fourth term captures the urban 'penalty' imposed when previous rural experience stops being used in a rural area.

Clearly, observed wage growth in move years captures a number of factors that are both static and dynamic. However, using data *only* for non-move years provides a direct estimate of Eq. (5) of δ_{CC} and δ_{RR}. That is, it tells us whether there is an urban wage growth premium such that wages grow faster for workers in cities than for those in rural areas.

So far, apart from ease of interpretation there is little to recommend our approach over estimation of the full dynamic model. This changes, however, if we now allow for the possibility that unobserved worker characteristics might influence wage *growth* as well as wage levels. That is, if we generalise Eq. (2) as follows:

$$w_{it} = \alpha_i + x_{it}^{'}\beta + d_{it}\gamma + \sum_{i=C,R}\delta_{jc}e_{ijt} + \delta_i e_{it} + \lambda_t + \varepsilon_{ict} \quad (6)$$

where everything is as before except for the inclusion of an individual-specific return to experience. If we assume that the unobserved individual return to experience is proportional to the unobserved individual effect on wage levels (i.e. $\delta_i = \alpha_i$) then we could follow De la Roca and Puga (2014) and estimate Eq. (6) using an iterative process.[14] In the first step, this process assumes values for the fixed effects and estimates parameters on the basis of these fixed effects. In the second step, the

[13] Including city characteristics makes little difference, with the estimated London premium falling slightly to 6.6%.

[14] In fact, De la Roca and Puga (2014) allow the unobserved return on experience to be city specific, but proportional to the individual fixed effect on wage levels. See their Eq. (11). Inspection of our Eq. (7) shows that this specification is impossible to estimate without imposing the assumption that $\delta_i = a_i$ because allowing for unobserved individual returns to be location specific means that δ_i cannot be separately identified from δ_{jj}.

process uses the resulting parameter estimates to re-calculate the fixed effects and plugs these in to the first stage, repeating this process until estimates converge. In contrast, we adopt a more standard specification, dropping the constraint on the proportionality of the fixed effects (i.e. allowing for the possibility that $\delta_i \neq \alpha_i$) and obtain consistent estimates of δ_{CC} and δ_{RR} by dropping move years and using a panel of individual wage growth rates to estimate[15]:

$$\Delta w_{ict} = \delta_i + \delta_{jj} + \Delta x_{it}^{'}\beta + \lambda_t + \Delta\varepsilon_{ict} \tag{7}$$

Note that this is preferable to dropping movers entirely (as in Wheeler, 2006) because that would only deal with the mobility effect and not sorting on unobservables. If we focus only on non-movers, then with fixed effects it is impossible to estimate the effects of city dummies and the identification of the effects of time-varying location characteristics comes only from time variation in those characteristics. By including movers, but dropping the years when workers move, identification of the effects of location characteristics comes from both time series and cross-section variation for movers, and it is possible to estimate the effects of city dummies from the movers. A similar logic – dropping move years and using fixed effects to control for unobservables – can be used to develop equivalent expressions for different city size categories (small, medium, large) or, indeed, to identify a full set of city dummies. As usual, identification relies on observing outcomes for movers in multiple time periods and across all city types. Given that we have data at annual frequencies, getting sufficient variation to identify the full set of dummies is a challenge and so we focus, instead, on identifying the urban premium for a small number of city size classes.

We begin by estimating Eq. (3) ignoring both the mobility effect and the possibility of sorting on unobservables. Results in column 1, of Table 4 (from a specification including only year dummies) show that working in London in year t is associated with wage growth between $t-1$ and t that is 1.4 percentage points greater than that experienced by workers living in rural areas. Working in a big city is associated with 0.4 percentage point higher growth, while working in a small city is associated with a 0.35 percentage point higher growth rate. Column 2 reports results when we introduce observable worker and job characteristics. The London wage growth premium drops to 0.6 percentage points, that of big cities to 0.1 and that of small cities to 0.12 percentage points. Column 3 allows for individual fixed effects to control for the sorting of individuals on the basis of unobserved characteristics. This substantially increases the estimated urban wage growth premium.

As discussed above, the estimated city coefficients combine both a mobility and pure growth effect. To control for the former, while continuing to allow for sorting on unobservables, we drop move years and estimate Eq. (7). Dropping observations for move years leaves us with a sample of both movers and non-movers, but we only use wage growth for years when workers remained in the same labour market. The number of observations drops by 9% from 520,000 to 473,000.[16] Once we both drop move years and include individual fixed effects we no longer detect any effect of city size on wage growth, as can be seen in column 4. In Britain, the higher wage growth rates observed in cities appear to be driven by the sorting of higher ability individuals experiencing 'one-off' higher wage growth in the year when they move into larger labour markets.[17]

These results help explain the large jump in coefficients that we see when we move between columns 2 and 3 (i.e. introduce fixed effects for the sample including observations for all years). With individual fixed effects, the identification of the coefficients on the city dummies comes only from movers and so observations from move years represent a high proportion of observations used to identify the city size effects. When there is a significant urban wage premium for levels, or when the returns to experience increase a lot when moving to a larger city, this biases estimated urban wage growth premiums upwards. In short, including fixed effects *and* dropping move years is necessary to isolate the pure growth effect from effects of mobility and sorting. In Britain, at least for the entire sample of workers, we find no evidence of a pure growth effect once we make both these corrections.

As usual, one concern in introducing fixed effects is that movers may differ systematically from non-movers. As discussed above, descriptive statistics in Table 1 suggest that differences are mostly small. One remaining concern, which has received some attention in the literature (Wheeler, 2006; Yankow, 2006) is that movers are on average younger and less experienced than non-movers. In addition, some studies, including Wheeler (2006) have focused on younger workers suggesting that the urban wage premium may be particularly pronounced for the young. For both reasons, it is interesting to repeat our analysis focussing only on younger workers and it is to this issue that we now turn.

We restrict the sample to male workers who were aged between 16 and 21 in 1998, the first year of our dataset.[18] Over the study period, these individuals provide us with 52,000 wage growth observations for workers aged between 16 and 32 (the mean age is 25 years old). As we would expect, given that movers are on average younger, the proportion of movers is slightly higher than that in the full sample, at 49%. The annual wage growth is much higher (11.25% vs. 7.09%).

Results in Table 5 replicate those in Table 4 using the sample of young workers. Consistent with the fact that annual wage growth is much higher for young workers, the OLS coefficients are consistently larger for this restricted sample. Further, when we include fixed effects and remove the move years the effects of cities remain significant in contrast to the results for the full sample. Small cities increase wage growth for young workers by 2.15 percentage points, big cities by 1.88 percentage points and London by 2.26 percentage points. So for younger workers, even after controlling for worker observable and unobservable characteristics, there is evidence of a pure effect on wage growth of working in cities compared to rural areas.

6. Between versus within-job wage growth

Given that we observe some workers in multiple jobs, we can distinguish two types of wage growth — within-job wage growth (when the worker stays in the same job) and between-job wage growth (when the worker changes jobs). The size of a labour market can have an effect on both types of wage growth. Wheeler (2006) argues that better learning in cities is more likely to be reflected in higher within-job wage growth, while better matching of workers and jobs, is more likely to be reflected in between-job wage growth. Even if one is not fully convinced by these assertions, the question of the impact on these different types of wage growth is still empirically interesting. In particular, in our context, it is interesting to consider whether our finding of no urban wage growth premium for the full sample disguises offsetting effects on within-job and between-job wage growth. We investigate this possibility by

[15] In Eq. (7) δ_i can be isolated in this way because we have assumed in Eq. (6) that the effect of individual experience is linear. We could allow for experience to enter non-linearly by imposing the assumption that the effect of any non-linear terms in experience was proportional to δ_i. This assumption is imposed in De la Roca and Puga (2014), in addition to the assumption that the unobserved individual return to experience (and its square) is proportional to the unobserved individual effect on wage levels.

[16] OLS estimates for this restricted sample of observations give results for the city dummies that are not significantly different from those for the full sample – small city has a coefficient of 0.123 (s.e. 0.048), big city 0.082 (s.e. 0.050) and London 0.522 (0.066) – mitigating concerns about the representativeness of the data when dropping move years. Full results are available on request.

[17] One worry might be that industry and occupation variables should be considered *city* rather than *individual* characteristics. However, replicating the fixed effects results in column 4 omitting these variables leaves results essentially unchanged.

[18] Wheeler (2006) uses a cohort panel which follows workers who were between 14 and 21 as of 31 December 1978, from 1978 until 1994.

Table 4
Urban wage growth premium by city size category.

	(1)	(2)	(3)	(4)
	OLS	OLS	FE	FE — no move year
Small city	0.352***	0.120**	0.536***	0.245
	(0.052)	(0.050)	(0.185)	(0.206)
Big city	0.402***	0.097*	0.650***	0.154
	(0.054)	(0.052)	(0.199)	(0.217)
London	1.378***	0.615***	2.117***	0.399
	(0.071)	(0.068)	(0.270)	(0.292)
Age		−0.640***	−0.572***	−0.567***
		(0.009)	(0.098)	(0.110)
Age^2		0.009***	0.012***	0.011***
		(0.000)	(0.000)	(0.000)
Part time		1.003***	3.056***	4.433***
		(0.133)	(0.282)	(0.292)
Collective agreement		−0.146***	−0.026	0.067
		(0.045)	(0.079)	(0.079)
Public sector		0.012	0.764***	0.455
		(0.072)	(0.295)	(0.289)
Year dummies	Yes	Yes	Yes	Yes
Occupation dummies	No	Yes	Yes	Yes
Industry dummies	No	Yes	Yes	Yes
Worker fixed effects	No	No	Yes	Yes
N	519,889	519,889	519,889	473,088
R^2	0.004	0.032	0.012	0.012
Number of workers			118,420	114,836

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variable is percentage annual growth in basic hourly earnings.

estimating the same models as in the previous section, replacing annual wage growth with measures for the two different types of wage growth.

In our data we are not able to assign each worker to a particular employer identifier. We therefore define a job change if the work postcode changes from one year to the next. Because postcodes in Britain are very small, often corresponding to a single building, this should provide a good indicator of a job change.[19] Since we only observe data annually, we face two further potential measurement issues. First, our within-job wage growth measure may miss some growth that occurs after the last time we observe the worker in a particular job but before they move to a new job. This would only affect our results, however, if wage growth differs towards the end of a job in different ways depending on area characteristics. This seems unlikely, although we cannot rule out this possibility. Second, our between-job wage growth includes the cumulative effect of all job changes in any given year. Again, it is not obvious that this creates any particular problems for us (other than the fact that it prevents us from studying the frequency of job changes).

We report results from separate regressions of within-job wage growth and between-job wage growth in Table 6. All specifications include worker fixed effects and drop observations corresponding to years when a worker moves across locations. For comparison, column 1 replicates the results for overall wage growth taken from column 4 in Table 4. Columns 2 and 3 then report the specifications using within-job wage growth and between-job wage growth as the dependent variable, respectively. We find no evidence of an urban premium for either within-job or between-job wage growth. Results which are available on request show a positive effect of all three city size categories on between-job growth if we include move years but, as for overall wage growth, it is the increase in wages when moving to larger cities that drives this effect. Once there, the results in column 3 show that between-job wage growth is no higher than it would have been in other areas.[20]

In short, when we include worker fixed effects and consider only the years without geographical moves, we find no evidence in favour of a pure effect of city size on either type of wage growth. This contradicts the results of Wheeler (2006) of positive effects of density on wage growth, particularly through between-job wage growth. That said, Wheeler's results of positive effects of density on wage growth, like ours, are not robust to the inclusion of worker fixed effects. As with overall wage growth, the absence of an effect once fixed effects are included suggests that the OLS results are due to the spatial sorting of more productive workers into larger markets, rather than the effects of larger markets per se.

Again, it is possible that these effects might be larger for younger workers who are more likely to switch jobs in the early years of their careers and to benefit from those job switches more than older workers (Topel and Ward, 1992; Chan and Stevens, 2004). When we focus on the subset of younger workers and remove the move years, as we did in Section 5, we find that small cities provide an advantage in terms of within-job wage growth. We also find some evidence that big cities and London have a positive effect on between-job wage growth (although all three coefficients are only significant at the 10% level).[21]

7. The long-term effects of city experience

The analysis so far shows that wage growth increases with city size but that, for most workers, this effect is a short-term one driven by wage increases in move years. For the sample as a whole there is no evidence that this growth premium persists beyond the first year and no evidence that results for overall wage growth hide offsetting effects on within and between-job wage growth. In this section we consider one final channel through which city living may affect longer-term wage growth by considering whether city 'experience' (i.e. having worked in a city at some point) affects wage growth.

[19] It is possible however, that in some instances a change in work postcode would falsely indicate a job change in the case of employer-induced relocation. See, for example, the discussion in Gutiérrez-i-Puigarnau and van Ommeren (2010).

[20] Note that identification in column 3 relies on the sub-sample of people who move between jobs multiple times (and more times than they move between the different sizes of cities during our study period).

[21] When we include two city characteristics (TTWA industrial diversity and high skill share), we no longer detect an effect of city size on overall wage growth (results available on request). This leads us to believe that the 'pure' urban wage growth premium enjoyed by younger workers is due to these two features of cities.

Table 5
Urban wage growth premium by city size for a sample of younger workers.

	(1)	(2)	(3)	(4)
	OLS	OLS	FE	FE — no move year
Small city	0.998***	0.635***	1.690**	2.146**
	(0.216)	(0.209)	(0.727)	(0.895)
Big city	0.796***	0.463**	1.847**	1.877**
	(0.222)	(0.216)	(0.762)	(0.940)
London	2.692***	1.898***	4.720***	2.256**
	(0.266)	(0.262)	(0.984)	(1.141)
Age		−1.871***	−1.675**	−2.085***
		(0.151)	(0.656)	(0.676)
Age^2		0.055***	0.088***	0.098***
		(0.008)	(0.012)	(0.012)
Part time		−2.134***	−2.449***	−0.210
		(0.341)	(0.690)	(0.734)
Collective agreement		−0.457***	−0.089	−0.112
		(0.177)	(0.303)	(0.310)
Public sector		0.047	1.978*	1.512
		(0.391)	(1.174)	(1.155)
Year dummies	Yes	Yes	Yes	Yes
Occupation dummies	No	Yes	Yes	Yes
Industry dummies	No	Yes	Yes	Yes
Worker fixed effects	No	No	Yes	Yes
N	51,789	51,789	51,789	45,496
R^2	0.016	0.044	0.035	0.035
Number of workers			17,037	16,043

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variable is percentage annual growth in basic hourly earnings. Workers included in the sample were aged between 16 and 21 in 1998.

From our results so far we know that working in a city brings no wage growth premium compared to working in a rural area but this does not rule out the possibility that city experience has an impact in the future. We cannot examine this possibility by looking at workers currently living in cities, but we can consider it by checking to see if rural workers with previous urban work experience have faster wage growth than rural workers with no previous urban work experience.

To do this we introduce an *Evercity* indicator which takes value one if an individual works in a rural area at time t and has at least one year of previous work experience in a city. In Table 7 we report results from wage growth regressions including the *Evercity* variable and the *City* indicator (for workers *currently* working in a city), so that the omitted category is workers who have always worked in a rural area. The *Evercity* dummy therefore indicates the effect of past urban experience on the wage growth of rural workers compared to having always worked in a rural area, while the *City* dummy indicates the effect of currently working in a city compared to having always worked in a rural area.[22]

The OLS without fixed effects results in column 1 indicate that past urban experience has a significant effect, increasing the wage growth relative to rural workers who have no city experience by 0.56 percentage points. We again explore the possibility that this effect may be due to worker heterogeneity or to the short-term effect of mobility out of cities by including worker fixed effects and dropping move years. Results, reported in column 2 show there is still a significant effect of 1 percentage point additional wage growth. That is, in the longer term there appears to be a wage growth premium for rural workers who have had past urban work experience, compared to those who have never had any city experience. In addition, we now find that compared to those having never had any city experience, current urban workers experience a wage growth premium of 0.9 percentage points. So when we consider this comparison group, there is an urban wage premium for all workers.

In fact, column 2 also provides some evidence of a hierarchy in wage growth: compared to rural workers with no city experience, those currently working in a city enjoy a wage growth premium which is lower than that of currently rural workers with some past urban experience (although these coefficients are not significantly different).

[22] In the specification without move years, the *Evercity* dummy is equal to 1 for wage growth observations in year t where the individual is in a rural area in both years $t-1$ and t and has city experience before year $t-1$. Since we include fixed effects, the *Evercity* and *City* effects are identified for movers from rural to urban to rural or from urban to rural to urban. We need to observe at least one growth rate in each location (remembering that we also drop move years).

Table 6
Urban within and between-job wage growth premium by city size category.

	(1)	(2)	(3)
	Overall wage growth FE — no move year	Within	Between — no move year
Small city	0.245	0.259	−1.677
	(0.206)	(0.214)	(1.567)
Big city	0.154	0.111	−2.621
	(0.217)	(0.225)	(1.629)
London	0.399	−0.023	−0.203
	(0.292)	(0.303)	(1.892)
Age	−0.567***	−0.460***	−1.248
	(0.110)	(0.132)	(0.772)
Age^2	0.011***	0.009***	0.020***
	(0.000)	(0.000)	(0.003)
Part time	4.433***	5.696***	0.037
	(0.292)	(0.316)	(1.310)
Collective agreement	0.067	0.029	0.128
	(0.079)	(0.080)	(0.484)
Public sector	0.455	−0.327	2.358
	(0.289)	(0.293)	(1.559)
Year dummies	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes
N	473,088	411,215	61,873
R^2	0.012	0.013	0.018
Number of workers	114,836	109,619	41,518

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variables are percentage annual growth, percentage within-job annual growth and percentage between-job annual growth in basic hourly earnings.

Table 7
Long-run effect of city experience.

	(1)	(2)
	OLS	FE — no move year
Evercity	0.556***	1.035***
	(0.105)	(0.322)
City	0.307***	0.878***
	(0.046)	(0.279)
Age	−0.641***	−0.568***
	(0.009)	(0.110)
Age^2	0.009***	0.011***
	(0.000)	(0.000)
Part time	1.011***	4.434***
	(0.133)	(0.292)
Collective agreement	−0.155***	0.069
	(0.045)	(0.079)
Public sector	0.017	0.452
	(0.072)	(0.289)
Year dummies	Yes	Yes
Occupation dummies	Yes	Yes
Industry dummies	Yes	Yes
Worker fixed effects	No	Yes
N	519,889	473,088
R^2	0.032	0.012
Number of workers		114,836

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variable is percentage annual wage growth.

This is consistent with the idea that in Great Britain "successful" urban workers relocate to rural areas. It also explains why using all rural workers as a comparison group, as we did in Section 5 in accordance with the rest of the urban wage premium literature, under-estimates the urban wage growth premium.

We now break down the past city experience of rural workers into three categories: London experience indicated by the variable *Everlondon* and experience in big and small cities indicated by *Everbigcity* and *Eversmallcity*, respectively. These categories are not distinct as rural workers can have past experience in more than one type of city however the correlation between these indicators is very low. We also include separate dummies for workers currently working in small cities, big cities and London. Again the omitted category consists of rural workers with no prior urban experience.

Results in Table 8 indicate that, in comparison with having never had any city experience, current and past experience in a small city brings about the highest wage growth premium: rural workers with past experience in a small city enjoy a 1 percentage point premium, while those currently working in small cities enjoy a 0.9 point premium. When we turn to big cities, we find that the wage growth premium is also significant but smaller: 0.7 point for rural workers with some past experience in a big city and about the same for those currently working in a big city. For London there is a wage growth premium of 0.9 point from currently working in London, but we do not find any effect of past experience in London.

We now turn to the effect of past city experience on the separate within-job and between-job components of wage growth. The first column of Table 9 replicates the second column of Table 7, where for currently rural workers the overall effect of past city experience on wage growth for the years when workers do not move across locations is a 1% higher wage growth. This effect comes through higher wage growth within jobs, as can be seen by the coefficient in column 2 indicating that wage growth within jobs is 0.9 points higher for rural workers with past city experience than for rural workers with no past city experience. We find no significant effect on between-job wage growth (column 3). We interpret these results as showing that, after controlling for time-invariant unobserved ability, rural workers with past urban experience have acquired skills and capabilities that enable them to achieve higher wage growth on the job once they relocate to rural areas.

Table 8
Long-run effect of city experience by city size category.

	(1)	(2)
	OLS	FE — no move year
Eversmallcity	0.480***	1.050***
	(0.130)	(0.358)
Everbigcity	0.667***	0.681*
	(0.158)	(0.397)
Everlondon	−0.230	0.166
	(0.300)	(0.642)
Small city	0.228***	0.861***
	(0.051)	(0.283)
Big city	0.208***	0.733**
	(0.053)	(0.291)
London	0.729***	0.949***
	(0.069)	(0.350)
Age	−0.641***	−0.569***
	(0.009)	(0.110)
Age^2	0.009***	0.011***
	(0.000)	(0.000)
Part time	1.004***	4.433***
	(0.133)	(0.292)
Collective agreement	−0.147***	0.069
	(0.045)	(0.079)
Public sector	0.011	0.450
	(0.072)	(0.289)
Year dummies	Yes	Yes
Occupation dummies	Yes	Yes
Industry dummies	Yes	Yes
Worker fixed effects	No	Yes
N	519,889	473,088
R^2	0.032	0.012
Number of workers		114,836

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variable is percentage annual wage growth.

8. Robustness checks

Given that the wages of public sector workers may be determined centrally, we replicate our core results using a sample excluding public sector workers, identified from the public sector indicator in the ASHE

Table 9
Long-run effect of city experience on within and between-job wage growth.

	(1)	(2)	(3)
	Overall wage growth FE — no move year	Within	Between — no move year
Evercity	1.035***	0.929***	3.041
	(0.322)	(0.339)	(2.150)
City	0.878***	0.764***	0.017
	(0.279)	(0.291)	(1.964)
Age	−0.568***	−0.461***	−1.239
	(0.110)	(0.132)	(0.766)
Age^2	0.011***	0.009***	0.020***
	(0.000)	(0.000)	(0.003)
Part time	4.434***	5.694***	0.044
	(0.292)	(0.316)	(1.309)
Collective agreement	0.069	0.029	0.136
	(0.079)	(0.080)	(0.484)
Public sector	0.452	−0.332	2.367
	(0.289)	(0.293)	(1.559)
Year dummies	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes
N	473,088	411,215	61,873
R^2	0.012	0.013	0.018
Number of workers	114,836	109,619	41,518

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variables are percentage annual growth, percentage within-job annual growth and percentage between-job annual growth in basic hourly earnings.

Table 10
Urban wage and wage growth premium for private sector workers.

	(1)	(2)	(3)	(4)	(5)
	FE	FE	FE	FE — no move year	FE — no move year
Evercity					1.270***
					(0.401)
City	0.021***				1.072***
	(0.003)				(0.341)
Small city		0.013***	0.369*	0.279	
		(0.003)	(0.210)	(0.242)	
Big city		0.024***	0.664***	0.186	
		(0.003)	(0.225)	(0.255)	
London		0.062***	1.610***	0.574*	
		(0.004)	(0.299)	(0.340)	
Age	0.028***	0.028***	−0.552***	−0.561***	−0.561***
	(0.002)	(0.002)	(0.116)	(0.128)	(0.128)
Age^2	−0.001***	−0.001***	0.012***	0.012***	0.012***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Part time	−0.011***	−0.011***	3.138***	4.413***	4.415***
	(0.004)	(0.004)	(0.333)	(0.347)	(0.347)
Collective agreement	0.002**	0.002**	0.077	0.176**	0.178**
	(0.001)	(0.001)	(0.085)	(0.086)	(0.086)
Year dummies	Yes	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes	Yes	Yes
N	412,834	412,834	412,834	374,485	374,485
R^2	0.495	0.496	0.012	0.012	0.012
Number of workers	100,963	100,963	100,963	97,318	97,318

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variable is log annual basic hourly earnings (columns 1 and 2) and percentage annual wage growth (columns 3–5).

data. When we do this we are left with 80% of our original sample size. Results are reported in Table 10. The first column reports results for the urban wage premium from a regression including individual fixed effects and should be compared to that reported in column 3 of Table 2. Focusing only on private sector workers suggests a slightly smaller urban wage premium (2.1% as opposed to 2.3% for the full sample). The second column reports results for the same specification, but using the three city size dummies as reported in column 3 of Table 3.

Table 11
Urban wage and wage growth premium for high-tech industries.

	(1)	(2)	(3)	(4)	(5)
	FE	FE	FE	FE — no move year	FE — no move year
Evercity					−0.955
					(1.646)
City	0.026*				−1.115
	(0.014)				(1.495)
Small city		0.021	−0.695	0.040	
		(0.014)	(0.913)	(1.008)	
Big city		0.030*	−0.748	−1.207	
		(0.016)	(0.915)	(1.009)	
London		0.040**	−0.397	−0.586	
		(0.019)	(1.206)	(1.261)	
Age	0.019***	0.019***	−1.308***	−0.841***	−0.833***
	(0.004)	(0.004)	(0.493)	(0.291)	(0.291)
Age^2	−0.001***	−0.001***	0.020***	0.017***	0.017***
	(0.000)	(0.000)	(0.002)	(0.002)	(0.002)
Part time	0.011	0.012	9.642***	9.382***	9.393***
	(0.029)	(0.029)	(2.245)	(2.299)	(2.299)
Collective agreement	−0.001	−0.001	0.095	0.382	0.379
	(0.004)	(0.004)	(0.292)	(0.294)	(0.295)
Public sector	−0.022*	−0.021*	0.108	0.340	0.373
	(0.011)	(0.011)	(0.939)	(0.932)	(0.929)
Year dummies	Yes	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes	Yes	Yes
N	30,321	30,321	30,321	27,519	27,519
R^2	0.481	0.481	0.032	0.029	0.029
Number of workers	9096	9096	9096	8591	8591

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variable is log annual basic hourly earnings (columns 1 and 2) and percentage annual wage growth (columns 3–5).

Table 12
Urban wage and wage growth premium for high-skilled workers.

	(1) FE	(2) FE	(3) FE	(4) FE — no move year
City	0.022***			
	(0.005)			
Small city		0.013**	0.119	0.240
		(0.006)	(0.363)	(0.402)
Big city		0.024***	0.198	0.097
		(0.006)	(0.397)	(0.431)
London		0.053***	0.906*	0.384
		(0.008)	(0.500)	(0.541)
Age	0.043***	0.044***	−1.005***	−1.141***
	(0.004)	(0.004)	(0.248)	(0.258)
Age^2	−0.001***	−0.001***	0.016***	0.015***
	(0.000)	(0.000)	(0.001)	(0.001)
Part time	0.119***	0.118***	11.574***	11.150***
	(0.010)	(0.010)	(0.834)	(0.842)
Collective agreement	0.000	0.001	−0.088	0.377
	(0.002)	(0.002)	(0.168)	(0.168)
Public sector	0.027**	0.014	0.321	0.316
	(0.010)	(0.011)	(0.618)	(0.597)
Year dummies	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes	Yes
N	114,928	114,928	114,928	104,401
R^2	0.518	0.517	0.024	0.024
Number of workers	25,180	25,180	25,180	25,180

Note: Robust standard errors in parentheses. Standard errors are clustered by worker. ***, **, * indicate significant at 1%, 5% and 10% level respectively. Dependent variable is log annual basic hourly earnings (columns 1 and 2) and percentage annual wage growth (columns 3 and 4).

All three coefficients are slightly smaller when using private sector workers with the most pronounced fall for London (the premium falls from 7.1% to 6.2%). Turning to wage growth, column 3 (which replicates column 3 of Table 4) indicates that the wage growth premium is lower for private sector workers, particularly in London, when we include the years when workers move across locations. Similarly to the main results (column 4 of Table 4), we find no "pure" growth effect in small and big cities when restricting attention to private sector workers (column 4). There does, however, seem to be a small effect in London where private sector workers enjoy annual wage growth 0.6 points higher than in rural areas in the years when they do not move, although the coefficient is only significant at the 10% level. Finally, column 5 reports results when we look at the long-run effects of city experience (compare to column 2 of Table 7). The results for the private sector sample confirm those obtained from the full sample, although the effect of current and past city experience on the wage growth of private sector workers is larger.

We might expect the urban wage premium and the wage growth premium to be greater for industries where knowledge spillovers are important (Audretsch and Feldman, 1996; Charlot and Duranton, 2004; Duranton and Puga, 2001, Henderson, 2003, Marshall, 1920). To consider this possibility we, once again, replicate our core results for a sub-sample of high-technology industries: SIC 2003 codes 30–33, 72 and 73. These correspond to the high-tech industries examined in Moretti (2010). This reduced sample has 30,321 observations and the proportion of move years is also 9%.

Results are reported in Table 11. Again the first column reports results for the urban wage premium from a regression including individual fixed effects and should be compared to that reported in column 3 of Table 2. Focusing only on high-tech sectors suggests a slightly larger urban wage premium (2.6% as opposed to 2.3% for the full sample). The second column reports results for the same specification, but using the three city size dummies as reported in column 3 of Table 3. We find no wage premium in small cities, a slightly larger effect in big cities (3% vs. 2.5% for the full sample) and a smaller effect in London (4% vs. 7.1%). The smaller sample sizes mean that all of these point estimates are less precisely estimated, although the difference for London between the full sample and the high-tech sample does appear to be (just) statistically significant. Turning to wage growth, columns 3 and 4 (which replicate columns 3 and 4 of Table 4) suggest that there is no wage growth premium for this sample, whether or not we include move years. A comparison of the standard errors suggests that these point estimates are very imprecisely estimated for this much smaller sub-sample. Finally, column 5 reports results when we look at the long-run effects of city experience (compare to column 2 of Table 7). Once again, the standard errors suggest that these coefficients are very imprecisely estimated.

We might expect the urban wage premium and the wage growth premium to be greater for high-skilled workers (Moretti, 2004; Rotemberg and Saloner, 2000). To consider this possibility we, once again, replicate our core results for a sub-sample of high-skilled workers.[23] This reduced sample has 114,928 observations. Results for this sample are reported in Table 12. The first column reports results for the urban wage premium from a regression including individual fixed effects and should be compared to that reported in column 3 of Table 2. Focusing only on high-skilled workers suggests a slightly smaller urban wage premium (2.2% as opposed to 2.3% for the full sample). The second column reports results for the same specification, but using the three city size dummies as reported in column 3 of Table 3. We find roughly similar wage premiums in small and big cities (1.3% vs. 1.4%; 2.4% vs. 2.5% respectively), and a smaller effect in London (5.3% vs. 7.1%). The smaller sample sizes mean that all of these point estimates are less precisely estimated, although all of the estimates are statistically significant. Turning to wage growth, columns 3 and 4 (which replicate columns 3 and 4 of Table 4) suggest that there is no wage growth premium for this sample, once we exclude move years. Unfortunately, we cannot estimate the long-run effects of city experience for this smaller sub-sample, because we have too few observations to allow us to separate out the evercity and fixed effects.[24]

9. Conclusion

Using micro-level data on British workers, we find no evidence that wages grow faster for workers living in cities once we allow for the possibility that workers sort on the basis of unobservable worker characteristics that influence both wage levels and wage growth. Wages do grow faster in the year that workers move to a city. That is, there is a mobility effect which may come from either a static urban wage premium (in the traditional sense) or from a change in the returns to *existing* experience when workers move (as emphasised by De la Roca and Puga, 2014). This finding of no urban wage growth premium for current urban workers, does not mean that city experience has no effect. When compared to rural workers who have *never* had any urban experience, we find an urban wage growth premium for all workers who have either current or past urban experience. In particular, rural workers with past urban experience enjoy higher annual wage growth within jobs. We view this as evidence in favour of the learning in cities hypothesis, where the acquired skills are highly transferable, so workers

[23] High-skilled workers are defined using their two-digit occupation codes according to the definition of high-skilled occupations in the SOC 1990 and SOC 2000 documentations. These include corporate managers and administrators as well as professional occupations (science and engineering professionals, health professionals, teaching and research professionals and business and public service professionals).

[24] Footnote 22 discusses the sample from which this effect is identified.

with past urban experience carry their acquired skills with them after relocating to rural areas. This also tells us that comparing currently urban to currently rural workers, as has been done widely in the literature on the urban wage premium under-estimates the urban wage growth premium.

We have addressed two main issues in the urban wage growth premium literature: the role of sorting of high ability individuals into larger locations and whether workers receive a wage growth premium immediately upon moving to a city or if there are long-lasting effects. To do this, we use standard panel data models for wage growth, dropping move years to help separate out mobility from pure growth effects. In contrast to the iterative methodology developed by De la Roca and Puga (2014) this approach does not impose the assumption that the effect of unobserved characteristics on wage growth is proportional to the effect on wage levels. Applying it, we find that understanding the urban wage growth premium requires us to recognise that both sorting *and* learning play a role in explaining higher wage growth in cities.

Appendix: Information on cities and occupations in our dataset

Table A1 provides a list of the urban TTWAs present in our dataset (with more than 100,000 workers). Our original data consists of 297 TTWAs, with average size of 91,000 workers. The TTWA names listed in Table A1 are those constructed from the 1991 Census.

Table A2 lists the job categories represented by the one-digit SOC classification.

Table A1
Lists of cities and their size by city size category.

Small Cities	Size	Small cities (cont.)	Size
Peterborough	102,561	Brighton	187,955
Warwick	104,683	Wigan & St Helens	200,208
Dundee	106,552	Oxford	204,280
Pontypridd & Aberdare	107,454	Hull	204,796
Poole	107,856	Sunderland & Durham	210,868
York	108,396	Stoke	213,546
Tunbridge Wells	108,538	Middlesbrough & Stockton	217,919
Chichester	110,929	Dudley and Sandwell	220,975
Huddersfield	113,680	Cardiff	221,505
Barnsley	115,306	Crawley	222,566
Crewe	121,324	Guildford & Aldershot	235,027
Swindon	123,106	Wolverhampton & Walsall	235,785
Ipswich	129,300	Bradford	240,386
Harlow	132,063	Portsmouth	241,156
Swansea	132,343	Wirral and Chester	242,895
Exeter	133,857	Reading	248,302
Milton Keynes	134,828	Coventry	249,331
Bolton	135,505		
Mansfield	137,628	Big cities	
Northampton	139,636	Southampton & Winchester	278,893
Blackburn	143,660	Leicester	283,809
Doncaster	145,846	Maidstone & North Kent	310,276
Luton	146,119	Southend	317,158
Cambridge	146,490	Leeds	336,464
Motherwell and Lanark	147,605	Nottingham	349,397
Blackpool	149,035	Bristol	353,477
Wakefield	153,724	Sheffield & Rotherham	363,643
Warrington	154,424	Edinburgh	399,116
Plymouth	159,050	Liverpool	443,340
Bournemouth	160,063	Tyneside	488,481
Stevenage	161,270	Slough & Woking	641,708
Derby	163,753	Glasgow	648,197
Colchester	164,193	Birmingham	808,982
Preston	166,868	Manchester	976,796
Aberdeen	167,386		
Norwich	180,881	London	
Aylesbury & Wycombe	181,544	London	3,462,107

Table A2
One-digit SOC classification.

Code	Description
1	Managers and senior officials
2	Professional occupations
3	Professional and technical occupations
4	Administrative and secretarial occupations
5	Skilled trades occupations
6	Personal service occupations
7	Sales and customer service occupations
8	Process, plant and machine operatives
9	Elementary occupations

References

Audretsch, D., Feldman, M., 1996. R&D spillovers and the geography of innovation and production. Am. Econ. Rev. 86 (3), 630–640.

Baum-Snow, N., Pavan, R., 2012. Understanding the city size wage gap. Rev. Econ. Stud. 79, 88–127.

Carlsen, F., Rattsø, J., Stokke, H., 2013. Education, experience and dynamic urban wage premium. Norwegian University of Science and Technology working paper.

Chan, S., Stevens, A., 2004. How does job loss affect the timing of retirement? The B.E. J. Econ. Anal. Policy 0 (1) 5..

Charlot, S., Duranton, G., 2004. Communication externalities in cities. J. Urban Econ. 56 (3), 581–613.

Combes, P.-P., Duranton, G., Gobillon, L., 2008. Spatial wage disparities: sorting matters! J. Urban Econ. 63, 723–742.

De la Roca, J., 2011. Selection in initial and return migration: evidence from moves across Spanish cities. IMDEA working paper.

De la Roca, J., Puga, D., 2014. Learning by working in big cities. http://diegopuga.org/papers/esurban.pdf.

Di Addario, S., Patacchini, E., 2008. Wages and the city. Evidence from Italy. Labour Econ. 15, 1040–1061.

Duranton, G., Puga, D., 2001. Nursery cities: urban diversity, process innovation, and the life cycle of products. Am. Econ. Rev. 91 (5), 1454–1477.

Duranton, G., Puga, D., 2004. Micro-foundations of urban agglomeration economies. In: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Urban and Regional Economics* 4. Elsevier, Amsterdam, pp. 2063–2117.

Fielding, A., 1989. Inter-regional migration and social change: a study of South East England based upon data from the longitudinal study. Trans. Inst. Brit. Geogr. 14 (1), 24–36 (New Series).

Fielding, A., 1992. Migration and social mobility: South East England as an escalator region. Reg. Stud. 26 (1), 1–15.

Fu, S., Ross, S., 2010. Wage premia in employment clusters: agglomeration or worker heterogeneity? Working Papers 10–04. Center for Economic Studies, U.S. Census Bureau.

Gibbons, S., Overman, H., Pelkonen, P., 2010. Wage disparities in Britain: people or place? SERC/LSE discussion paper 0060.

Glaeser, E., 1999. Learning in cities. J. Urban Econ. 46 (2), 254–277.

Glaeser, E., Maré, D., 2001. Cities and skills. J. Labor Econ. 19 (2), 316–342.

Gould, E., 2007. Cities, workers, and wages: a structural analysis of the urban wage premium. Rev. Econ. Stud. 74 (2), 477–506.

Gutiérrez-i-Puigarnau, E., van Ommeren, J.N., 2010. Labour supply and commuting. J. Urban Econ. 68, 82–89.

Henderson, J. Vernon, 2003. Marshall's scale economies. Journal of Urban Economics, Elsevier 53 (1), 1–28.

Marshall, A., 1920. Principles of Economics. MacMillan, London.

Melo, P., Graham, D., 2009. Agglomeration economies and labour productivity: evidence from longitudinal worker data for GB's travel-to-work areas. SERC/LSE Discussion Paper 31.

Mion, G., Naticchioni, P., 2009. The spatial sorting and matching of skills and firms. Can. J. Econ. 42 (1), 28–55.

Moretti, E., 2004. Human capital externalities in cities. In: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Urban and Regional Economics* 4. Elsevier, Amsterdam, pp. 2243–2291.

Moretti, E., 2010. Local multipliers. Am. Econ. Rev. Pap. Proc. 100, 1–7.

Office for National Statistics, 2012. *Annual Survey of Hours and Earnings, 1997–2011: Secure Data Service Access* [computer file], Colchester, Essex: UK Data Archive [distributor], June 2012. SN: 66893rd Edition. http://dx.doi.org/10.5255/UKDA-SN-6689-2.

Puga, D., 2010. The magnitude and causes of agglomeration economies. J. Reg. Sci. 50 (1), 203–219.

Rosenthal, S., Strange, W., 2004. Evidence on the nature and sources of agglomeration economies. In: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Urban and Regional Economics* 4. Elsevier, Amsterdam, pp. 2119–2171.

Rotemberg, J., Saloner, G., 2000. Competition and human capital accumulation: a theory of interregional specialization and trade. Reg. Sci. Urban Econ. 30 (4), 373–404.

Topel, R., Ward, M., 1992. Job mobility and the careers of young men. Q. J. Econ. 107 (2), 439–479.

Wheeler, C., 2006. Cities and the growth of wages among young workers: evidence from the NLSY. J. Urban Econ. 60, 162–184.

Yankow, J., 2006. Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. J. Urban Econ. 60, 139–161.

Zenou, Y., 2009. Search in cities. Eur. Econ. Rev. 53 (6), 607–624.