



Why does birthplace matter so much?

LSE Research Online URL for this paper: <http://eprints.lse.ac.uk/100201/>

Version: Published Version

Article:

Bosquet, Clément and Overman, Henry G. (2019) Why does birthplace matter so much? *Journal of Urban Economics*, 110. pp. 26-34. ISSN 0094-1190

<https://doi.org/10.1016/j.jue.2019.01.003>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>



Why does birthplace matter so much?

Clément Bosquet^{a,*}, Henry G. Overman^{b,**}

^a ThEMA (University of Cergy-Pontoise) and CEP (LSE), France

^b LSE, CEP and WWC, London, United Kingdom

ARTICLE INFO

JEL codes:

J61
J62
R23
J31

Keywords:

Place of birth
Spatial sorting
Intergenerational transmission
Lifetime mobility

ABSTRACT

We consider the link between birthplace and wages. Using a unique panel dataset, we estimate a raw elasticity of wages with respect to birthplace size of 4.2%, two thirds of the 6.8% raw elasticity with respect to city size. Part of this effect simply reflects intergenerational transmission and the spatial sorting of parents, part is explained by the role that birthplace size plays in determining current city size. Lifetime immobility explains a lot of the correlation between birthplace and current city size: we show that 43.7% of individuals only ever work while living in the place they were born. Our results highlight the importance of intergenerational and individual sorting in helping explain the persistence of spatial disparities.

1. Introduction

The question of links from birthplace to outcomes has long been a concern of the literature that looks at the impact of growing up in a disadvantaged neighbourhood (see e.g. Oreopoulos, 2003; Durlauf, 2004; Topa and Zenou, 2015; Chetty et al., 2016). This paper asks a similar question, but at a larger spatial scale (the local labour market). It contributes to a small, but growing, literature that considers the impact of ‘initial conditions’ on labour market outcomes (see e.g. Aslund and Rooth, 2007; Almond and Currie, 2011). The emphasis on birthplace and intergenerational sorting means the paper is also related to recent works on the geography of intergenerational mobility (Chetty et al., 2014; Chetty and Hendren, 2018a,b).

We consider the link between birthplace size and wages using a unique panel data set (the British Household Panel Survey) which provides information on current location and birthplace for a sample of UK individuals and households questioned annually between 1991 and 2009.¹ We estimate a raw elasticity of wage with respect to birthplace size of 4.2%, two thirds of the 6.8% raw elasticity with respect to current city size. The BHPS also provides information on individual characteristics and a limited set of parental characteristics which allows us to consider why this effect occurs.

Why could birthplace size matter? One possibility is that individual characteristics vary with birthplace size because of the spatial sorting of parents and the intergenerational transmission of characteristics. A second possibility is that birthplace size affects the accumulation of human capital. A third possibility is that birthplace influences past and current city sizes – either through immobility or an effect on migration choices – and, thus, local labour market opportunities. Indeed, in the extreme case of no mobility, birthplace size directly determines labour market size and it makes little sense to try to distinguish between the role of birthplace and of current location.

Our results suggest that intergenerational transmission and the effect of birthplace on current location both play a role in explaining the link between wages and birthplace size. We find no direct role for differences in childhood educational outcomes, other than through the sorting of parents. This highlights the importance of intergenerational transmission and parental sorting in helping explain the persistence of spatial disparities. Low lifetime mobility reinforces the link between the location decisions of generations, which suggests that there is a geographic component to the inheritance of inequality at birth in addition to intergenerational transmission through parental characteristics. We provide descriptive evidence on lifetime mobility that suggests this is an important consideration in the UK: in our data around 43.7% of individuals only ever work while living in the same area as they were born.

* Correspondence to: ThEMA, UMR 8184, Université Cergy-Pontoise, CNRS, F-95000 Cergy-Pontoise. 33 boulevard du Port, 95011 Cergy-Pontoise cedex, France.

** Correspondence to: London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom.

E-mail addresses: clement.bosquet@u-cergy.fr (C. Bosquet), h.g.overman@lse.ac.uk (H.G. Overman).

¹ After cleaning, we have data on around 7000 workers. Given sample size, we follow the literature and focus on the link from city size – birthplace and current location – to wages, rather than on the full set of area effects.

Our paper is closely related to the literature that considers the role of agglomeration economies in explaining spatial disparities. In the urban economics literature sorting – the concentration of more productive workers in more productive locations – plays an important role in understanding disparities across space. For example, [Combes et al. \(2008\)](#) show that, for wages in France, the correlation between average individual fixed effects and area fixed effects is somewhere around 0.3. [Mion and Naticchioni \(2009\)](#) find qualitatively similar results for Italy. Such positive correlation can explain a large part of overall spatial disparities. For example, [Gibbons et al. \(2014\)](#) show that between 85% and 88% of area wage disparities in the UK are explained by individual characteristics (including individual fixed effects). [Combes and Gobillon \(2015\)](#) provide a recent survey and further discussion.

Because this literature uses individual level panel data to estimate area effects from movers across areas, there is a tendency to assume that the ‘sorting’ that explains the concentration of more productive workers in more productive locations is predominantly driven by the mobility decisions of workers. However, it is equally possible that the sorting that explains this concentration is predominantly the result of birthplace variation in individual characteristics combined with low levels of mobility. Indeed, both [Mion and Naticchioni \(2009\)](#) and [Combes et al. \(2012\)](#) show that selective migration accounts for little of the skill differences between dense and less dense areas and suggest a role for ‘sorting at birth’. These birthplace effects could occur directly (e.g. if birthplace size helps determine educational outcomes) or indirectly via the sorting of parents (e.g. if parental characteristics help determine educational outcomes and parental characteristics are correlated with city size). We present evidence consistent with the latter, rather than the former explanation.

2. Data and descriptive statistics

We use the British Households Panel Survey (BHPS) which is a non-balanced panel of households and individuals questioned in 18 waves from 1991 to 2009. The BHPS is based on a nationally representative sample of households recruited in 1991. Panel members comprise all individuals resident at sampled addresses at the first wave of the survey. Subsequent surveys re-interview these individuals annually, following any individuals who split-off from original households (e.g. because of family break-up or because a child enters adulthood and leaves home). All adult members of new households are interviewed, as are new members joining sample households. Children are interviewed once they reach the age of 16. The panel has several advantages. In addition to being representative, it also provides both labour market and geographical information (including birthplace) at a fine level of detail for individuals observed over a relatively long period of time.²

The full sample consists of 32,380 individuals observed on average 7.4 times for a total of 238,996 observations. Available variables cover a variety of topics including education, labour market outcomes, income, health, personal values, labour and life conditions (e.g. workplace characteristics, union membership, family commitments, relationship status, wellbeing). In terms of outcome variable, we focus on total gross pay constructed from self-reported data on ‘usual gross pay per month in current job’. Basic control variables – gender and age – are available for all individuals.³ For parental characteristics we use a measure of social class based on self-reported parental occupations ranging from unskilled to professional occupation with the parents’ highest social class

² Wave-to-wave retention rate averages 91.2%. For fully interviewed individuals (our sample), only 1% on average are not re-interviewed because of loss of contact – which might result from moves within or outside the area. Most of the (relatively small) attrition in the panel is due to refusal, out of scope, non-eligibility or death (BHPS user manual, volume A, section IV.20).

³ We use age based on date of interview and date of birth (as reported age is inconsistent across waves).

constructed as the maximum rank of mother and father.⁴ For individual educational outcomes we construct a measure of qualification based on reported highest educational and academic qualifications. We end up with seven educational dummies: no qualifications; apprenticeship; age 16 qualification low grades; age 16 qualification higher grades; A-level; HNC/D or teaching qualifications; 1st or higher degree.⁵ These are mapped to years of education based on the modal education leaving age for each category. We also have information on the individual’s current occupation classified according to one-digit SOC.⁶

In addition to information on these family and individual characteristics, the data set also provides information on both place of residence and birth. For place of residence we have very precise geographical coordinates (eastings and northings), while place of birth is recorded at the Local Authority District level.⁷ To study spatial sorting across cities we follow much of the existing literature and map these two geographies to local labour markets (constructed from Travel to Work Areas as described in [Gibbons et al., 2014](#)). Given sample sizes – the mean number of workers by area and year is just under 39 – we focus on the effect of birthplace and current city size, measured using the number of people in employment.⁸

One disadvantage of the data is that we only have information on where people live, rather than where they work. This is unfortunate, because the existing agglomeration literature is mainly concerned with the link from workplace size to wages. In practice, this is not a major problem because Travel to Work Areas, our underlying geography, are constructed to ensure that at least 75% of the area’s resident workforce work in the area and at least 75% of the people who work in the area also live in the area. Consistent with this, as we report below, we get estimates of the elasticity of wages with respect to current city size that are broadly in line with the existing literature.

Given small sample sizes, we drop individuals who were born outside of Great Britain (including those born, or currently located, in Northern Ireland). As our main focus is on wage disparities, we also drop observations corresponding to years in which the individual is studying, unemployed or retired. Concerns over self-reported hours lead us to focus on the total pay for full-time workers, although our results are robust to considering all workers. To allow us to include a reasonable set of observable characteristics, we drop individuals with missing occupation,

⁴ From the lowest to the highest social class the categories of occupation are as follows: unskilled, partly skilled, skilled manual, armed forces, skilled non-manual, managerial and technical, and professional occupations.

⁵ Age 16 qualifications are taken at the end of compulsory schooling, A-levels at the end of schooling (age 18). HNC is a Higher National Certificate, HND a Higher National Diploma, usually involving one or two year’s study post-18, respectively. Most UK 1st degrees involve three years post-18 study.

⁶ Managers and Senior Officials, Professional Occupations, Professional and Technical Occupations, Administrative and Secretarial Occupations, Skilled Trades Occupations, Personal Service Occupations, Sales and Customer Service Occupations, Process, Plant and Machine Operatives, and Elementary Occupations.

⁷ Eastings and northings are the reference system used in the British National Grid providing location information rounded to the nearest one meter. Local Authorities Districts are the 326 sub-national divisions of England used for local government and the equivalent districts in Scotland and Wales.

⁸ Birthplace and current city sizes are from the closest census year (1971, 1981, 1991, 2001, 2011) aggregated from TTWA level data constructed by [Amior and Manning \(2017\)](#). Online appendix Table O7 provides descriptives for local labour market size. Historical city sizes (used as instruments) are constructed by mapping LAs to TTWA and using historical LA data from the Vision of Britain project which uses the UK census mapped to stable-across-time LA boundaries. See <http://www.visionofbritain.org.uk/data/>. LAs are mapped to TTWAs using area share weightings. More information is available on request. Results available on request show that all results in the paper are robust to matching to specific years with linear interpolations between census years, to only one specific year or to using population instead of employment as the measure of city size.

Table 1
Descriptive statistics for full-time workers.

| Variable | (1) | (2) | (3) | (4) | (5) |
|----------------------------|---------|---------|---------|---------|---------|
| Women (%) | 45.9 | 46.1 | 46.1 | 45.9 | 44.7 |
| Age | 35.1 | 34.9 | 34.9 | 37.6 | 38.3 |
| Gross pay | 1487 | 1490 | 1490 | 1586 | 1649 |
| Occupation (%) | | | | | |
| Managers/Senior Officials | 14.1 | 14.1 | 14.1 | 15.3 | 16.1 |
| Professional Occupations | 9.7 | 9.9 | 9.9 | 10.9 | 11.5 |
| Professional & Technical | 11.6 | 11.6 | 11.6 | 12.3 | 12.7 |
| Admin & Secretarial | 17.8 | 17.9 | 17.9 | 17.5 | 17.1 |
| Skilled Trades | 11.7 | 11.7 | 11.7 | 11.2 | 11.3 |
| Personal Service | 11.3 | 11.2 | 11.2 | 10.3 | 9.7 |
| Sales and Customer Service | 6.6 | 6.6 | 6.6 | 5.7 | 5.4 |
| Machine Operatives | 10.5 | 10.3 | 10.3 | 10.6 | 10.4 |
| Elementary | 6.7 | 6.6 | 6.6 | 6.2 | 5.7 |
| Location | | | | | |
| Resident city size | 504,951 | 507,542 | 507,731 | 488,481 | 475,576 |
| Live in city (%) | 70.7 | 70.6 | 70.7 | 69.6 | 69.6 |
| Live in London (%) | 7.8 | 7.9 | 7.9 | 7.5 | 7.1 |
| Birth city size | 587,361 | 586,162 | 585,722 | 596,379 | 603,166 |
| Born in city (%) | 75.0 | 74.9 | 74.9 | 74.2 | 74.4 |
| Born in London (%) | 9.4 | 9.4 | 9.4 | 9.5 | 9.7 |
| Number of observations | 72,580 | 70,045 | 70,025 | 57,125 | 55,382 |
| Number of individuals | 12,698 | 12,370 | 12,364 | 9,243 | 7,500 |

Source: Authors own calculation based on BHPS. Notes: Gross pay data are monthly and have been deflated using a consumer price index (base year = 2005). Occupations classified according to one-digit SOC.

education and parents' highest social class, after extrapolating and interpolating from existing data where appropriate. This leaves us with 57,125 observations for 9243 individuals. Finally, when using the panel dimension of the data (with individual fixed effects), we keep only workers observed at least twice. This leaves us with 55,382 observations for 7500 individuals. This is our minimum sample size although, as will become clear below, we can use larger samples in some of our estimations when the full set of restrictions need not apply.

Descriptive statistics are provided in Table 1. Column (1) presents descriptive statistics for the sample of full-time workers restricted based on country of birth (dropping those born outside Great Britain, including in Northern Ireland) and dropping individuals who are studying, unemployed or retired. The focus on full-time workers leads to women being slightly under-represented in the total sample. Gross (monthly) pay figures deflated to 2005 base year look broadly in line with those reported from the Annual Survey of Hours and Earnings (and before that from the New Earnings Survey). Average city size is larger for birthplace than for current residence – explained by the fact that both birth-rates and immigrant shares are higher in larger cities (so more UK citizens are born in big cities, than live in those cities).⁹ Column (2) shows what happens when we drop individuals with missing education, column (3) additionally drops those with missing occupation and column (4) those with missing parent's highest social class. Finally, column (5) keeps only full-time workers observed at least twice – the sample that we use when including fixed effects to exploit the panel dimension of the data. As is to be expected, these restrictions slightly skew the sample towards those with higher incomes and occupations associated with higher education levels – particularly when dropping individuals with missing highest parent social class and individuals observed only once. But none of the changes are particularly large. In short, to the extent the initial sample is representative, restricting on observable characteristics does not significantly affect the representativeness of our final sample.

⁹ <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/populationdynamicsofukcityregionssince2011/2016-10-11>.

3. Econometric strategy

We now outline the way in which we estimate the effect of both current location and birthplace on individual wages. As discussed above, given sample sizes, our focus is on estimating the effect of city size, rather than the full set of birthplace and current city effects. As birthplace size is fixed we use a two-step estimator which allows us to control for individual heterogeneity using fixed effects in the first step.

To see why this is necessary, start by considering a simple 'one-step' method for assessing how outcomes vary with birthplace size. Denoting the (log of the) wage of individual i at date t as $w_{i,t}$ we can regress:

$$w_{i,t} = \gamma B_{a(i)} + \rho' PX_i + \beta' X_{it} + \lambda R_{a(i),t} + \varepsilon_{i,t} \quad (1)$$

where $B_{a(i)}$ is the (log of the) size of area a where individual i is born (calculated as described in Section 2), PX_i are parental characteristics, X_{it} are time varying individual characteristics, $R_{a(i),t}$ measures the (log of the) size of the area a of residence of individual i at time t , γ , λ , ρ and β are (vectors of) coefficients. The coefficient γ is the main object of interest and captures the elasticity of wages with respect to birthplace size (λ captures the same elasticity for current city size). As discussed in Section 2, we have relatively limited data on parental characteristics – we use a measure of social class based on self-reported parental occupations. For individual controls, we have data on individual age, gender, educational outcomes and occupation. Finally, we include current place of residence to account for the effect of city size on wages as documented in the agglomeration literature.

While this 'one-step' estimator is intuitive, it leads to inconsistent estimates of γ , ρ , β and λ if individual unobserved characteristics are correlated with current city size, a fact that is well established in the economic geography literature (see Combes and Gobillon (2015) for a review). Even if these unobserved characteristics are uncorrelated with birthplace size (after conditioning on parental characteristics), correlation between current city size and birthplace size will still render estimates of γ inconsistent. This is, of course, true more generally for any correlation between unobserved characteristics and the included right-hand side variables. Our emphasis on the correlation with current or birthplace city size simply reflects the fact that this fits with the

substantive focus of the paper and that these correlations are central to, and well-evidenced in, much of the recent urban economics literature on sorting and agglomeration effects.

To overcome this problem, we adopt a two-step econometric strategy in the spirit of Combes et al. (2008). In the first step, we regress wage of individual i living in area a at date t on an individual fixed effect θ_i , time-varying observable characteristics X_{it} , an area size effect $R_{a(i),t}$ and a time fixed effect δ_t :

$$w_{i,t} = \theta_i + \beta' X_{it} + \lambda R_{a(i),t} + \delta_t + \varepsilon_{i,t} \quad (2)$$

In the second step, we then regress the estimated individual fixed effects on time-invariant characteristics (Z_i) and birthplace size:

$$\hat{\theta}_i = \gamma B_{a(i)} + \alpha' Z_i + \eta_i \quad (3)$$

where $B_{a(i)}$ is defined as in Eq. (1), Z_i includes gender, education and parental characteristics, η_i is the error term and γ and α are (vectors of) coefficients. As with Eq. (1) the coefficient γ is the main object of interest and captures the elasticity of wages with respect to birthplace size.

If time variant unobserved shocks are uncorrelated with $R_{a(i),t}$, we can use the panel dimension of our data to estimate (2) to provide a consistent estimate of the coefficient on $R_{a(i),t}$. If we further assume that $E[\eta_i | B_{a(i)}, Z_i] = 0$ then this two-step procedure also provides us with consistent estimates of the effects of birthplace and parental characteristics. Note that, if we use a fixed measure of city size λ is estimated from movers in Eq. (2) (current city size is perfectly correlated with individual fixed effects for stayers), but conditional on having an estimate $\hat{\lambda}$, γ in Eq. (3) is estimated from all workers. It captures the elasticity of wages with respect to birthplace size after controlling for parental and individual characteristics, including current city size.

In a recent paper, De la Roca and Puga (2017) suggest that we should be careful to distinguish between static and dynamic agglomeration economies when estimating wage equations of the kind we use in our first step (i.e. Eq. (2)). If adult learning is important, De la Roca and Puga show that we should control for the whole labour market history when assessing the impact of current city size. In their estimation, they consider a full set of area effects so allowing for the effect of adult learning involves the introduction of city-specific experience variables in their estimated equation. In our specification with only city size on the right hand side, this equates to including a variable that captures accumulated city size (up to and including the period before the current observation) in the first-step estimation. That is, we can estimate:

$$w_{i,t} = \theta_i + \beta' X_{it} + \lambda R_{a(i),t} + \theta \sum_{t=t_0}^{t-1} R_{a(i),t} + \varepsilon_{i,t} \quad (2a)$$

where the summation captures accumulated city size from the time that the individual entered the labour market (t_0) until the period before the current observation. Following De-la-Roca and Puga, we restrict the summation to periods where the individual is working so that it has the interpretation of accumulated experience. As a result, this variable is equal to the product of the average past city size multiplied by the number of years the individual is present in the panel and working.

We present results using both the static and dynamic first-step specifications in what follows. As we discuss further below, once we recognise that birthplace size can affect outcomes, and that mobility rates are low, this further increases the difficulty of separately identifying the effect of current city size from accumulated experience.

If all the effect of birthplace size is explained by observed characteristics (of parents and individuals, including workplace location) then γ (in Eq. (3)) should be 0. It is important to note, however, that if differences in these observed outcomes are partly explained by birthplace size then controlling for them will lead us to underestimate the coefficient on birthplace size. In contrast, omitting these variables may lead us to overestimate the coefficient on birthplace size – e.g. if spatial sorting

of parents based on unobservable characteristics leads to variations in educational outcomes by city size.

An additional estimation issue arising for both the static and dynamic specifications concerns the endogeneity of city size. To address this concern, we follow the existing literature and use long lags of historical population from 1801 to instrument for current city sizes (measured using constant (2001) city size to reflect the fact that we only have one instrument).

4. Results

We start with results for standard agglomeration regressions of wages on current city, rather than birthplace, size.¹⁰ These results, reported in Table 2, are interesting in two regards. First, because they provide an estimate of the elasticity of wages with respect to city size based on our BHPS data. Second, because they constitute the first-step estimates that we use in our two-step analysis.

The estimate of the elasticity of wages with respect to city size is around 6.8% when we control only for gender and age, falling to 4.6% as we add individual level controls for, education (column 2) and occupation (column 3). Results reported in column (4) show that this coefficient is roughly halved once we use the panel dimension of our data and include individual fixed effects. Both the point estimates, and the changes in coefficients as we include observable and unobservable characteristics, are broadly in line with the findings from the existing agglomeration literature.

Column (5) shows what happens when we follow De la Roca and Puga (2017) and distinguish between static and dynamic agglomeration economies, by including variables to capture accumulated experience.¹¹ Finally column (6) reports estimates when we instrument for city size and learning using long lags of historical population dating from 1801. First stage regressions are reported in Table O1 of the online appendix. The instrumenting for current city size is completely standard. For learning, we construct the instrument by aggregating historical city sizes from the time that the individual entered the labour market – analogous to the expression for learning based on contemporaneous city sizes introduced in Eq. (2a).

To obtain the elasticity of wages with respect to birthplace size, we switch to two-step estimation. The one-step results, available in Table O2 of the online appendix suggest that the elasticity of wages with respect to birthplace ranges from around 1.2% to 3.8%. As explained in Section 3, however, while the one-step results are easy to interpret, estimates of the coefficient on birthplace size are biased if unobserved ability is correlated with current city size. Switching to two-step estimation allows us to (partially) address this concern subject to the

¹⁰ All results reported in this section are robust to changes in the sample, including: using all workers or only lifetime movers, trimming top and bottom 1% of wages, only estimating on workers born 1966 onwards (as our city size data begin in 1971 and we match to nearest census year). As we only have one measure of historical city sizes, IV uses a time invariant measure of current city size based on total employment in 2001. Instrumenting does not change the city size only coefficient (column 4) and changing the timing of instruments does not change the results when instrumenting for city size and learning (column 6). For comparability, we use the same time invariant measure of city size for OLS. Results are robust to using time varying city size, a time invariant measure of birthplace size based on total employment in 1971, or using total population, instead of employment, as a measure of birthplace and current city size. Finally, results are robust to only estimating on individuals for whom we observe birthplace or learning. All these results are available on request.

¹¹ The number of individuals is smaller because learning is accumulated city size until $t-1$, so (with individual fixed effects) we need to observe individuals at least 3 times for them to be included in the sample used to estimate columns (5) and (6). We also lose the first observation for these individuals as, by definition, learning is not defined in the first period in which the individual is observed. Results available on request show that columns (1)–(4) are robust to the restriction of the sample to observations for which learning is observed.

Table 2
First-step regressions of (log) gross total wage on city size and controls (full-time workers only).

| | (1) OLS | (2) OLS | (3) OLS | (4) OLS | (5) OLS | (6) IV |
|-----------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| (log) City size | 0.068*** (0.003) | 0.048*** (0.003) | 0.046*** (0.003) | 0.025*** (0.003) | 0.008** (0.004) | 0.009** (0.005) |
| Learning | | | | | 0.060*** (0.004) | 0.061*** (0.004) |
| Time FE | X | X | X | X | X | X |
| Gen., Age, Age ² | X | X | X | X | X | X |
| Education | | X | X | X | X | X |
| Occupation | | | X | X | X | X |
| Individual FE | | | | X | X | X |
| Observations | 74,242 | 74,242 | 74,242 | 74,242 | 63,317 | 63,317 |
| R-squared | 0.320 | 0.445 | 0.512 | 0.846 | 0.853 | 0.853 |
| Number of ind. | 10,545 | 10,545 | 10,545 | 10,545 | 8,923 | 8,923 |

Source: Authors own calculation based on BHPS. *Notes:* Standard errors clustered at the current city by time level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. City size is current city size measured using total employment in 2001; learning is log accumulated city size and calculated as explained in the text. In column (6), city size and learning are instrumented with city size in 1801 and accumulated 1801 city sizes. The Kleibergen-Paap rank Wald F statistics is 3571, suggesting no weak-instrument concerns. Education is defined using seven educational dummies, while occupation uses nine dummies based on one-digit standard occupational classification (SOC). See Section 2 for details. In columns (4)–(6), gender and education are time invariant and absorbed by individual fixed effects.

Table 3
Second-step regressions for gross total wage; individual fixed effects on birthplace size and controls (full-time workers only).

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|-----------------------|---------------------|---------------------|---------------------|---------------------|--------------------|------------------|-------------------|
| (log) Birthplace size | 0.042*** (0.009) | 0.035*** (0.008) | 0.033*** (0.010) | 0.033*** (0.009) | 0.023** (0.009) | 0.000 (0.008) | −0.001 (0.008) |
| 1st-step controls | | | | | | | |
| Time FE | X | X | X | X | X | X | X |
| Age, Age ² | X | X | X | X | X | X | X |
| Occupation | | | | X | X | X | X |
| (log) City size | | | | | X | X | IV |
| Learning | | | | | | X | IV |
| 2nd-step controls | | | | | | | |
| Gender, birth year | X | X | X | X | X | X | X |
| HPSC | | X | X | X | X | X | X |
| Education | | | X | X | X | X | X |
| Observations | 7,500 | 7,500 | 7,500 | 7,500 | 7,500 | 6,472 | 6,472 |
| R-squared | 0.146 | 0.201 | 0.333 | 0.322 | 0.319 | 0.466 | 0.468 |

Source: Authors own calculation based on BHPS. *Notes:* Standard errors clustered at the birthplace level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Birthplace size is measured using the number of people in employment from the closest census year. Occupation, education, city size and learning are all defined as in the notes to Table 2. HPSC is Highest Parental Social Class. In column (7), city size and learning are instrumented with city size in 1801 and accumulated 1801 city sizes in the first step. See Section 2 for further details. For these second-step estimates, the number of observations corresponds to the number of individuals because the dependent variable is the individual fixed effects estimated in the first step.

caveats discussed in Section 3.¹² Results for the first-step regressions have already been reported in Table 2, whilst results for the second-step regressions are reported in Table 3. Comparing the birthplace elasticity in column (4) of Table 3 (0.033), with its one-step equivalent (0.026) shows that we underestimate the coefficient on birthplace size if we ignore the correlation between unobserved ability and current city size.¹³

¹² We put time varying variables – time fixed effects, age, age squared, occupation, current and accumulated city size (learning) in the first stage. Time invariant variables – gender, highest parent social class (HPSC), education and the cohort effect (birth year) go in the second stage. We can control for age in the first stage because individuals are not interviewed at the same date every year. Results available on request show that the findings are robust to controlling only for age squared in the first stage and to using Weighted Least Squares with the inverse of individual fixed effect variance as the weights in the second stage.

¹³ To provide a clean comparison, Table O3 in the online appendix reproduces Table O2 for the sub-sample of individuals used in this second-step regression (Table 3), controlling for birth year. Results are robust.

The raw elasticity of wages with respect to birthplace size, controlling only for demographic characteristics (column (1) of Table 3), is 4.2%. This suggests that individuals born in London in 1971 make on average 6.6% more than individuals born in Manchester (the second largest TTWA in the UK) and 9.3% more than individuals born in Liverpool (the sixth largest).¹⁴ Columns (2)–(6) of Table 3 allow us to consider how different mechanisms explain this correlation between birthplace size and wages. The interpretation of these results obviously depends on the order of introduction of these variables, an issue we discuss in depth in Section 6. In this section, controls are introduced in the natural ordering of life events. We hence start by including controls for parental social class – a family characteristic that is clearly pre-determined for individuals in the sample used for estimation. Results are reported in column (2) and show that the coefficient on birthplace size is reduced by around 20%, reflecting the fact that some of the correlation between birthplace size and wages is explained by the sorting

¹⁴ $(4,084,810/882,333)^{0.042} - 1$ and $(4,084,810/493,218)^{0.042} - 1$, respectively.

of *parents* across places of different sizes.¹⁵ Given what we know about intergenerational transmission (see, e.g., Black and Devereux, 2011 for a review), this suggests that higher parental social class must be positively correlated with city size. Table O4 in the online appendix reports several descriptive statistics that suggest that this is indeed the case.

Column (3) shows what happens once we introduce individual education as an additional control. The coefficient on birthplace size is almost unchanged, suggesting that the correlation between birthplace size and wages does not work through own educational outcome (once we control for parental characteristics). Consistent with this, results in Table O5 in the online appendix show that there is a positive significant correlation between birthplace size and years of education but that this effect disappears once we control for parental social class. These results only look at education quantities, which leaves open the question of whether variations in school quality across different sized cities could help explain our results. To consider this, we tested whether the returns to schooling varied by birthplace size by interacting the qualification dummies with birthplace size. We generally found no significant effects for qualifications obtained at the end of compulsory schooling.¹⁶ Coefficients are also insignificant if we, instead, include the interaction of years of education with birthplace size.

Controlling for own occupation (column 4) similarly has little effect. In contrast, controlling for current city size (column 5) has a substantial impact on the birthplace size elasticity reducing it further from 3.3% to 2.3%.¹⁷

In column (6) of Table 3 we allow for adult learning by introducing cumulated experience. As is clear from results in column (5) of Table 2, allowing for learning makes a big difference in terms of the estimated effect of current city size on wages. In turn, this makes a big difference to our estimates of the coefficient on birthplace size, as shown in the second-step results reported in column (6) of Table 3.¹⁸ This suggests a third mechanism that explains the correlation between wages and birthplace size: specifically, it determines the amount of time spent in large cities which increases wages via the effect of adult learning in big cities. As discussed in Section 3, the fact that the elasticity of wages with respect to birthplace size is now 0 suggests that all the effect of birthplace size is explained by the included observed characteristics. Finally, once again, column (7) shows that instrumenting in the first-step for current city size and learning using historical population (in levels and aggregated from first labour market entry, respectively) makes no substantive difference to our results.

The results in Table 3 are estimated using all workers. As discussed above while the effect of current city size is identified from movers, *conditional* on having an estimate of current city size the effect of birthplace is identified from all workers for whom we can calculate individual fixed effects (i.e. those with more than one observation). Appendix A shows what happens when we estimate Table 3 only for lifetime movers. The broad pattern of results in terms of changes to the coefficient on birthplace size is in line with those reported in Table 3. Two differences do emerge, however. First, the initial correlation between birthplace size and individual fixed effects (i.e. before controlling for any individual or family characteristics) is somewhat weaker for lifetime movers than for the sample as a whole. Second, as we move across columns the coefficients reduce less quickly for lifetime movers and some residual effect

¹⁵ A Wald test suggests that the change in coefficient from 0.042 (0.009) to 0.035 (0.008) is statistically significant at the 1% level.

¹⁶ These are CSE or O-level qualifications (depending on cohort) taken at age 16. In some specifications, we find evidence of a positive effect of birthplace size on the returns for some post-compulsory schooling qualifications (A levels and degrees) and negative effects for other post-compulsory schooling qualifications (apprenticeships). Results are available upon request.

¹⁷ A Wald test suggests that the change in coefficient from 0.033 (0.009) to 0.023 (0.009) is statistically significant at the 1% level.

¹⁸ A Wald test suggests that the change in coefficient from 0.023 (0.009) to 0.000 (0.008) is statistically significant at the 1% level.

of birthplace size remains. That is, the overall effect of birthplace size is less important for movers. But in contrast to stayers observable characteristics of individuals cannot completely explain the birthplace size effect. As we will see below, both these differences can be explained by the fact that birthplace perfectly determines current city size for stayers (making the overall effect of birthplace larger for stayers, but also consistent with the fact that controlling for current city size explains more of that bigger birthplace coefficient).

To summarise, results so far suggest an elasticity of wages with respect to birthplace size of around 4.2%. Some of this correlation is capturing the sorting of *parents* across places of different sizes. Once we control for this parental sorting, own educational outcome does not play much of a role in explaining the correlation between wages and birthplace size, and neither does occupation. In contrast, the fact that birthplace determines current city size plays an important role via the effect of static and dynamic agglomeration economies on wages. This pattern is broadly similar for both movers and non-movers. These results suggest an important role for lifetime mobility in explaining the link between birthplace and wages. We consider this in more detail in the next section.

5. Geography and lifetime mobility patterns

The results in Table 3 make clear that the most substantial reduction in the coefficient on birthplace size occurs when we control for current and accumulated city size. Once we control for these, along with individual controls, we completely account for the link between birthplace size and wages. Consistent with the agglomeration literature, we know from Table 2 that current and accumulated city size both have a positive effect on wages. This suggests that the reduction in the coefficient on birthplace size when current and accumulated city size are added as controls occurs because of a positive correlation between birthplace size and the size of cities where individuals work as adults. Consistent with this, results in Table O6 of the online appendix, show a strong positive relationship when regressing current city size on birthplace size. For movers the correlation is still positive, albeit weaker than for the full sample. This helps explain why the coefficient on birthplace size is similar (although slightly smaller) when we focus only on lifetime movers. Thus, while low lifetime mobility does not fully explain the positive coefficient on birthplace size it is an important mechanism through which birthplace, via its effect on current and accumulated city size, affects wages.

Because the BHPS provides information on both current location and place of birth, we can use it to assess the extent of lifetime mobility in Britain. We ignore mobility for non-work-related reasons – such as study or retirement – and focus on the share of workers who have only ever worked while living in the same place as they were born. The first row in Table 4 shows the overall figures and then broken down by qualification. As the table shows, over 40% of workers have only ever worked in the place where they were born. The breakdown by qualification shows that these figures are decreasing with education level – consistent with the wider literature on the relationship between education and mobility.¹⁹

The next 4 rows show the figures broken down by the type of area in which the individual was born. The figures provide evidence that mobility also varies with birthplace size – although the major difference is observed in the larger lifetime mobility away from rural areas. The pattern with respect to qualifications is repeated across area types. The final two rows consider similar figures but now focus on whether someone was born in the same place of birth as their parents (these figures are calculated for a sub-set of the 5361 individuals for whom we observe both parent and individual birthplace). These figures are higher than for the percentage of individuals who have always worked where they

¹⁹ For example, Diamond (2016) documents that 67% of US citizens live in their birth state, the figure being only 50% for college graduates.

Table 4
Lifetime mobility: Various indicators based on area where work, area where parents born, area at different ages, by skills.

| % always worked where born | Total | No quals. | GCSE eq. | A-level eq. | Degree |
|---|-------|-----------|----------|-------------|--------|
| Total | 43.7 | 51.8 | 48.7 | 45.8 | 30.5 |
| Born in | | | | | |
| Rural | 33.2 | 40.7 | 37.9 | 32.9 | 21.4 |
| Small city | 46.5 | 52.0 | 53.5 | 51.7 | 29.2 |
| Medium city | 45.1 | 57.1 | 49.5 | 48.6 | 28.9 |
| Large city | 48.8 | 57.2 | 53.8 | 50.3 | 37.2 |
| % born same place as (all individuals): | | | | | |
| Mother born | 53.8 | 63.0 | 56.2 | 50.5 | 49.9 |
| Father born | 52.8 | 56.7 | 56.7 | 50.1 | 48.8 |
| % live in area where born at age | | | | | |
| 16 | 61.0 | 59.9 | 60.7 | 69.8 | 69.1 |
| 18 | 57.5 | 62.5 | 60.5 | 52.4 | 63.4 |
| 21 | 45.9 | 59.4 | 53.2 | 40.7 | 38.1 |
| 65 | 44.2 | 52.8 | 40.8 | 41.8 | 27.8 |

Source: Authors own calculation based on BHPS. *Notes:* Areas correspond to Local Labour Market Areas and are classified either as rural or urban with urban further divided in to large cities (employment greater than 260,000), medium cities (employment 130,000–260,000) and small cities (employment smaller than 130,000). See online appendix Table O7 for further details. Education is classified based on an aggregation of the highest educational and academic qualifications variables. GCSE qualification includes those with O-level and CSE; A-level includes those with HND, HNC or teaching qualifications; Degree includes both 1st and higher degree. First panel based on all workers; second panel based on all individuals for whom we observe at least one parent's birthplace; third panel based on all individuals.

were born. This is partly explained by the fact that lifetime mobility is increasing with age (and that people tend to have children when they are younger). But the degree of intergenerational persistence in place of birth is still striking.

Consistent with this, the bottom panel of Table 4 shows that the aggregate lifetime mobility figures hide substantial heterogeneity with respect to age. The table shows overall lifetime mobility at four particular cut-offs – age 16 (compulsory schooling age), age 18 (end of schooling), age 21 (the age at which most university graduates complete their course) and age 65 (retirement).²⁰ The figures show that nearly 61% of 16 years olds live in the same places as they were born, 57.5% of 18 year olds and 46% of 21 year olds. The full set of figures (available on request) show a gradual decline until age 56, with figures increasing slightly afterwards, suggesting some return migration for retirement.

To summarise, both lifetime immobility and the positive correlation between current city and birthplace sizes for movers play an important role in helping explain the link between birthplace and current city sizes.²¹ In the next section, we consider the relative importance of this effect on current city size and other factors that help explain the role for birthplace size on wages.

6. Decompositions

The previous section considered the role of different observable variables in explaining the correlation between birthplace size and wages. As discussed above, the order in which variables are introduced and the partial correlation between explanatory variables will have implications for the changes in the magnitudes of the birthplace coefficients as we move from specification to specification. Our ordering above was justified by what we know about the sequencing of the different de-

terminants.²² Specifically, for an individual, parental social class is determined 'at birth' and before educational outcomes are determined. In turn, educational outcomes tend to be determined before occupation and city of residence. In this section, we ask what happens if we ignore this information on the sequencing of determinants and instead decompose the correlation between wages and birthplace size in to the contribution from different observable variables.

To assess the relative importance of all the observables of interest (i.e. parental social class, education, occupation, current city size and eventually learning), we implement the decomposition proposed by Gelbach (2016). This allows us to calculate how much of the change in the birthplace coefficient can be attributed to particular observables as we move from a specification which controls for only basic observables (age, gender) to the full specification that includes all observables. The simple decomposition procedure uses the omitted variable bias formula to calculate the share of each observable (or group of variables) in explaining the total change in the coefficient of interest.

Gelbach's methodology is designed for standard one-step regressions so we adapt it to our two-step specification. The technical details are provided in the online appendix. Intuitively, the decomposition works as follows: A variable will explain a large share of the change in the birthplace coefficient if it is 1) highly correlated with wages in a 'full' regression including all control variables and 2) highly correlated with birthplace size in a partial regression where the variable is regressed on birthplace size and basic controls (such as gender and age). For instance, current city size will explain a large share of the change in the birthplace coefficient if it is highly correlated with wages in the full regression (conditional on individual fixed effects and observable characteristics; as shown in Table 3) and with birthplace size in the partial regression (allowing for basic controls; as shown in Table O6 in the online appendix). Conversely, occupation will not explain a large share of the change in the birthplace coefficient if, as is the case, it is weakly

²⁰ Note that these figures are calculated for all individuals, rather than focusing on mobility for work (which would make no sense for many 16–21 year olds who are still in education and thus outside the labour force).

²¹ Including learning effects places a much stronger weight on the full set of adult local labour market decisions and reduces estimates of the coefficient on birthplace size. The correlation of current and birth city size for movers becomes more important once we allow for accumulated city size. This highlights the difficulties of separately estimating dynamic (i.e. learning) and static agglomeration economies in situations where a relatively large proportion of workers are immobile. See D'Costa and Overman (2014) for further discussion.

²² Results available upon request show that some changes in birthplace coefficient are robust to the order of introduction of variables. For instance, occupation has little effect on the coefficient (even if introduced first), while current city size has roughly the same effect on the coefficient if we introduce it before education and parental social class. Interestingly, swapping the order of education and parental social class does make a difference, with the change in the coefficient larger if education is introduced first (as opposed to second). This is consistent with the discussion in the text on the decomposition of the change in the birthplace coefficient.

Table 5
Decomposition of the total effect of birthplace size, with OLS in the first step.

| Variables | Without learning | | | With learning | | | From Table 3 |
|-----------------|------------------|---|-----------------------------------|----------------|---|-----------------------------------|--|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | $\hat{\delta}$ | $\hat{\delta}/(\hat{\beta}_{base} - \hat{\beta}_{tot})$ | $\hat{\delta}/\hat{\beta}_{base}$ | $\hat{\delta}$ | $\hat{\delta}/(\hat{\beta}_{base} - \hat{\beta}_{tot})$ | $\hat{\delta}/\hat{\beta}_{base}$ | $(\hat{\beta}_{c-1} - \hat{\beta}_c)/\hat{\beta}_{base}$ |
| HPSC | 0.0032 | 16.8% | 7.6% | 0.0026 | 6.5% | 6.5% | 16.7% |
| Education | 0.0048 | 25.6% | 11.6% | 0.0043 | 10.9% | 10.9% | 3.7% |
| Occupation | 0.0005 | 2.8% | 1.3% | 0.0005 | 1.3% | 1.3% | 0.7% |
| City size | 0.0102 | 54.2% | 24.6% | 0.0034 | 8.7% | 8.6% | 24.2% |
| Learning | | | | 0.0258 | 65.4% | 65.2% | 54.3% |
| 1st step contr. | 0.0001 | 0.7% | 0.3% | 0.0028 | 7.2% | 7.2% | |
| Total | 0.0189 | 100.0% | 45.4% | 0.0394 | 100.0% | 99.7% | 99.7% |

Source: Authors own calculation based on BHPS. Notes: Occupation, education, city size and learning are all defined as in the notes to Table 2. HPSC is Highest Parental Social Class. See Section 2 for further details. 1st step controls include time fixed effects, age and its square.

correlated with wages conditional on individual fixed effects and observables characteristics or if it is weakly correlated with birthplace size allowing for basic controls. The technique fully explains the change in the birthplace coefficient that is due to all the variables included to capture different mechanisms (i.e. parental social class, education, occupation, current city size and learning), so the sum of all the shares of the change in the birthplace coefficient is equal to one. One might also be interested in decomposing the total effect of birthplace size (rather than the total change in its coefficient once all control variables are included), in which case the sum of all terms in the decomposition would be less than one if the birthplace coefficient is not driven to zero.

Since the decomposition is based on the full specification of wages on all controls, plus on partial regressions of the mechanism variables on birthplace size and basic controls, results do not depend on the order of introduction of control variables in the wage equation. On the other hand, it completely ignores potential causal influences of some variables on other variables (e.g. parental social class might have a causal impact on education). As a consequence, the decomposition will tend to underestimate the importance of variables that might influence other variables (e.g. parental social class).

In Table 5 below, the last row of columns (1) and (4), reports $\hat{\delta}_{total}$ the total change in the birthplace coefficient as we move from a regression including only basic controls (gender and age) to a regression including full controls. For each row, columns (1) and (4) report $\hat{\delta}_i$ the contribution of each mechanism variable i (or group of dummies in the case of HPSC, education and occupation) to the total change in the birthplace coefficient. Columns (2) and (5) report this as the share of each mechanism in the total change in the coefficient $(\hat{\beta}_{base} - \hat{\beta}_{tot})$ where $\hat{\beta}_{base}$ is the raw birthplace coefficient of 4.2% reported in column (1) of Table 3 and $\hat{\beta}_{tot}$ is the birthplace coefficient once we control for all mechanisms. In the left hand side of Table 5, which does not include learning, this is 2.3% from column (5) of Table 3, while in the right hand side of Table 5, including learning, this is 0.0, from columns (6) of Table 3.²³ Finally, columns (3) and (6) report the decomposition of the total effect of birthplace size.

For comparison, column (7) shows how much of the change in the birthplace coefficient we attribute to each set of observables if we use information on the sequencing of the determinants: each row reports the change of the birthplace coefficient between consecutive columns of Table 3 relative to the raw birthplace coefficient $((\hat{\beta}_{c-1} - \hat{\beta}_c)/\hat{\beta}_{base})$. This method naturally attributes more weight to parental social class which is controlled for first and less weight to education that does not play a big role conditional on parental social class as explained in Section 4.

The left-hand side of Table 5 shows that, without controlling for learning, current city size (i.e. living place) explains the biggest share

(around one fourth) of the birthplace elasticity, which is consistent with the existence of significant agglomeration effects combined with low mobility documented in the previous subsection. Educational attainment and parental social class explain 11.6% and 7.6% of the raw birthplace elasticity, respectively. Occupation plays a much smaller role.

As noted above, the decomposition tends to underestimate the importance of variables that influence other variables. As shown in Table O5 in the online appendix, education is not correlated with birthplace size, conditional on parental social class. But because education is more correlated with wages than is parental social class, the decomposition (which ignores the role of parental social class in explaining education) attributes education greater explanatory power for changes in the birthplace coefficient. In contrast, the results in column (7) impose restrictions on the sequencing of determinants. As parental social class likely impacts educational attainment, whereas the reverse is not possible, we view 11.6% (column 3) as the upper bound for the share of birthplace coefficient explained by education and 3.7% (column 7) as the lower bound. Similarly, the upper and lower bound for the share of birthplace coefficient explained by parental social class are 16.7% (column 7) and 7.6% (column 3).

Note that, in these regressions, where we do not control for learning, around 55% of the total birthplace coefficient is left unexplained. The right-hand side of Table 5 shows that controlling for learning (experience accumulated in larger cities) reduces the unexplained part of the total birthplace coefficient to zero. Learning itself, explains around 66% of the total birthplace coefficient. The estimated shares of the birthplace coefficient due to HPSC and education are quite stable when learning is introduced, but the share due to agglomeration economies (current city size) more than halves to less than 10%, consistent with the findings of De La Roca and Puga (2017).

7. Conclusions

This paper considers the link between birthplace size and wages. We show that there is a positive correlation between birthplace size and wages and that the magnitude is similar to that of current city size. A number of mechanisms appear to explain (most of) this link between wages and birthplace size. First, birthplace size is linked to parental social class so that some of the link between wages and birthplace size is explained by the sorting of parents. Once we control for parental social class, there appears to be no additional role for education. Second, current city size is correlated with birthplace size creating a link from birthplace to current location. As current city size influences wages (as a result of agglomeration economies) the effect of birthplace on current city size is the second mechanism through which the effect operates. Third, because adult learning matters, the effect on current location provides an additional mechanism because it determines the amount of time spent in large cities which increases wages via the effect of adult learning in big cities. Inertia explains some of these findings: around 40% of workers only ever work while living in the area that they were

²³ Results available upon request show that the decomposition is robust to instrumenting for city size and learning in the first step and that first-step estimates are not affected by this instrumentation.

Table A1
2nd step regressions of individual fixed effects (gross total wage) on birthplace and controls (full-time workers only, lifetime movers).

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|-----------------------|---------------------|---------------------|---------------------|---------------------|---------------------|--------------------|--------------------|
| (log) Birthplace size | 0.029*** (0.005) | 0.021*** (0.004) | 0.022*** (0.004) | 0.021*** (0.004) | 0.019*** (0.004) | 0.009** (0.004) | 0.009** (0.004) |
| 1st-step controls | | | | | | | |
| Time FE | X | X | X | X | X | X | X |
| Age, Age ² | X | X | X | X | X | X | X |
| Occupation | | | | X | X | X | X |
| (log) City size | | | | | X | X | IV |
| Learning | | | | | | X | IV |
| 2nd-step controls | | | | | | | |
| Gender, birth year | X | X | X | X | X | X | X |
| HPSC | | X | X | X | X | X | X |
| Education | | | X | X | X | X | X |
| Observations | 4,393 | 4,393 | 4,393 | 4,393 | 4,393 | 3,839 | 3,839 |
| R-squared | 0.132 | 0.182 | 0.322 | 0.311 | 0.309 | 0.442 | 0.443 |

Notes: See Table 3 in the main text.

born. For at least 60% of individuals, place of birth also identifies the area in which a person grows up. But birthplace also plays a role in determining the future location of movers and our results are not fully explained by inertia.

Further work remains to be done on understanding the mechanisms that explain the link between birthplace size and labour market outcomes and the implications for our understanding of spatial disparities. But, whereas the existing literature has focussed on the role of sorting in adulthood, our results point to the importance of considering other kinds of sorting if we want to fully understand the causes and consequences of spatial disparities.

Acknowledgements

This work was supported by the Economic and Social Research Council grant numbers ES/J021342/1 and ES/G005966/1. This research has also benefited from the project Labex MME-DII (ANR11-LBX-0023-01). We are very grateful to Michael Amior for providing us with the city level census data and to Neil Lee and Yi Liu for help in constructing the historical population instruments. Helpful comments have come from Wolfgang Dauth, Gilles Duranton, Steve Gibbons, Ed Glaeser and from seminar participants at EGIT, SERC, LSE, Geneva, Cergy-Pontoise, Lancaster, NARSC, Nancy, LAGV, PSE RUES, Mannheim, Lille and UEA. The use of these data does not imply the endorsement of the data owner or the UK Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jue.2019.01.003.

Appendix A. Results for lifetime movers

As discussed in the text, our main results are broadly robust to restricting the sample to lifetime movers (footnote 10), although with some differences in the size of the overall birthplace coefficient and the extent to which observable characteristics explain the changes in the coefficient. Results for birthplace size for life-time movers are reported in Table A1 and should be compared to those reported in Table 3 of the main text.

References

- Almond, Douglas, Currie, Janet, 2011. Human capital development before age 5. In: Card, David, Ashenfelter, Orley (Eds.). In: Handbook of Labor Economics, 4B. Elsevier Science, Amsterdam, pp. 1315–1486.
- Amior, Michael, Manning, Alan, 2017. The Impact of Local Demand Shocks on Labour Markets: An Integrated Approach to Commuting and Migration. Mimeo.
- Aslund, Olof, Rooth, Dan-Olof, 2007. Do when and where matter? Initial labour market conditions and immigrant earnings. *Econ. J.* 117 (518(03)), 422–448.
- Black, Sandra E., Devereux, Paul J., 2011. Recent developments in intergenerational mobility. In: Ashenfelter, Orley, Card, David (Eds.). In: Handbook of Labor Economics, 4B. Elsevier Science, Amsterdam, pp. 1487–1541.
- Chetty, Raj, Hendren, Nathan, 2018a. The impacts of neighborhoods on intergenerational mobility I: childhood exposure effects. *Q. J. Econ.* 133 (3), 1107–1162.
- Chetty, Raj, Hendren, Nathan, 2018b. The impact of neighborhoods on intergenerational mobility II: county-level estimates. *Q. J. Econ.* 133 (3), 1163–1228.
- Chetty, Raj, Hendren, Nathan, Katz, Lawrence F., 2016. The effects of exposure to better neighborhoods on children: new evidence from the moving to opportunity experiment. *Am. Econ. Rev.* 106 (4), 855–902.
- Chetty, Raj, Hendren, Nathan, Kline, Patrick, Saez, Emmanuel, 2014. Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Q. J. Econ.* 129 (3), 1553–1623.
- Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, 2008. Spatial wage disparities: sorting matters!. *J. Urban Econ.* 63 (2), 723–742.
- Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, Roux, Sebastien, 2012. Sorting and local wage and skill distributions in France. *Reg. Sci. Urban Econ.* 42 (6), 913–930.
- Combes, Pierre-Philippe, Gobillon, Laurent, 2015. The empirics of agglomeration economies. In: Duranton, Gilles, Henderson, J.Vernon, Strange, William C. (Eds.). In: Handbook of Urban and Regional Economics, 5B. Elsevier Science, Amsterdam, pp. 247–348.
- D’Costa, Sabine, Overman, Henry G., 2014. The urban wage growth premium: sorting or learning? *Reg. Sci. Urban Econ.* 48, 168–179.
- De La Roca, Jorge, Puga, Diego, 2017. Learning by working in big cities. *Rev. Econ. Stud.* 84 (1), 106–142.
- Diamond, Rebecca, 2016. The determinants and welfare implications of US workers’ diverging location choices by skill: 1980–2000. *Am. Econ. Rev.* 106 (3), 479–524.
- Durlauf, Steven N., 2004. Neighbourhood effects. In: Vernon Henderson, J., Thisse, Jacques-François (Eds.). In: Handbook of Urban and Regional Economics, 4. Elsevier Science, Amsterdam, pp. 2173–2242.
- Gelbach, Jonah B., 2016. When do covariates matter? And which ones, and how much? *J. Labor Econ.* 34 (2), 509–543.
- Gibbons, Stephen, Overman, Henry G., Pelkonen, Panu, 2014. Area disparities in Britain: understanding the contribution of people vs. place through variance decompositions. *Oxford Bull. Econ. Statist.* 76 (5), 745–763.
- Mion, Giordano, Naticchioni, Paolo, 2009. The spatial sorting and matching of skills and firms. *Can. J. Econ.* 42 (1), 28–55.
- Oreopoulos, Philip, 2003. The long run consequences of growing up in a poor neighbourhood. *Q. J. Econ.* 118 (4), 1533–1575.
- Topa, Giorgio, Zenou, Yves, 2015. Neighbourhood and network effects. In: Duranton, Gilles, Henderson, J.Vernon, Strange, William C. (Eds.). In: Handbook of Urban and Regional Economics, 5B. Elsevier Science, Amsterdam, pp. 561–624.