

Home versus hometown: what do we mean by spatial sorting?

Clement Bosquet and Henry G. Overman

17/03/2015



ONS disclaimer for data analysis

- “This work is based on data from the British Household Panel Survey produced by the Office for National Statistics (ONS) and supplied by the Secure Data Service at the UK Data Archive. The data are Crown Copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the data in this work does not imply the endorsement of ONS or the Secure 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.”



Overview

Agglomeration economies: What kind of sorting matters?

- Agglomeration economies
 - Productivity increases with city size
- Most productive people live in the most productive places → sorting matters
- But what kind of sorting?
 - Sorting of adults across labour markets
 - Sorting at birth (which generates link between personal characteristics and city size)
- More generally, what is the effect of birthplace on labour market outcomes?

What do we do?

- Use a representative sample of circa 5,000 households followed over 20 years
- Panel data on
 - Labour market outcomes
 - Large number individual and family characteristics
 - Geographical location (current and birth place)
- Two step econometric approach
 - Panel data to estimate agglomeration economies
 - Fixed effects on birth place and other characteristics

Findings

- ~40% only ever work while living in area where born
- For >60%, birthplace is area where person grows up.
- Positive significant effect of birth place size
- Similar in magnitude to effect current city size
- Mechanisms
 - Parental social class → sorting of parents
 - Current city size → link to location decisions
 - *Not* education (once control for parental sorting)
 - Learning matters

Agglomeration and sorting

Agglomeration: Sorting matters

- Agglomeration economies
 - Productivity increases with city size
- Sorting matters [CDG, JUE 2008]
 - Controlling for individual characteristics (observable and unobservable) halves city size effect
 - Positive correlation between average individual fixed effects and area fixed effects

Agglomeration: Does sorting matter?

- Learning matters [De-la Roca; Puga, 2014]
 - Static regression muddles sorting and learning
 - If control for full labour market history then sorting on ability (i.e. unobservables) doesn't matter
- Learning & sorting [D'Costa; Overman 2015]
 - Growth regressions show no link with city size once control for sorting
 - But some evidence labour market history does affect future growth [city experience increases future growth]

What do we mean by sorting?

- Literature focuses on sorting of adults across local labour markets
- But sorting also happens at birth because people grow up in different places
- How does sorting at birth change our understanding of agglomeration economies?

Birth place effects

- Are people disadvantaged because of where they were born?
- Economics of education: focus on neighbourhood effects
- The effect of growing up in different local labour markets (larger spatial scale) has received much less attention

Why could birthplace matter?

- Intergenerational transmissions: spatial sorting of parents
- Better schools and universities in denser cities can make workers born there more productive
- Mobility patterns: birth place affects access to labour market opportunities

Data and Empirical approach

Data: BHPS

- British Households Panel Survey
 - 18 waves from 1991 to 2009
 - Representative sample of households (follow all individuals who leave or join household)
- Full sample: 32,058 individuals, observed on average 7.4 times

Data: Sample restrictions

- Sample restrictions
 - focus on full time workers born in Britain; with characteristic data; living in ‘big enough’ places
 - 56,268 observations on for 9,019 individuals.
- Sample sizes mean that we focus on estimating city size effects rather than area effects

Local labour markets



Wage equation

- Wage of individual i living in area a at date t is:

$$w_{i(a)t} = \theta_i + x'_{it}\beta + \lambda RP_{i(a)t} + \lambda_t + \varepsilon_{it}$$
$$\theta_i = \gamma BP_i + \eta_i$$

with η_i unobserved ability s.t. $E[\eta_i | BP_i] = 0$

One-step estimation

- Inference based on

$$w_{i(a)t} = \gamma BP_i + \rho' PX_i + \beta' X_{it} + \lambda RP_{i(a)t} + \varepsilon_{i(a)t}$$

is biased because

$$E[RP_i | BP_i] \neq 0 \text{ (low lifetime mobility)}$$

$$E[\eta_i | RP_i] \neq 0 \text{ (sorting across labour markets)}$$

Two step estimation

- 1st step:

$$w_{i(a)t} = \theta_i + X_{it}\beta + \lambda RP_{i(a)t} + \delta_t + \varepsilon_{i(a)t}$$

- 2nd step:

$$\theta_i = \gamma BP_i + \eta_i$$

$$\theta_i = \gamma BP_i + \rho' PX_i + \eta_i$$

$$\theta_i = \gamma BP_i + \rho' PX_i + \beta' X_i + \eta_i$$

Identification and data issues

- Parental characteristics and education may be bad controls (if affected by area)
 - If BP affects education, controlling for education will underestimate the effect of BP
- Spatial sorting of parents based on unobservables (too much weight on BP?)

Descriptive statistics

Variable	Full time	Miss Ed.	Miss Occ.	Miss Parent	Small place
Women (%)	46.0	46.2	46.2	45.9	46.0
Age	34.9	34.7	34.7	37.4	37.5
Gross pay	1,487	1,493	1,493	1,589	1,588
Occupation (%)					
Managers / Senior Officials	14.1	14.2	14.2	15.4	15.3
Professional Occupations	9.7	10.0	10.0	11.0	10.9
Professional & Technical	11.6	11.7	11.7	12.4	12.4
Admin & Secretarial	17.8	18.0	18.0	17.6	17.7
Skilled Trades	11.7	11.6	11.6	11.1	11.1
Personal Service	11.3	11.2	11.2	10.3	10.3
Sales and Customer Service	6.6	6.6	6.6	5.7	5.7
Process/Plant/Machine Op.	10.5	10.2	10.2	10.5	10.5
Elementary	6.7	6.6	6.6	6.2	6.2
Location					
Resident city size	496,500	500,392	500,646	483,709	489,827
Birth city size	570,962	570,875	570,484	574,807	574,872
Number of observations	72,565	70,026	70,006	57,101	56,268
Number of individuals	12,699	12,266	12,257	9,153	9,019

Table 2: Lifetime mobility (% always worked where born)

	Total	No quals.	GCSE eq.	A-level eq.	Degree
Total	43.7	51.7	48.6	45.8	30.5
Born in					
Rural	33.1	40.5	37.6	32.7	21.5
Small city	44.7	49.2	52.2	53.8	27.1
Med city	48.1	59.5	53.2	50.4	31.3
Large city	47.1	55.2	51.8	49.3	35.5
% born same place as:					
Mother born	53.8	63.0	56.2	50.5	49.9
Father born	52.8	56.7	56.7	50.1	48.8

Table 3: Lifetime mobility (% always worked where born)

	Total	No quals.	GCSE eq.	A-level eq.	Degree
Total	43.7	51.7	48.6	45.8	30.5
At age:					
16	60.7	59.3	60.4	65.3	69.4
18	55.5	59.5	59.1	50.5	61.4
21	46.0	59.3	53.1	41.5	37.1
65	44.3	53.1	40.8	41.6	28.1

Descriptive statistics: headlines

- Restricted sample broadly representative
- Life time mobility
 - Around 40% only ever work while living in area where born
 - Increasing with qualifications
 - Higher for rural than urban
 - Increases with age up to ‘retirement’
- For at least 60%, birthplace identifies area where person grows up.

Regression results: Effect of current and birth place size

Table 4: One-step (log) gross total wage on birth place (full time workers)

	(1)	(2)	(3)	(4)
(log) Birth size	0.050***	0.044***	0.040***	0.034***
	(0.005)	(0.004)	(0.004)	(0.003)
Wave FE	X	X	X	X
Gender, Age⁽²⁾		X	X	X
Education			X	X
Occupation				X
Observations	68,080	68,080	68,080	68,080
R-squared	0.116	0.310	0.441	0.510

Table 5: 1st stage (log) gross total wage on city size (full time workers)

	(1)	(2)	(3)	(4)	(5)
(log) City size	0.064***	0.071***	0.050***	0.047***	0.029***
	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)
Wave FE	X	X	X	X	X
Gender Age⁽²⁾		X	X	X	X
Education			X	X	X
Occupation				X	X
Individual FE					X
Observations	76,294	76,294	76,294	76,294	76,294
R-squared	0.127	0.325	0.447	0.513	0.855
# Individuals					13,599

Table 6: 2nd-stage (log) gross wage; individual fe on birthplace (full time)

	(1)	(2)	(3)	(4)	(5)
(log) Birth size	0.056***	0.048***	0.049***	0.047***	0.034***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
1st-step					
Wave FE	X	X	X	X	X
Occupation				X	X
City size					X
2nd-step					
Gender Age⁽²⁾	X	X	X	X	X
Parent Class		X	X	X	X
Education			X	X	X
Occupation			X	X	X
Observation	7,107	7,107	7,107	7,107	7,107
R-squared	0.142	0.195	0.326	0.309	0.305

Birth place regressions: headlines

- Positive significant effect of birth place size
- Similar in magnitude to effect current city size
- Mechanisms
 - Parental social class → sorting of parents
 - Current city size → link to location decisions
 - *Not* education (once control for parental sorting)

Regression results: Education

Table 7: 2nd step regressions of years of education

	(1)	(2)	(3)
(log) Birth place size	0.027	0.075***	0.016
	(0.020)	(0.020)	(0.019)
Women		X	X
Year of birth		X	X
Parent Social Class			X
Observations	13,354	13,354	13,354
R-squared	0.000	0.070	0.172

Regression results: Inertia and Movers

Table 8: Current city size on birth place and controls

	(1)	(2)	(3)	(4)
Birth place size	0.414***	0.416***	0.415***	0.415***
	(0.014)	(0.014)	(0.014)	(0.014)
Wave FE	X	X	X	X
Gender, Age⁽²⁾		X	X	X
Parent social class			X	X
Education				X
Observations	72,801	72,801	72,801	72,801
R-squared	0.220	0.222	0.223	0.232

Table 9: Current city log size on birth place and controls (movers)

	(1)	(2)	(3)	(4)
Birth place size	0.042***	0.045***	0.037***	0.039***
	(0.010)	(0.010)	(0.011)	(0.010)
Wave FE	X	X	X	X
Gender, Age⁽²⁾		X	X	X
Parent social class			X	X
Education				X
Observations	42,379	42,379	42,379	42,379
R-squared	0.013	0.015	0.027	0.058

Table 10: 2nd step individual fe (gross total wage) on birthplace (full time, movers)

	(1)	(2)	(3)	(4)	(5)
(log) Birth size	0.041***	0.031***	0.035***	0.033***	0.030***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
1st-step					
Wave FE	X	X	X	X	X
Occupation				X	X
Current city size					X
2nd-step					
Gender, Av. Age	X	X	X	X	X
Parent Class		X	X	X	X
Education			X	X	X
Observations	4,061	4,061	4,061	4,061	4,061
R-squared	0.130	0.180	0.312	0.295	0.294

Mechanisms: headlines

- Positive significant effect of birth place size
- Mechanisms
 - Parental social class → sorting of parents
 - Current city size
 - Inertia
 - Some effect on location decisions
 - *Not* education (once control for parental sorting)

**Static vs dynamic
agglomeration economies
[preliminary]**

Incorporating dynamic agglom.

De la Roca and Puga - If learning matters:

$$w_{i(a)t} = \theta_i + X_{it}\beta + d_{it}\gamma + \sum_{j=C} \delta_{jc} e_{ijt} + \lambda_t + \varepsilon_{ict}$$

where d_{it} are city dummies; e_{ijt} is total experience for worker i in city j

In our setting \rightarrow 1st step:

$$w_{i(a)t} = \theta_i + X_{it}\beta + \lambda_1 RP_{i(a)t} + \lambda_2 \sum_{t=0}^{t-1} RP_{i(a)t} + \delta_t + \varepsilon_{i(a)t}$$

Table 11: (log) gross total wage, full time workers only, with learning, lifetime movers only once they have moved at least once

	(1)	(2)	(3)	(4)	(5)
(log) City size	-0.040***	0.023**	-0.013	-0.010	0.014***
	(0.010)	(0.009)	(0.008)	(0.007)	(0.005)
Log sum size until t-1	0.134***	0.077***	0.074***	0.063***	0.057***
	(0.007)	(0.007)	(0.006)	(0.006)	(0.004)
Wave FE	X	X	X	X	X
Gender, Age⁽²⁾		X	X	X	X
Education			X	X	X
Occupation				X	X
Individual FE					X
Observations	33,571	33,571	33,571	33,571	33,571
R-squared	0.158	0.284	0.415	0.500	0.860
# individuals					5,592

Table 12: 2nd step regressions of individual fe; ft workers; lifetime movers

	(1)	(2)	(3)	(4)	(5)	(6)
(log) Birth size	0.041***	0.031***	0.035***	0.033***	0.030***	0.011**
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
1st-step controls						
Wave FE	X	X	X	X	X	X
Occupation				X	X	X
Current city size					X	X
Cumul. city size						X
2nd-step controls						
Gender, Av. age	X	X	X	X	X	X
Parent Class		X	X	X	X	X
Education			X	X	X	X
Observations	4,061	4,061	4,061	4,061	4,061	3,485
R-squared	0.130	0.180	0.312	0.295	0.294	0.294

Learning: headlines

- Evidence of static and dynamic economies
 - Allowing for learning significantly reduces coefficient on birth place size
- Additional mechanism via effect on future (labour market) learning

Conclusions

Conclusions

- Sorting at birth matters
- Need to do more work to
 - Understand results!
 - Understand mechanisms
 - Assess implications for spatial disparities [e.g. contribution to variance]