

Agglomeration, accessibility and productivity: Evidence for large metropolitan areas in the US

Urban Studies

1–17

© Urban Studies Journal Limited 2016

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0042098015624850

usj.sagepub.com



Patricia C Melo

The James Hutton Institute, Aberdeen, Scotland, UK

Daniel J Graham

Imperial College London, UK

David Levinson

University of Minnesota, Minneapolis, USA

Sarah Aarabi

MBA Candidate at Harvard Business School, Boston, MA, USA

Abstract

This paper estimates the productivity gains from agglomeration economies for a sample of the largest metropolitan areas in the United States using measures of urban agglomeration based on employment density and employment accessibility. The latter is a more accurate measure of economic proximity and allows testing for the spatial decay of agglomeration effects with increasing travel time. We find that the productivity gains from urban agglomeration are consistent between measures, with elasticity values between 0.07 and 0.10. The large majority of the productivity gains occur within the first 20 minutes, and do not appear to exhibit significant nonlinearities.

Keywords

metropolitan areas, productivity, spatial decay, transport accessibility, urban agglomeration economies

Received June 2015; accepted December 2015

Introduction

The existence of urban agglomeration externalities implies that the allocation of resources to cities delivers greater productivity gains than non-urban areas. This has policy implications, particularly the rationale

Corresponding author:

Patricia C Melo, Social, Economic and Geographical Sciences, The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, Scotland, UK.

Email: patricia.melo@hutton.ac.uk

for investing in major infrastructure. The design of agglomeration-based policies requires knowledge about the magnitude and the spatial decay of the productivity benefits from urban agglomeration across different regions. In spite of abundant research on the size of productivity gains from spatial agglomeration, there has been insufficient research on the spatial decay pattern of agglomeration effects and the presence of nonlinearities. This study helps towards finding an answer to some of the key remaining questions in the literature, namely: How far and wide do the productivity effects of spatial agglomeration spread and how quickly do they attenuate over space? Are there significant discontinuities in these effects within the urban hierarchy?

Previous studies have investigated how agglomeration economies attenuate over space by using accessibility type measures such as *market potential*, *economic mass* and *effective density*. These measures can be described as a distance weighted sum of opportunities (e.g. employment, population) between pairs of locations and have been used to incorporate the notion of distance decay into the measurement of agglomeration economies. They improve on the more conventional measures based on population and employment densities by providing a better representation of spatial proximity. However, the majority of these studies measure accessibility in terms of physical distance and hence cannot account for the role of (changes in) transport networks on improved connectivity and the subsequent positive effect on productivity. To our knowledge, Graham (2007), Holl (2012), Lall et al. (2004), Le Néchet et al. (2012) and Rice et al. (2006) are the only studies using accessibility type measures based on travel times derived from actual road networks. These studies, however, generally assume a constant rate of decline in agglomeration effects with increased distance/travel time.

Only a few studies allow for varying rates of decay with increased distance and overall they suggest a steep decay of agglomeration effects, although they can extend as far as the boundaries of labour markets (e.g. Di Addario and Patacchini, 2008; Rice et al., 2006; Rosenthal and Strange, 2008).

Another limitation of the literature is the generalised adoption of a linear relationship between urban agglomeration and productivity, and hence the assumption that the productivity effects of agglomeration increase in a proportional way across the urban hierarchy. Some studies have allowed for variable returns to urban agglomeration by using a quadratic function (e.g. Carlino and Voith, 1992; Graham, 2007; Kawashima, 1975; Moomaw, 1983), but only very few have relaxed the assumption of a linear parametric fit. Exceptions include Graham and Dender (2011) and Le Néchet et al. (2012) who used semiparametric techniques on firm level data for the UK and France, respectively, and found evidence suggestive of nonlinearities in the relationship between productivity and urban agglomeration.

This paper addresses the limitations above by examining the productivity gains from urban agglomeration for the 50 largest metropolitan statistical areas (MSA) in the United States using a transport-based measure of agglomeration (i.e. employment accessibility by automobile) to investigate the spatial decay of the productivity-agglomeration effects and semiparametric techniques to test for nonlinearities in agglomeration effects across the urban hierarchy.¹ To our knowledge this is the first attempt to estimate productivity-agglomeration effects for the United States using a transport-based measure of agglomeration and semiparametric techniques.

The findings for the effect of urban agglomeration on labour productivity are very similar regardless of whether we use

employment density or employment accessibility to measure agglomeration. This suggests that, in the context of our study, metropolitan density seems to have a stronger role than road network speed in the realisation of urbanisation externalities. The preferred estimates indicate an elasticity value between 0.07 and 0.10, suggesting that a 10% increase in urban agglomeration is associated with a 0.7%–1% increase in labour productivity. The results further indicate that the productivity effects of urban agglomeration can extend up to 60 minutes driving time, although the large majority occur within the first 20 minutes and hence are spatially very localised. The semiparametric analysis does not reveal any significant ‘threshold effects’ within our sample of large and very large metropolitan areas, therefore indicating that a linear parametric fit seems to be a reasonable assumption.

The remainder of the paper is organised as follows. We next present the empirical methodology and the main estimation issues, while the following section describes the data and variables used in the analysis. The main results are then presented and discussed, and the final section draws together the main conclusions.

Empirical methodology

Under the standard assumption that input factors are paid the value of their marginal products, local input factors will have higher prices in more productive areas. In this context, wage rates should reflect, even if partially, labour productivity. Wage models can be used to investigate spatial variation in labour productivity. The literature offers several explanations for the differences in labour productivity across space. Under spatial equilibrium real wages should be equal across space, however this is rarely observed in the real world. Spatial inequalities in labour productivity and wages can result from exogenous and endogenous factors.

These include differences in human capital (e.g. Glaeser and Mare, 2001; Moretti, 2004b; Rauch, 1993); differences in the cost of living and in the availability and quality of local amenities (e.g. Glaeser and Gottlieb, 2009; Roback, 1982); and agglomeration externalities (e.g. Fujita and Thisse, 2002). To measure the effect of urban agglomeration on labour productivity we therefore need to specify wage models that control for other determinants of labour productivity. This is illustrated in the general wage model:

$$w_{it} = f(U_{it}, E_{it}, K_{it}, H_{it}, \lambda_t, \mu_i) \quad (1)$$

where the subscripts i and t denote the MSA and year, respectively. w_{it} is the average real wage, U_{it} represents urban agglomeration, E_{it} measures educational attainment, K_{it} denotes the Krugman index of relative industrial specialisation and H_{it} is a proxy for local cost of living. To account for unobserved shocks that are common to all areas but vary across time we include year specific effects λ_t . To account for additional MSA specific unobserved heterogeneity (e.g. valuable natural resources) we also include metropolitan area effects μ_i .

The log-linearised wage models are estimated using the two separate measures of urban agglomeration (U_{it}): employment density ($U_{it} = D_{it}$) and employment accessibility ($U_{it} = A_{it}$). The term ε_{it} in equation (2) is the residual error, which is assumed to be normally distributed while allowing for heteroskedasticity and spatial clustering on MSA.

$$\begin{aligned} \ln(w_{it}) = & \alpha + \beta \ln(U_{it}) + \sigma \ln(E_{it}) \\ & + \delta K_{it} + \gamma H_{it} + \lambda_t + \mu_i + \varepsilon_{it} \end{aligned} \quad (2)$$

To investigate the spatial decay of urban agglomeration effects we separate the measure of employment accessibility into a series of contiguous travel time bands k , as shown in equation (3).

$$\ln(w_{it}) = \alpha + \sum_k \beta_k \ln(A_{it}^k) + \sigma \ln(E_{it}) + \delta K_{it} + \gamma H_{it} + \lambda_t + \mu_i + \varepsilon_{it} \quad (3)$$

Finally, to examine the presence of nonlinearities in the productivity effects of urban agglomeration across MSA, we use a semi-parametric partially linear model (e.g. Ruppert et al., 2003; Wood, 2006), where urban agglomeration enters the equation nonlinearly according to a smooth function f estimated using penalised spline regression techniques and the other terms are defined as in the equations above. This is illustrated in equation (4).

$$\ln(w_{it}) = \alpha + f(U_{it}) + \sigma \ln(E_{it}) + \delta K_{it} + \gamma H_{it} + \lambda_t + \mu_i + \varepsilon_{it} \quad (4)$$

There are important identification issues that need to be considered in the estimation of the models above, namely possible endogeneity of urban agglomeration due to omitted variable bias or unobserved heterogeneity, and reverse causality between urban agglomeration, human capital and labour productivity. Individual unobserved heterogeneity is generally addressed through the use of the random-effects (RE) and fixed-effects (FE) estimators. The RE estimator assumes that unobserved heterogeneity is not correlated with the observed covariates, which if false leads to inconsistent and biased parameter estimates. On the other hand, although the FE estimator allows for correlation between unobserved heterogeneity and the model covariates, it can result in a great loss of efficiency in the presence of highly persistent data, as is typically the case of spatial agglomeration. A more flexible approach is the correlated random-effects (CRE) estimator, which includes the mean of the covariates and allows testing the assumption of no correlation between unobserved heterogeneity and model covariates (Chamberlain, 1982;

Mundlak, 1978). In the case of reverse causality, the concern is that more productive metropolitan areas may attract workers with higher education, thus further increasing urban agglomeration and educational attainment. The remedial strategy adopted in the literature is to use instrumental variable (IV) estimators. The most common instruments used in the agglomeration literature include long-lagged values of urban agglomeration (Ciccone and Hall, 1996; Combes et al., 2010; Rice et al., 2006) and geological and geographical instruments (Ciccone, 2002; Combes et al., 2010; Glaeser and Gottlieb, 2009; Rosenthal and Strange, 2008). Given the changes in MSA boundaries over time, it is difficult to construct an instrument for urban agglomeration based on deep time lags. Alternatively, we follow an approach similar to that of Fingleton (2003, 2006) and use a five-group method which ranks the endogenous variable into one of five quintiles according to its size and then defines the instrument as the rank order (i.e. quintiles). The rationale is that there is strong association between the rankings of urban agglomeration and urban agglomeration size, but there is no relation between the relative rankings and labour productivity other than through the value of urban agglomeration. To instrument human capital, we follow Moretti (2004a) who used a deep time lag for the presence of colleges and universities as created in the 19th century by The Land Grant Movement.

Data

The dataset consists of a balanced panel of the 50 largest MSAs in the United States with a population of at least 1 million. The time horizon of the study is limited by data availability for employment accessibility, which is available for 1990, 1995, 2001 and 2009. The data sources and variables used in the study are described in the following paragraphs.

Labour productivity

To represent labour productivity, we use data for average MSA wage per job, available from the Regional Economic Information System (REIS) of the Bureau of Economic Analysis (BEA). To calculate real average wage per job (w) we use the GDP deflator with base year for 1990.

Urban agglomeration: Employment density vs. employment accessibility

We consider two measures of urban agglomeration: employment density and employment accessibility. The former has been extensively used in the empirical literature and although it is generally preferred to simple measures of total population and employment, its main limitation is that it assumes a uniform distribution of people across space and it does not account for the role of the transport network on actual spatial proximity, which our measure of employment accessibility does. Employment accessibility is defined in equation (5) as the number of jobs a representative traveller can reach within a certain travel time threshold.

$$A_{k,t} = \pi \left[\frac{V_{n,t}k}{C_t} \right] D_t \quad (5)$$

where,

- D_t = metropolitan area employment density at year t
- k = time band (in minutes)
- $V_{n,t}$ = average network speed in km/h at year t
- C_t = average circuitry at year t

Because we do not have microdata on speeds by link going back to 1990 (this is only widely available, for a fee, in recent years with the advent of GPS), average metropolitan speeds estimated from a variety of data by the Texas A&M Transportation

Institute (TTI) in their Urban Mobility Report are used, comprising a weighted average of arterial and freeway speeds. Similarly, micro level employment (at the block level) is not available consistently before the Census LEHD datasets became available in 2003, so urban employment density is used. The average circuitry is the ratio of the shortest path network distance to the Euclidean distance and has been computed in a separate study (Levinson and El-Geneidy, 2009). For more details about the calculation of employment accessibility see Levinson (2012).

The representative traveller experiences uniform metropolitan average employment density, metropolitan average network speed and circuitries that vary by trip length. All variables change with year. The maximum distance (or time) travelled is constrained so that once the representative traveller reaches a band where s/he passes all metropolitan jobs, the accessibility is capped (i.e. the region does not go on infinitely, only so long as all jobs at average density are available).

The distribution of employment across the travel time bands between 10 and 90 minutes indicates that 15% of employment is accessible within the first 10 minutes, and this value increases to 54% for 20 minutes, 83% for 30 minutes, 94% for 40 minutes, 98% for 50 minutes and 99% for 60 minutes. As a result, on average, the full majority of employment is accessible within one hour's travel time. Unfortunately, our measure of employment accessibility is solely based on the road transport network and hence does not account for the role of public transport, which is likely to be important in the largest urban areas.² The average mode share for work trips in the US for 2013 is 76.4% drive alone and 9.4% car-pool, compared with 73.2% and 13.4% in 1990. It is lower in the largest metro areas, but even there automobile dominates (e.g. 59% for metropolitan New York, 73.6% for San Francisco) (McKenzie, 2015).

Economic structure

To account for differences in the economic structure of metropolitan areas that might also affect their wage composition we use a measure of relative economic specialisation based on the Krugman Specialisation Index (K) (Krugman, 1991).

$$K_{it} = \sum_{o=1}^O |s_{oi} - \bar{s}_{oi}| \quad (6)$$

where s_{oi} is the local employment share in industry o and \bar{s}_{oi} is the average employment share in industry o for the group of metropolitan areas. Industry employment data are obtained from the BEA REIS database for 14 different industries. The index values range between $K = 0$ (no specialisation) and $K = 2*(I-1)/I$ (maximum specialisation). The higher the index the more the economic structure of the local economy deviates from that of the reference group (i.e. sample average) and the more it is considered to be specialised.

Human capital

Human capital (E) is measured by the percentage of population aged 25 years and over holding a bachelor's degree or higher. Data for educational attainment were obtained from the decennial Census (for 1990 and 2000) and from the American Community Survey (ACS) for 2009. Using 2000 data for 2001, we had data for 1990, 2001 and 2009. Unfortunately, data were not available for 1995 as the ACS was conducted for the first time in 1996 in four counties only. To avoid losing a quarter of the observations we interpolated the 1995 educational data using a compound annual growth rate between 1990 and 2000.

Cost of living and local amenities

Wages are generally higher in larger urban areas, even after correcting for differences in

housing costs. Spatial variation in real wages can reflect differences in the availability and quality of urban amenities. To account for differences in the cost of living we use the Federal Housing Finance Agency (FHFA) house price index (H) for each MSA. We adjust the original index to be centred on the mean value of the index for our sample of metropolitan areas and to have 1990 as the baseline year. Lower real wages might be more acceptable where climate is pleasurable, while areas with hostile climate might have to offer higher real wages to attract people (Glaeser et al., 2001). To account for this we use controls for the climate region a given MSA belongs to based on the nine climate regions identified by the National Climatic Data Center (Karl and Koss, 1984).

Table 1 provides basic descriptive statistics for the variables, while Figure 1 shows scatter plots for the relationship between metropolitan area average wage and the measures of urban agglomeration (i.e. employment density and employment accessibility) and human capital. There is a strong positive correlation between average wages and employment density (0.60) and educational attainment (0.74), and a relatively less strong correlation for employment accessibility, especially within 20 minutes' travel time (0.15). The latter might result from greater congestion effects in the core urban areas of metropolitan areas. The correlation between average wages and employment accessibility is very similar across the travel time bands for 30, 60 and 90 minutes – 0.45, 0.54 and 0.55, respectively. This is expected because the number of jobs accessible within 30 and 60 minutes' travel time corresponds to approximately 83% and 99% of the MSA employment accessible within 90 minutes, respectively.

Results and discussion

This section presents and discusses the main findings. The results obtained from the IV

Table I. Descriptive statistics of variables.

Variable	Min	Mean	Median	CV	Max	SD _B /SD _W
Real average wage (<i>w</i>)	20,268	27,918	26,995	0.20	53,182	1.15
Emp. accessibility, 10 min (<i>A</i> ₁₀)	48,012	126,331	117,833	0.38	275,360	1.19
Emp. accessibility, 20 min (<i>A</i> ₂₀)	192,046	478,833	456,453	0.33	1,101,442	1.19
Emp. accessibility, 30 min (<i>A</i> ₃₀)	371,128	834,306	792,804	0.43	2,333,749	1.83
Emp. accessibility, 40 min (<i>A</i> ₄₀)	371,128	1,069,420	884,454	0.71	3,377,059	3.08
Emp. accessibility, 50 min (<i>A</i> ₅₀)	371,128	1,205,735	884,454	1.02	5,276,655	4.27
Emp. accessibility, 60 min (<i>A</i> ₆₀)	371,128	1,293,838	884,454	1.33	7,598,383	5.17
Emp. accessibility, 90 min (<i>A</i> ₉₀)	371,128	1,342,600	884,454	1.55	8,550,473	7.33
Employment density (<i>D</i>)	16	138	114	0.91	560	6.99
Krugman specialisation index (<i>K</i>)	0.07	0.18	0.16	0.47	0.45	5.53
Educational attainment (<i>E</i>)	14	27	26	0.23	47	1.54
House price index (<i>H</i>)	0.80	1.45	1.28	0.39	3.19	0.40

Notes: CV: coefficient of variation (i.e. ratio of standard deviation to the mean). SD_B/SD_W: ratio of between-MSA standard deviation over within-MSA standard deviation. Wages are expressed in real USD; employment accessibility is expressed in number of people; employment density is expressed in terms of people per square kilometre; Krugman's index of relative specialisation ranges between 0 (no specialisation) and 2 (maximum specialisation); Educational attainment consists of the percentage of population with bachelor's or higher degree; the cost of living index is based on the MSA house price index with 1990 as the reference year.

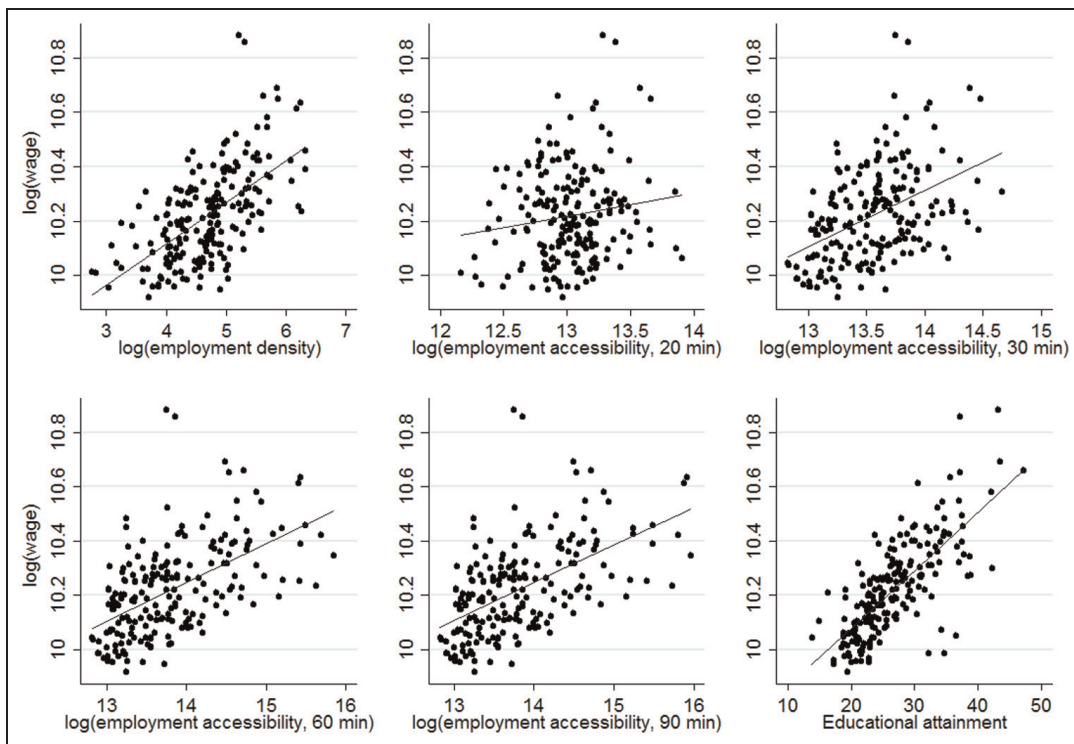


Figure 1. The relationship between real average wages, urban agglomeration and educational attainment.

and non-IV models using RE, FE and CRE estimators are presented in Tables 2 to 4, respectively. Table 2 shows the models in

which urban agglomeration is measured by employment density, while Table 3 reports the results for the models based on

Table 2. Results of the wage models using employment density.

Model	RE	FE	CRE (1)	CRE (2)	RE-IV	FE-IV	CRE-IV (1)	CRE-IV (2)
Constant	8.9950***	10.0361***	8.6019***	8.5865***	9.1973***	10.9582***	8.5751***	8.4855***
ln(D)	0.1139***	0.0302	0.0302	0.0336	0.1453***	0.0258	0.0579	0.0990***
ln(E)	0.1373***	-0.0541	-0.0541	-0.0441	0.0253	-0.3350*	-0.2908	-0.2476**
K	0.1509	-0.1828	-0.1828	0.0763	0.1803	-0.2831	-0.2615	-0.2372
H	0.0998***	0.1137***	0.1137***	0.1126***	0.0989***	0.1192***	0.1151***	0.1066***
mean of ln(D)			0.0760*	0.0716*			0.0483	
mean of ln(E)			0.3580***	0.3610***			0.5946***	0.5421***
mean of K			0.4150				0.4938*	0.4665*
mean of H			-0.1201*	-0.1189			-0.1214	
Observations	200	200	200	200	200	200	200	200
Hausman test (p-value)	19.17 (0.0039)				10.30 (0.2445)			
First stage Partial/Shea R ²								
- ln(E) ⁽¹⁾					0.85/0.80	0.34/0.12	0.28/0.06	0.27/0.26
- ln(D) ⁽¹⁾					0.84/0.79	0.27/0.10	0.15/0.03	0.80/0.79
Underidentification test ⁽²⁾					50.087***	11.631***	4.816**	9.183***
Weak identification test ⁽³⁾					284.97	5.968	2.476	7.203
Weak-instrument-robust inference ⁽⁴⁾					81.73***	14.20***	2.38	24.042***
R ² - overall	0.8218	0.4365	0.8481	0.8443	0.7884	0.1529	0.8444	0.8384
R ² - within	0.9355	0.9410	0.9410	0.9403	0.9356	0.9315	0.9331	0.9329
R ² - between	0.7299	0.0170	0.7696	0.7632	0.6639	0.3097	0.7696	0.7586
Year control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: ***, ** and * indicate significance at 1%, 5% and 10%, respectively. Standard errors are cluster-robust. Clustering is on metropolitan areas. Dependent variable: log of MSA real average wage rate. The models also include control variables for climate regions, but for the FE models these variables drop due to multicollinearity with the metropolitan area fixed-effects.

(1) Underidentification test (Kleibergen-Paap rk LM statistic).

(2) Weak identification test (Kleibergen-Paap rk Wald F statistic). Stock-Yogo weak identification test critical values are available from Stock-Yogo (2005).

(3) Weak-instrument-robust inference (Stock-Wright LM S statistic).

Table 3. Results of the wage models using employment accessibility.

Model	RE	FE	CRE (1)	CRE (2)	RE-IV	FE-IV	CRE-IV (1)	CRE-IV (2)
Constant	8.3152***	10.2467***	7.6229***	7.5497***	8.2277***	10.1331***	7.5939***	7.5263***
ln(A _{E,60})	0.0804***	-0.0013	-0.0013	-0.0007	0.1154***	0.0652	0.0688	0.0953***
ln(E)	0.1620***	-0.0720	-0.0720	-0.0710	0.0320	-0.3196***	-0.3148***	-0.2979**
K	0.2209	-0.1948	-0.1948	-0.1936	0.2948**	-0.2740	-0.2718	-0.2615
H	0.1040***	0.1173***	0.1173***	0.1151***	0.1039***	0.1147***	0.1143***	0.1091***
mean of ln(A _{E,60})			0.0954***	0.0958***			0.0253	
mean of ln(E)			0.4033***	0.3870***			0.6462***	0.6147***
mean of K			0.5403*	0.5408*			0.6173**	0.6091**
mean of H			-0.0771				-0.0740	
Observations	200	200	200	200	200	200	200	200
Hausman test	39.60 (0.0000)				17.28 (0.0273)			
First stage Partial/Shea R ²								
- ln(E) ⁽¹⁾					0.85/0.80	0.32/0.31	0.27/0.12	0.27/0.27
- ln(A _{E,60}) ⁽¹⁾					0.86/0.81	0.34/0.33	0.04/0.02	0.51/0.51
Underidentification test ⁽²⁾					49.185***	45.886***	2.953*	10.454***
Weak identification test ⁽³⁾					346.837	31.292	1.379	7.979
Weak-instrument-robust inference ⁽⁴⁾					80.99***	15.60***	9.47***	42.04***
R ² - overall	0.8137	0.3054	0.8509	0.8485	0.7732	0.2385	0.8449	0.8416
R ² - within	0.9302	0.9408	0.9408	0.9407	0.9284	0.9277	0.9278	0.9249
R ² - between	0.7353	0.2181	0.7750	0.7706	0.6454	0.0516	0.7750	0.7713
Year control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

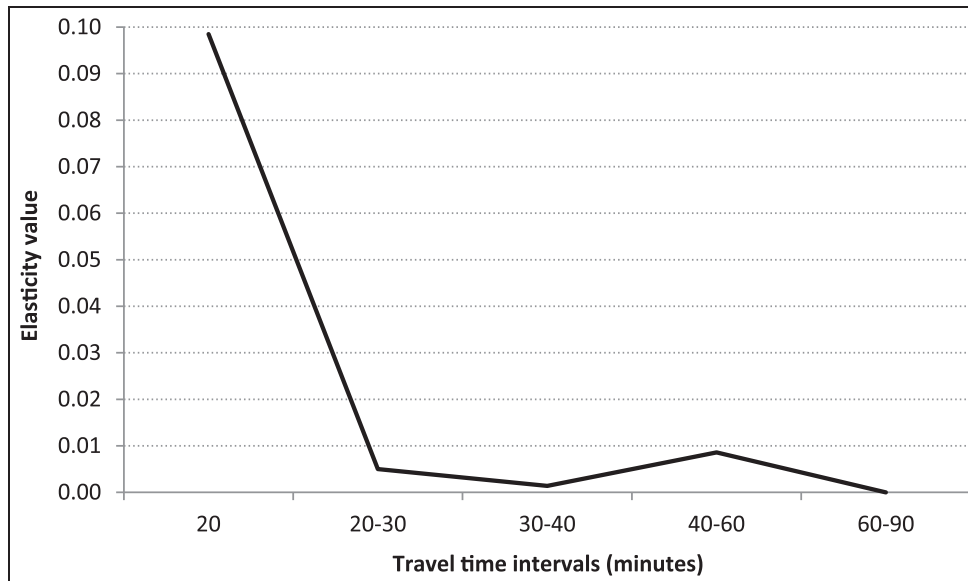
Notes: ***, ** and * indicate significance at 1%, 5% and 10%, respectively. Standard errors are cluster-robust. Clustering is on metropolitan areas. Dependent variable: log of MSA real average wage rate. The models also include control variables for climate regions, but for the FE models these variables drop due to multicollinearity with the metropolitan area fixed-effects.

- (1) Underidentification test (Kleibergen-Paap rk LM statistic).
- (2) Weak identification test (Kleibergen-Paap rk Wald F statistic). Stock-Yogo weak identification test critical values are available from Stock-Yogo (2005).
- (3) Weak-instrument-robust inference (Stock-Wright LM S statistic).

Table 4. Results of the wage models testing the spatial decay of urban agglomeration.

Model	RE	FE	CRE
Constant	8.6951***	10.3749***	7.5708***
$\ln(A_{E,20})$	0.0530***	-0.0106	-0.0106
$\ln(A_{E,20-30})$	0.0011	0.0003	0.0003
$\ln(A_{E,30-40})$	0.0021**	-0.0017	-0.0017
$\ln(A_{E,40-60})$	0.0039**	0.0001	0.0001
$\ln(A_{E,60-90})$	0.0023*	-0.0002	-0.0002
$\ln(E)$	0.1628***	-0.0727	-0.0727
K	0.2017	-0.2162	-0.2162
H	0.1040***	0.1214***	0.1214***
mean of $\ln(A_{E,20})$			0.1091**
mean of $\ln(A_{E,20-30})$			0.0047
mean of $\ln(A_{E,30-40})$			0.0031
mean of $\ln(A_{E,40-60})$			0.0085**
mean of $\ln(A_{E,60-90})$			-0.0007
mean of $\ln(E)$			0.4035***
mean of K			0.5316*
mean of H			-0.0881
Observations	200	200	200
Hausman test	54.15***		
R ² - overall	0.7933	0.2622	0.8570
R ² - within	0.9286	0.9417	0.9417
R ² - between	0.7175	0.3518	0.7855
Year control variables	Yes	Yes	Yes

Notes: ***, ** and * indicate significance at 1%, 5% and 10%, respectively. Standard errors are cluster-robust. Clustering is on metropolitan areas. Dependent variable: log of MSA real average wage rate. The models also include control variables for climate regions, but for the FE models these variables drop due to multicollinearity with the metropolitan area fixed-effects.

**Figure 2.** Spatial decay of urban agglomeration effects based on employment accessibility.

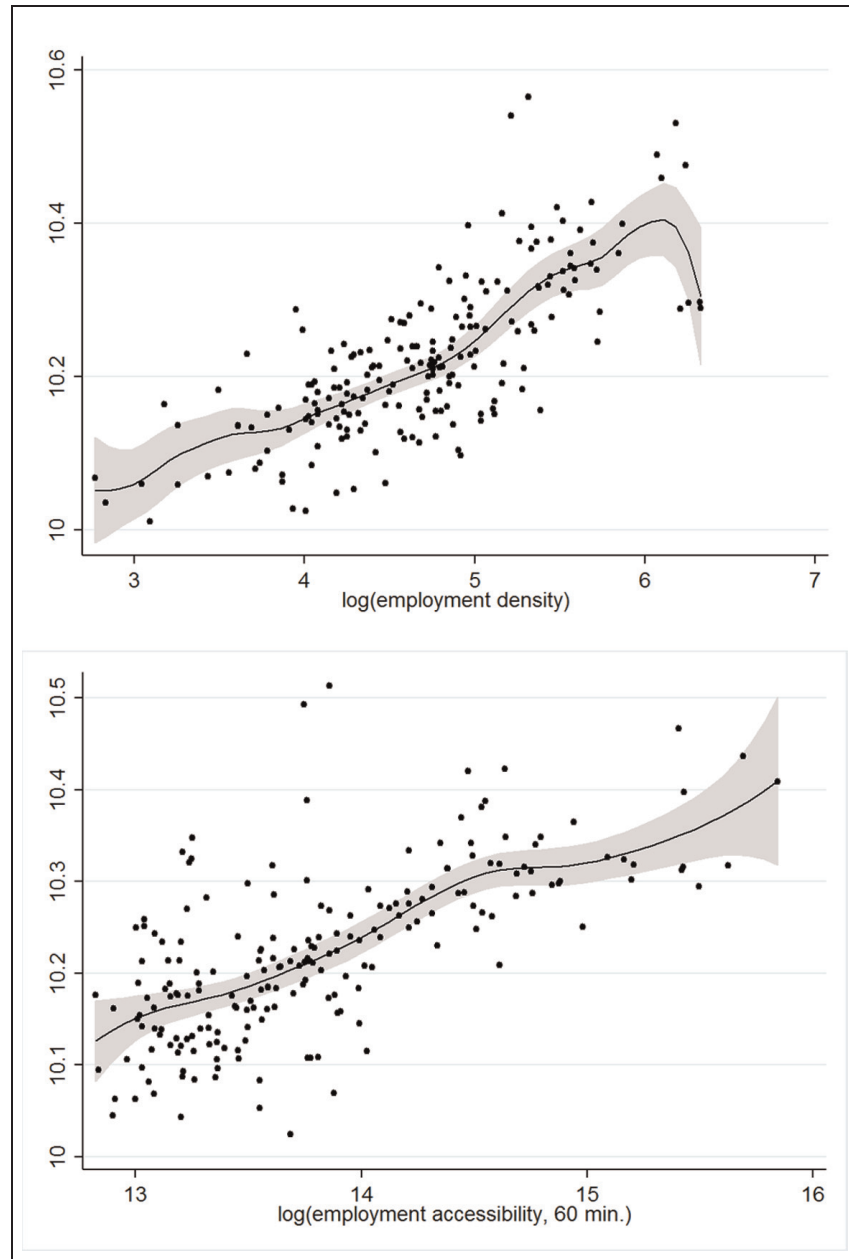


Figure 3. Nonparametric fit of the relation between metropolitan area wages (vertical axis) and urban agglomeration (horizontal axis): employment density (top) and employment accessibility (bottom).

employment accessibility. Table 4 and Figure 2 refer to the models testing the spatial decay of the productivity gains from urban agglomeration across a set of successive travel time bands. The last part of the section discusses the analysis of nonlinearities in the relation between productivity and urban agglomeration (Figure 3).

Productivity gains from urban agglomeration

We first discuss the results in Table 2. Overall the models have good explanatory power with generally high coefficients of determination (R^2). The first four columns of the table refer to the non-IV models,

while the last four columns refer to the IV models.³ The discussion focuses on the preferred models selected using appropriate tests for model comparison.⁴ The elasticity of wage with respect to employment density obtained from the preferred model CRE (2), which includes correlated random-effects for employment density, educational attainment and housing costs index, is 0.072; this indicates that increasing employment density by 10% increases wages, all else equal, by 0.72%. Correcting for reverse causality appears to have a minor impact on the effect of urban agglomeration, which increases from 0.072 to 0.099 (CRE-IV (2)).

The models in Table 3 are based on the 60 minutes' travel time band accounting for the large majority of metropolitan employment. The results obtained from the preferred model CRE (2), which includes correlated random-effects for employment density, educational attainment and relative industrial specialisation, are in line with those obtained in Table 2 for employment density.⁵ Raising employment accessibility by 10% is associated with an increase in wages of 0.96%. Similarly, correcting for reverse causality does not seem to change the elasticity estimate for employment accessibility, which is equal to 0.095 (CRE-IV (2)).

The findings reported above for the effect of urban agglomeration on average labour productivity are consistent between the two measures used and suggest that incorporating information on network speeds does not seem to affect the results much, which seem to be mainly driven by density effects. This may be partly due to the fact that employment accessibility is based on average metropolitan speeds, which do not vary as much as density. Our preferred estimates indicate an elasticity value between 0.07 and 0.10. Corroborating other studies (e.g. Ciccone, 2002; Ciccone and Hall, 1996), we find that correcting for reverse causality does not much affect the magnitude of the estimates.

Overall, the magnitudes of the urban agglomeration elasticity estimates appear to be slightly higher than those obtained in previous studies using aggregate regional level data for the United States; however, these differences are considered reasonable because our analysis is focused on the 50 largest metropolitan areas for which we expect to observe stronger returns to urban agglomeration. Ciccone and Hall (1996), using data for 1988, found that doubling county employment density increased productivity by about 6%. More recent evidence for the United States based on metropolitan area data suggests that the urban agglomeration elasticity ranges between 5 and 8%. Glaeser et al. (2001) and Glaeser and Resseger (2010) found that doubling metropolitan area population leads to an increase in average labour productivity of about 5% for 1980 and 8% in 1990 and 2001.

Table 4 presents the results for the spatial decay models using a series of consecutive travel time bands between 20 and 90 minutes. The CRE model is selected as the preferred model because it provides consistent and more efficient parameter estimates than the FE model.⁶ The results suggest that the spatial scope of the productivity effects of agglomeration can extend up to 60 minutes' driving time, although the bulk of the effects occur within the first 20 minutes. The magnitude of the effect reduces dramatically for the following travel time bands and tends to be statistically insignificant (with the exception of the travel time band between 40 and 60 minutes). This pattern of a steep spatial decay is illustrated in Figure 2, and supports the view of very localised benefits from spatial agglomeration while at the same time reaching as far as, but not necessarily in a continuous way, the boundaries of labour market areas.

In a report describing commuting patterns in the United States, McKenzie and Rapino (2011) show that the average one-

way travel time to work was about 22 minutes in 1990 and 25 minutes in 2000, and remained at 25 minutes in 2009. This commuting pattern is in line with stronger urban agglomeration effects within 20 and 30 minutes of travel time, as found in our analysis. In addition, similar findings of a steep spatial decay of urban agglomeration effects were also obtained from wage models based on driving time accessibility measures for Great Britain (Rice et al., 2006) and physical distance accessibility measures for the US (Rosenthal and Strange, 2008), Italy (Di Addario and Patacchini, 2008) and the UK (Melo and Graham, 2009).

Finally, we consider the results obtained from the semiparametric analysis. This approach does not make any assumptions on the functional form of the relationship between productivity and urban agglomeration, and allows the shape of the curve to be drawn as much as possible from the data. This provides a good way of investigating if there are ‘size’ thresholds above which the benefits from additional agglomeration become disproportionately greater or smaller.

The sample used consists of the 50 largest metropolitan areas, and hence contains considerably less variation than a full sample of large, medium and small metropolitan areas. As a result, the findings should be considered in the context of large and very large urban areas. Nonetheless, there is still more than an order of magnitude (about 20-fold) variation in the size of the metropolitan areas in the sample. Taking the year 2009 as reference, the latest year in our dataset, we observe that 60% of the sample consists of metropolitan areas with a population between 1–2.5 million, 22% of the sample consists of areas with a population between 2.5–5 million, 14% of the sample consists of areas with a population between 5–9 million and only 4% of the sample consists of metropolitan areas with a population of 10 million or more.

Figure 3 illustrates the nonparametric fit of the metropolitan area average real wage and urban agglomeration, where urban agglomeration is measured by employment density (top panel) and employment accessibility (bottom panel).⁷ The vertical axis shows the value of the nonparametric fit of the dependent variable and the horizontal axis shows the values of urban agglomeration. The shaded area corresponds to the interval determined by the two standard error lines above and below the estimate of the smooth curve. The shape of the curves does not reveal significant nonlinear effects, and indeed we cannot reject the test with null hypothesis that the nonparametric fit can be approximated by a parametric linear fit.⁸ This indicates that it is overall reasonable to assume a linear relation between spatial agglomeration and productivity for urban areas with and above 1 million people, and that there are no significant ‘threshold effects’ across urban areas of these sizes. It also suggests that there is no ‘wider economic benefits’ or ‘agglomeration economies’ rationale for allocating disproportionately more public investment to the top larger metropolitan areas compared to other large and very large metropolitan areas.

The main difference between the two curves is the downward sloping trend for the metropolitan areas with the highest employment density, although this effect appears to be statistically insignificant. One partial explanation for this difference might be the inability to disentangle congestion from high densities. It has been shown elsewhere (Levinson, 2012) that there is often a trade-off between high densities and lower mobility resulting from increased travel times and congestion. However, accessibility levels may vary between high density urban areas as a result of differences in the spatial organisation (e.g. urban form) and the planning and quality of transport systems. This

implies that transport-based measures of urban agglomeration such as our measure of employment accessibility may be better able to disentangle, even if only partially, the effect of increased congestion from that of increased density.

Human capital, economic structure and cost of living

As expected, educational attainment helps explain spatial differences in average wages across metropolitan areas. The effect is consistent across the models reported in Tables 2 to 4 and suggests that increasing educational attainment by 10% is associated with an increase in wages between 3–4%. As for economic structure, our analysis based on the Krugman index of relative specialisation could not identify any conclusive effect for its relationship with average metropolitan area wage. Future analysis on this topic could attempt to combine this indicator with a measure of the type of industrial specialisation as a way to further explore issues relating to industry mix. However, for the purpose of our analysis we were mainly interested in separating the effects of urban clustering from those of industrial clustering. The models also show that an increase of 1 point in the house price index is associated with a less than proportional increase in average wage.

Conclusions

This paper contributes to ongoing research on the productivity gains from urban agglomeration by empirically examining the spatial decay and nonlinearity of these effects using a new transport-based measure of urban agglomeration and semiparametric techniques. To the best of our knowledge this paper provides the first attempt to test for potential nonlinearities in the productivity gains from urban agglomeration for the

United States using semiparametric techniques.

The findings suggest that the productivity-agglomeration effects obtained for our sample of MSA are generally consistent between the two measures of urban agglomeration used – employment density and road transport-based employment accessibility. This suggests that, in the context of our study, metropolitan areas' densities seem to have a stronger role than road network speeds in the realisation of urban agglomeration externalities. The second main finding is that the large majority of the productivity-agglomeration effects occur within the first 20 minutes, which is in agreement with previous evidence on very localised benefits, while remaining significant but small at wider distances encompassing the boundaries of labour markets. Furthermore, the analysis suggests that the returns to additional spatial agglomeration for our sample of large and very large metropolitan areas appear to be constant; however, this result should not be generalised to the more heterogeneous population of metropolitan areas in the United States, for which there may be significant 'threshold effects'.

Some policy implications can be drawn from our findings. Evidence of a localised nature of spatial agglomeration effects suggests that within metropolitan areas immediate access (within 20 minutes) to jobs is very important for productivity, which in turn highlights the importance of investing in efficient transport networks. In addition, the lack of evidence in favour of nonlinearities in the relationship between productivity and urban agglomeration suggests there is no economic argument for allocating disproportionately more public money, to improve density and (road) transport accessibility, to the larger regions in our sample of metropolitan areas with 1 million people or more.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. MSA can be defined as functional economic areas with a strong degree of economic integration and are generally centred around a core urban area of at least 50,000 people.
2. Data for public transport revenue vehicle miles are available from the National Transit Database for urbanised areas, but not for metropolitan areas. In a recent study Chatman and Noland (2014) look at the relationship between transit services, physical agglomeration and productivity for urbanised areas in the US. Huang and Levinson (2015) have measured transit circuitry for recent years, and Owen and Levinson (2014) have measured accessibility.
3. The first stage partial R^2 and Shea R^2 have moderate to high values and tend to be higher for urban agglomeration. Moreover, the Kleibergen-Paap rank LM statistic test rejects the null hypothesis of model underidentification, indicating that our models are identified. In addition, the Kleibergen-Paap Wald rank F weak identification test statistic is generally higher than the Stock and Yogo critical values, suggesting that the instruments are not weak. Finally, the Stock-Wright LM S statistic also generally indicates that the instruments are valid.
4. The Hausman test rejects the null hypothesis of no correlation between MSA specific unobserved heterogeneity and the model covariates, suggesting the FE estimator should be preferred over the RE estimator as only the former provides consistent estimates. However, the FE estimator can result in great loss of efficiency for variables with relatively small time variation, as tends to be the case of education and agglomeration. The CRE estimator allows for greater flexibility; it provides consistent but more efficient parameter estimates than the FE estimator. CRE (1) model suggests there is correlation for employment density, educational attainment and housing cost index. Hence, model CRE (2) allows for correlation between MSA specific unobserved heterogeneity and these covariates. Turning to the IV models, model CRE-IV (1) indicates there is correlation for educational attainment and the relative industrial specialisation index. Model CRE-IV (2) allows for correlation between these variables and MSA specific unobserved heterogeneity.
5. For both the non-IV and IV models, the Hausman test suggests there is correlation between metropolitan area specific unobserved heterogeneity and the observed covariates. Model CRE (1) indicates this is the case for employment accessibility, educational attainment and relative industrial specialisation. Model CRE (2) allows for correlation for these variables and produces an elasticity estimate for employment accessibility equal to 0.096. Model CRE-IV (1) suggests there is correlation for educational attainment and relative industrial specialisation, which is allowed for in model CRE-IV (2) and produces an elasticity estimate for employment accessibility equal to 0.095.
6. The Hausman test rejects the null hypothesis of consistency of the RE model. Issues of little time variability in the measures of educational attainment and employment accessibility render most of the coefficients insignificant in the FE model. To avoid the loss of efficiency of the FE model, we estimate a CRE model which we select as our preferred model. IV models were not estimated because of the many time bands but given the results in Tables 2 and 3 we would not expect to find great differences between the IV and non-IV estimates.
7. The full set of results can be obtained from the authors upon request.
8. Similar results were obtained for the IV models, which can be obtained from the authors upon request.

References

- Carlino GA and Voith R (1992) Accounting for differences in aggregate state productivity. *Regional Science and Urban Economics* 22: 597–617.

- Chamberlain G (1982) Multivariate regression models for panel data. *Journal of Econometrics* 1: 5–46.
- Chatman DG and Noland RB (2014) Transit service, physical agglomeration and productivity in US metropolitan areas. *Urban Studies* 51: 917–937.
- Ciccone A (2002) Agglomeration effects in Europe. *European Economic Review* 46: 213–227.
- Ciccone A and Hall RE (1996) Productivity and the density of economic activity. *American Economic Review* 86: 54–70.
- Combes P-P, Duranton G, Gobillon L, et al. (2010) Estimating agglomeration economies with history, geology, and worker effects. In: Glaeser EL (ed.) *Agglomeration Economics*. Chicago, IL: University of Chicago Press, 15–65.
- Di Addario S and Patacchini E (2008) Wages and the city. Evidence from Italy. *Labour Economics* 15: 1040–1061.
- Fingleton B (2003) Increasing returns: Evidence from local wage rates in Great Britain. *Oxford Economic Papers* 55: 716–739.
- Fingleton B (2006) The new economic geography versus urban economics: An evaluation using local wage rates in Great Britain. *Oxford Economic Papers* 58: 501–530.
- Fujita MM and Thisse J-F (2002) *Economics of Agglomeration-Cities, Industrial Location and Regional Growth*. Cambridge: Cambridge University Press.
- Glaeser EL and Gottlieb JD (2009) The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature* 47: 983–1028.
- Glaeser EL, Kolko J and Saiz A (2001) Consumer city. *Journal of Economic Geography* 1: 27–50.
- Glaeser EL and Mare DC (2001) Cities and skills. *Journal of Labor Economics* 19: 316–342.
- Glaeser EL and Resseger MG (2010) The complementarity between cities and skills. *Journal of Regional Science* 50: 221–244.
- Graham DJ (2007) Variable returns to urbanization and the effect of road traffic congestion. *Journal of Urban Economics* 62: 103–120.
- Graham DJ and Dender KV (2011) Estimating the agglomeration benefits of transport investments: Some tests for stability. *Transportation* 38: 409–426.
- Holl A (2012) Market potential and firm-level productivity in Spain. *Journal of Economic Geography* 12: 1191–1215.
- Huang J and Levinson DM (2015) Circuitry in urban transit networks. *Journal of Transport Geography* 48: 145–153.
- Karl TR and Koss WJ (1984) Regional and national monthly, seasonal, and annual temperature weighted by area, 1895–1983. *Historical Climatology Series 4–3*. Asheville, NC: National Climatic Data Center.
- Kawashima T (1975) Urban agglomeration economies in manufacturing industries. *Papers in Regional Science* 34: 157–175.
- Krugman P (1991) *Geography and Trade*. London: MIT Press.
- Lall SV, Shalizi Z and Deichmann U (2004) Agglomeration economies and productivity in Indian industry. *Journal of Development Economics* 73: 643–673.
- Le Néchet F, Melo PC and Graham DJ (2012) The role of transport induced agglomeration effects on firm productivity in mega-city regions: Evidence for Bassin Parisien. *Transportation Research Record: Journal of the Transportation Research Board* 2307: 21–30.
- Levinson D (2012) Network structure and city size. *PLoS One* 7(1): e29721. DOI:10.1371/journal.pone.0029721.
- Levinson D and El-Geneidy A (2009) The minimum circuitry frontier and the journey to work. *Regional Science and Urban Economics* 39: 732–738.
- Mckenzie B (2015) Who drives to work? Commuting by automobile in the United States: 2013. *American Community Survey Reports*. Washington, DC: US Bureau of the Census.
- Mckenzie B and Rapino M (2011) Commuting in the United States: 2009. *American Community Survey Reports, ACS-15*. Washington, DC: US Census Bureau.
- Melo PC and Graham DJ (2009) Agglomeration economies and labour productivity: Evidence from longitudinal worker data for GB's travel-to-work areas. *SERC Discussion Paper 31*, pp. 1–48.
- Moomaw RL (1983) Is population scale a worthless surrogate for business agglomeration economies? *Regional Science and Urban Economics* 13: 525–545.

- Moretti E (2004a) Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics* 121: 175–212.
- Moretti E (2004b) Human capital externalities in cities. In: Henderson JV and Thisse JF (eds) *Handbook of Regional and Urban Economics*. Amsterdam: North-Holland, pp. 2243–2291 .
- Mundlak Y (1978) On the pooling of time series and cross section data. *Econometrica* 46: 69–85.
- Owen A and Levinson DM (2014) Access Across America: Transit 2014. Report no. CTS 14-11, 1-104. Accessibility Observatory, University of Minnesota.
- Rauch JE (1993) Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics* 34: 380–400.
- Rice P, Venables AJ and Patacchini E (2006) Spatial determinants of productivity: Analysis for the regions of Great Britain. *Regional Science and Urban Economics* 36: 727–752.
- Roback J (1982) Wages, rents, and the quality of life. *Journal of Political Economy* 90: 1257–1278.
- Rosenthal SS and Strange WC (2008) The attenuation of human capital spillovers. *Journal of Urban Economics* 64: 373–389.
- Ruppert D, Wand MP and Carroll RJ (2003) *Semiparametric Regression*. Cambridge: Cambridge University Press.
- Wood SN (2006) *Generalized Additive Models: An Introduction with R*. Boca Raton, FL: Chapman and Hall/CRC.