

Does Agglomeration Have a Positive and Statistically Significant Impact on Regional Labour Productivity in Great Britain?

Rachel Hunt
Professional Economist BSc and Apprenticeship Level 6
School of Economics
University of Kent, 2026

Executive Summary

Since 2008 Great Britain has faced slow productivity growth alongside large disparities in labour productivity between metropolitan and non-metropolitan regions. This paper tests the hypothesis that agglomeration can increase regional labour productivity in Great Britain through the positive externalities it creates. This paper adds to the small amount of literature using a lag of urbanisation to capture the delayed impact of past urbanisation on labour productivity, it utilises a panel dataset covering 160 regions at an International Territorial Level 3 from 2007 to 2021. A two-way fixed effect model which captures current total employment density and past employment density is used, the findings show increasing current total employment density leads to a decrease on gross value added per employee whilst increasing past total employment density leads to an increase.

AI Statement

I acknowledge the use of generative AI in code development for this paper. However, the work reported remains my own.

Acknowledgements

I would like to thank my family and friends for their support and encouragement during my time researching and writing this dissertation. Their support during this time was immense and did not go unnoticed. I would also like to thank my team at Arup for their support during the apprenticeship, during my time at Arup I have had the privilege to learn from so many intelligent and capable people, as well as make many memories I'll look back on fondly. Finally, special thanks to the University of Kent for their support during my studies and my dissertation supervisor Maria Kyriacou.

Introduction

Since 1890 economists theorised that increasing density of firms and employees can promote greater labour productivity through positive agglomeration externalities such as knowledge spillovers, asset sharing and larger labour pools (Marshall, 1890). Whilst many economists acknowledge the role highly agglomerated cities and urban centres play in promoting greater productivity, a growing schism in the literature would begin to form. More recent literature differentiates between agglomeration focused on increasing specialisation, referred to as localisation, and agglomeration which increases diversification also called urbanisation. With competing ideas on which form of agglomeration has a positive and statistically significant impact on labour productivity. This paper seeks to test the hypothesis that agglomeration has a positive and statistically significant impact on regional labour productivity, whilst employing controls for both urbanisation and localisation to reflect the more comprehensive and nuanced methodologies used within the modern literature. The research in this paper will provide policy recommendations and empirical evidence for how the built environment can help increase labour productivity in Great Britain.

The motivation for this research is to address the productivity puzzle in Great Britain from the perspective of a spatial economist. The productivity puzzle is characterised by the growing discrepancy between actual output per employee and anticipated output per employee. The slow growth of labour productivity has long-lasting consequences as increasing productivity is needed to improve living standards and wages within the United Kingdom (ONS, 2015). This research revisits the question of whether agglomeration can increase labour productivity using recent data and a dynamic model. If agglomeration externalities have the potential to increase labour productivity, then policy makers should seek to use regulation which incentivise agglomeration whether that is greater localisation and or diversification. However, if agglomeration externalities are no longer effective in increasing labour productivity due to the diseconomies effect of agglomeration, notably increasing pressure on public services, a lack of affordable housing and greater environmental degradation then policy makers should consider how they can mitigate these congestion externalities (NIESR, 2022).

This research adds to the small amount of literature using a dynamic two-way fixed effects model to address agglomeration's impact on labour productivity, it captures urbanisation and increasing localisation for the financial and insurance as well as manufacturing services. The

dataset in this paper covers 160 International Territorial Level 3 regions in Great Britain from 2007 to 2021, this paper has a larger sample size and covers a more recent period compared to other papers on this topic. The findings from this empirical model will be discussed in relation to how they align with the ongoing debates within the literature on this topic. A two-stage-least square model will be used to test for the presence of endogeneity. In addition, a sub-sample analysis will be conducted to determine the impact of agglomeration in both metropolitan and non-metropolitan regions. This paper will then provide policy recommendations, a discussion of the limitations of this research and how future research can improve upon this work.

Literature Review

Theoretical background behind agglomerations impact on labour productivity

The economist Alfred Marshall's in his 1890 book 'Principles of Economics' states that increasing the density of employees and firms in the same industry, also referred to as increasing localisation, can create positive externalities that promote greater economic development for regions and cities (Marshall, 1890). He states that localised economies increase the amount of knowledge spillovers between employees in the same industry thereby improving worker productivity. Additionally, localised economies can create larger labour pools which make recruiting workers with relevant industry knowledge and skillsets easier for firms, thereby improving efficient allocation of employees and reducing training costs for firms (Corradini et al, 2025). He suggests localised economies can promote asset sharing of specialised machinery for firms within the same industry which keep capital costs per employee low (Van der panne, 2015). Positive externalities from localised economies are referred to as Marshallian externalities.

Urban designer Jane Jacobs would put forward a different theory that agglomeration across different industries, also referred to urbanisation, can add new kinds of goods and services to cities which can lead to greater economic growth (Jacobs, 1969). Jacobs believed that the 'cross-fertilisation' of ideas that occurs in urbanised economies can promote innovation and experimentation within firms. This is referred to as a Jacobian externality, which increases productivity through the diversified economies created by urbanisation (Beaudry and Schiffauerova, 2009).

Figure 1: Different agglomeration externalities



Whilst both Marshall and Jacobs provide theoretical frameworks for how different types of agglomeration can increase labour productivity. Both hinge on the concept that the positive externalities of agglomeration are larger than the negative congestion externalities that agglomeration may generate. Congestion externalities include increasing land and housing prices, environmental degradation, increase competition which can crowd out weaker firms and even decrease innovation as firms leave highly agglomerated locations to avoid knowledge leakages to competitors (Zheng, 2001).

Empirical evidence for the relationship between agglomeration and labour productivity

Ciccone and Hall (1996) used a fixed effects model with log-log transformations and cross-sectional data from 1988, they concluded that doubling employment density per acre lead to a 6% increase in average output per employee, even when controlling for average years of schooling. Additionally, they found that employment density explains 55% of the variation in labour productivity per employee across the USA. Ciccone (2002) followed upon his prior research and looked at the impact of agglomeration on labour productivity in Europe. Using cross-sectional data for 1986,1987 and 1988 covering NUTS-3 regions in France, Germany, Italy, Spain, and the UK. He found that doubling employment density as measured by employment-by-acre increased average labour productivity by 5% for their log linear model with NUTS-2 region fixed effects employed. Both papers were highly influential for future research and outlined how urbanisation could increase output per employee, moreover it popularised the use of fixed effects when constructing empirical models due to its ability to

control for time invariant characteristics between regions. However, there are limitations in his research. Firstly, Ciccone stated that productivity gains maybe driven by the presence of different industry compositions across the USA and Europe. Later research would control for both urbanisation and localisation across industries to comprehensively measure the impact of both types of agglomeration on labour productivity.

Kluge and Lehmann (2012) looked at the way localisation and urbanisation in both manufacturing and financial services impacts GVA per employee across 70 German cities from 1998-2008. They used a location quotient to measure localisation and a two-way fixed effect model (TWFE). They found that increasing localisation of manufacturing led to increases in labour productivity by 0.20%, thereby providing evidence of Marshallian externalities. Moreover, increasing localisation for advance services defined as 'Financing, Insurance activities and Professional, Scientific, and Technical Activities' increased labour productivity by 0.012%. On the other hand, increasing urbanisation as measured by the Hirschman-Herfindahl index (HHI) lead to a negative impact on GVA per employee. These results contrast with Ciccone's work as it suggests increases in urbanisation and thus diversified economies can decrease labour productivity. Moreover, they excluded other industries such as agriculture, utilities, and mining as they are unlikely to produce knowledge spillover and firms in these industries lack autonomy over selecting where they are established. Whilst location quotients are used in some literature to measure localisation research from Beaudry and Schiffauerova (2009) highlighted its limitations. Notably it does not account for the absolute size of agglomeration instead it acts as a relative measure of localisation within a country. Some agglomeration externalities such as better labour pool match making requires large labour pools for the positive externality to increase labour productivity. As a result, many researchers still use employment density as it can better capture the absolute size of agglomeration.

Brühlhart and Mathys (2008) used NUTS-2 data across 20 European countries from 1980 to 2003. They estimated multiple models including a pooled OLS model, TWFE, Two-Stage Least Squares (2SLS) and a system generalised method of moments (GMM). Urbanisation was measured by employment density per Km^2 and lag of employment density per Km^2 . Their pooled OLS model showed a net positive impact of increasing urbanisation on gross value added per employee. However, their TWFE specification had a negative elasticity of 0.59% for current employment on labour productivity, whilst their lagged employment density had a positive elasticity of 0.44%. They concluded that congestion externalities may

materialise sooner than the Jacobean externalities from urbanisation suggesting future research should utilise dynamic models.

Their GMM model used employment density for finance and manufacturing along with other firm controls. They found that a 1% increase in employment density for manufacturing lead to a decrease of 0.2% for gross value added per employee in the long run, suggesting congestion externalities of agglomeration offset Marshallian externalities, the opposite finding to Kluge and Lehmann's (2012). On the other hand, a 1% increase in employment density for financial services lead to an increase of 0.26% showing that financial services benefit from localisation. The researchers stated that other industries such as construction, transportation, communication, and market services tend to be affected by government interference in addition their location is dictated by the presence of other industries. Thereby, the literature only focuses on the financial and manufacturing industries.

Ahrend et al (2017) used data across 430 cities across 5 OECD countries from 2005-2007, and a TWFE model. They found increasing population density per km² in high-tech manufacturing led to an increase of 1.1% whilst medium-low-tech manufacturing had no statistically significant results. Moreover, increases in localisation for financial services led to an increase of 0.87% in wages. The researchers found that increasing the HHI¹ led to a negative impact on wages suggesting diversification increase wages. However, the authors note that these results vary over different model specifications leaving them to conclude they may be unreliable. The researchers suggested that localisation increases wages in manufacturing industries which are knowledge intensive as well as financial services, even when controlling for the share of university graduates. A critique of this methodology made by Cingano and Schivardi (2004) is its use of wages as a proxy for labour productivity, as wages maybe subject to mean reversion as locations with low (high) wages may see faster (slower) wage growth over time. Instead, they suggest researchers should use more direct measures of labour productivity when such data is available.

Finally, some researchers have used sub-sample analysis to test the robustness of their findings and to see how agglomeration impact metropolitan and non-metropolitan regions differently. Anastassova (2006) used a fixed effect model and cross-sectional data and conducted an analysis of both metropolitan and non-metropolitan regions in Great Britain. They found that increasing employment density per Km² lead to an increase of 0.08 for

¹ A higher HHI suggests a higher level of localisation.

average wages in their metropolitan sample, and 0.04 in their non-metropolitan sample. As a result, they suggest metropolitan regions may experience increasing returns to agglomeration compared to non-metropolitan locations.

The presence of endogeneity in the literature

The meta-analysis conducted by Melo, et al (2009) highlighted how endogeneity remains a source of bias within the literature, as highly productive firms and or employees may relocate to areas with high levels of agglomeration thereby increasing their density. This bidirectionality leads to simultaneous causality between labour productivity and agglomeration, which violates the assumption of a zero conditional mean. Common methods to address this include using lags of agglomeration as instrumental variables, as past levels of agglomeration are correlated with current levels of agglomeration which meet the condition of instrument relevancy. But do not dictate current levels of productivity, thereby maintaining instrument exogeneity (Wooldridge, 2009, p. 508).

Kluge and Lehmann (2012) used 1-to-2-year lags of agglomeration as instrumental variables and estimated a 2 Stage Least Squares model (2SLS). Their instrumental variables had a high Cragg-Donald Wald F statistic leading them to conclude their instruments are not weak. They failed to reject the Hansen J test null hypothesis suggesting that their instruments are not correlated with the error term meaning their valid. Brülhart and Mathys (2008) used lags of employment density for their system GMM. Unlike, Kluge and Lehmann (2012) they chose to use deeper lags of 2 to 3-years for employment density. Their reasoning was that productivity shocks in the immediate past may persist into the current error term thereby violating the exclusion restriction. Using lags further back in time would ensure that the instruments are not correlated with the current error term.

Whilst much of the literature tests for endogeneity between agglomeration and labour productivity, papers using education as a proxy for human capital assume that it is exogenous. This paper will follow the literature and will only instrument the explanatory variables measuring agglomeration; however, this does reflect a possible limitation in the research if endogeneity is present in the control for human capital.

Summary of the Literature

The literature highlights two types of agglomeration which impact labour productivity: urbanisation and localisation. Urbanisation has mixed results with work from Ciccone finding it has a positive impact on labour productivity, whilst Brühlhart and Mathys (2008) found negative impacts of urbanisation on regional labour productivity. Moreover, many researchers suggested that localisation for manufacturing and financial services are statistically significant. With increasing localisation in the financial industry showing a positive and statistically significant impact on labour productivity across all papers reviewed, with mixed results for increasing localisation in the manufacturing industry. The model in this paper will include a measure of total employment density to capture urbanisation as well as the absolute size of agglomeration mirroring Ciccone and Hall (1996) and Ciccone (2002). Employment density for the financial and the manufacturing industries will be used to capture the impact of localisation on labour productivity similar to Brühlhart and Mathys (2008) paper. Furthermore, Brühlhart and Mathys (2008) found that positive externalities took time to materialise, thereby a static model may not be the best fit. This paper will test whether a lag of the agglomeration explanatory variables improve the fit of the model by using the Akaike and Bayesian information criterions. Researchers have used a control for education to help prevent omitted variable bias and isolate the impact of agglomeration on labour productivity. Therefore, a control variable for economically active people with a degree or equivalent will be included in this paper's model like the control used in Ahrend et al (2017). The most common model specification was a fixed effect model with time dummies as it can control for time invariant characteristics and shocks to labour productivity overtime. As a result, a TWFE model will be estimated along with the relevant diagnostic tests to determine if this model is the best fit. Moreover, a 2SLS will be estimated using past agglomeration as an instrumental variable to test for endogeneity. Finally, this paper will split the panel dataset between metropolitan and non-metropolitan regions to undertake a robustness testing similar to that used in Anastassova (2006).

Data sources

Dependent variable

The measure of labour productivity is regional Gross Value Added (GVA) it is at an International Territorial Level 3 (ITL3) which is the United Kingdom's replacement for NUTS 3. The dataset includes GVA across 18 SIC 2007 1-digit codes across 182 ITL 3 from 1998 to 2023, it is provided by the Office for National Statistics (ONS) as outlined in Table 1. The dataset includes regional deflators for each industry SIC code, ITL3 and year, as such the GVA data is reported in 2021 real prices.

Explanatory variables

Employment data from the Business Register and Employment Survey (BRES) was used to capture agglomeration. The BRES includes responses from approximately 87,000 businesses with the data being projected and re-weighted to represent the employment and industry composition of Great Britain. It is disaggregated across 21 SIC codes at a Local Authority District (LAD) level. For employment data between 2007 and 2009 the Annual Business Inquiry (ABI) was used as it is the precursor to the BRES. Like BRES it surveys business and extrapolates the data to a national level. It splits the employment data by SIC codes for both men, women, full time employees and part-time employees. For this research total employment includes employment for Females, Males, full-time and part-time employees (Office for National Statistics, 2013).

Control variable

A control for the impact of human capital on labour productivity is included that being the number of economically active people between ages 16-64 whose highest level of education is a degree or equivalent. This data is from the Labour Force Survey (LFS) which collects data on 175,000 respondents in the United Kingdom. The data is then reweighted using the most recent population projections estimates (Office for National Statistics, 2012).

Table 1: Data sources

Variable	Description	Source
Gross Value Added	Gross Value Added, ITL3, 1998-2023	Nominal and real regional gross value added (balanced) by industry (2025) - Office for National Statistics
Employment	Employment count by local authority districts	Business register and employment survey: open access (2015 to 2024)
		Business register and employment survey: open access (2009 to 2015)
	All employees (part-time, full time, Male, and Female) by local authority district prior to 2015	Annual Business Inquiry Employee Analysis (SIC2007): open access (2007 to 2008)
Area of Local Authority Districts	Area (M ²)	Open Geography, Local Authority Districts Shapefile
Number of degree holders or equivalent as their highest level of education	Highest level of education for economically active people aged 16-64 by local authority districts	Annual Population Survey (2007-2021)

Aggregating the data

To aggregate the LAD data to the ITL3 level a master lookup table was created using multiple lookup tables from open geography's website from 2018 to 2020. Some LAD names have changed between the publication dates of the lookup tables, the old and new names can be found by looking at the LADs councils' websites, all lookup tables and council websites are listed in the bibliography. The employment data and highest level of education data were both aggregated to the ITL3 level using the lookup table.

Organising and cleaning the data

The employment data includes 21 SIC codes at a LAD level, SIC codes A and B were combined to become 'AB' as well as D and E which became 'DE' as the GVA datasets reports these SIC codes together. SIC code 'U' and 'T' were excluded from the panel dataset used in this paper, as they represent activities of extraterritorial organisations and activities of households making them redundant for capturing agglomeration (HESA, n.d). The GVA dataset included 182 ITL3 regions, 11 ITL3 were removed as they were in Northern Island. Berkshire East and Berkshire West were combined to form Berkshire as BRES's employment

data is aggregated to the county of Berkshire. And 10 ITL3 were dropped due to there being incomplete employment data for 2021 leaving a total of 160 ITL3 regions shown in Figure 2.

Figure 2: Map of the 160 ITL3 regions



Manipulating the data

QGIS was used to extract the area of each LAD from the open geography shapefile, it was converted from M² to Km² and aggregated to the ITL3 level using the lookup table. The data was then normalised to show GVA per employee by dividing GVA by the total employment for the same ITL3 and year. The employment data for each SIC code was then divided by the ITL3 area in Km² to create an employment density for a given industry per Km² to capture localisation by different industries. The employment data for all 18 SIC codes were combined and divided by the ITL3 area to capture total employment density in Km² for each ITL3 across 2007-2021. Equation 1 shows the precise formulas where 'i' represents a ITL3 region, 't' represents a given time period, and 'z' represents an industry. Once the data had been

cleaned and transformed it was uploaded to Stata where the regressions and diagnostic tests were conducted.

Equation 1: GVA per employee (Left), Employment density by industry Z per Km² (Middle) Total Employment density by Km² (Right)

$$Y_{it} = \frac{GVA_{it}}{Employment_{it}} \quad X_{itz} = \frac{Employment_{itz}}{Area_i} \quad X_{it} = \frac{Employment_{it}}{Area_i}$$

Selecting industries to include in the model

To control for the impact of localisation on labour productivity employment density measures for specific industries will be included in the model. Including many industries in the same model will induce high levels of multicollinearity which can inflate standard errors and make the coefficients less precise (Wooldridge, 2009, p. 148). The literature has outlined that localisation within manufacturing and finance services have a statistically significant impact on labour productivity, therefore they are included in the model and tested for the presence of multicollinearity. Their variance inflation factors (VIF) are shown in Table 2 alongside their variable notations, the VIF are far below the severe level of multicollinearity that being 10 or above (Shalizi, 2025, p 256).

Table 2: Variance inflation factors for different industries

Explanatory variables	Explanatory variable notations	VIF
Employment density for manufacturing per Km ²	M	1.39
Employment density for financial and insurance per Km ²	F&I	1.65
Employment density per Km ²	EMP	3.10
Amount of economically active (Age 16-65) who have a degree or equivalent	DEGREE	1.73

Summary statistics

The correlation coefficient matrix in Table 3 shows a moderate relationship between ‘EMP’ and GVA per employee of 0.53, ‘F&I’ has a stronger relationship between ‘EMP’ of 0.60. This suggests a stronger correlation between localisation within the financial and insurance

industries and labour productivity then between urbanisation and labour productivity. ‘M’ has a weak correlation of 0.1 with GVA per employee. There is a moderate to weak correlation between ‘DEGREE’ and GVA per employee at 0.3 and a weak negative correlation between ‘DEGREE’ and ‘M’ of -0.06.

Table 3: Correlation coefficient Matrix

Correlation coefficient matrix	GVA per employee	EMP	F&I	M	DEGREE
GVA per employee	1				
EMP	0.53	1			
K	0.60	0.60	1		
M	0.10	0.47	0.22	1	
DEGREE	0.30	0.31	0.08	-0.06	1

Table 4 shows variables have a strong rightward skew as demonstrated by the skewness being above 1 (University of Southampton, 2024). This suggests outliers in the data are contributing to an uneven distribution between variables. Employment density for ‘F&I’ has the highest kurtosis and skewness, reflecting strong asymmetry in the data, this can be addressed by transforming the data with natural logarithms.

Table 4: Summary statistics

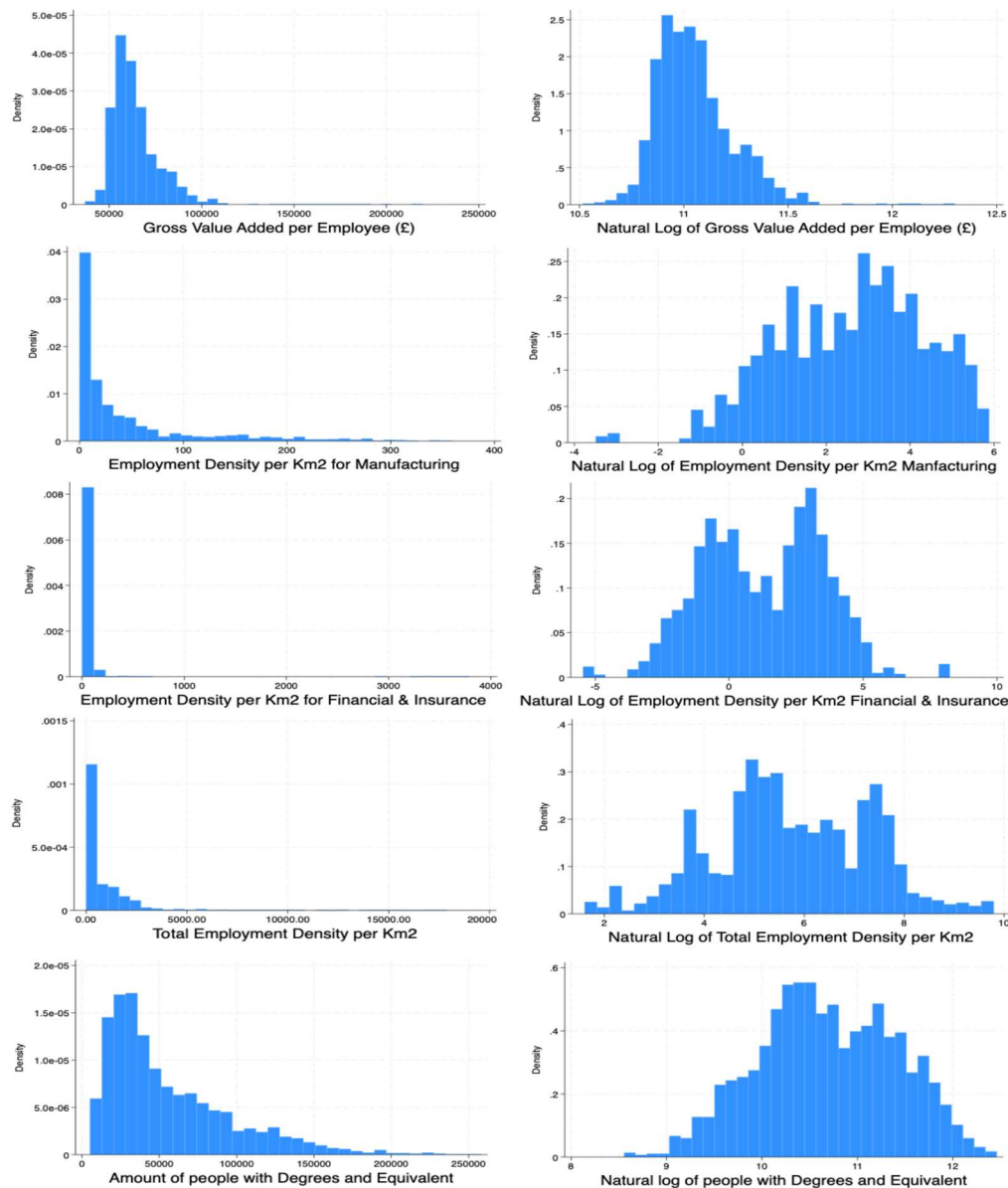
Variable	GVA per Employee	EMP per Km ²	F&I per Km ²	M per Km ²	DEGREE
25%	£55,227	116.8	0.57	3.5	26,500
50%	£61,200	268.0	3.7	16.5	42,200
75%	£68,976	1,184.0	21.8	50.8	75,900
Mean	£64,345	945.7	42.2	43.1	56,160
Standard Deviation	£14,590	1,917.0	276.6	63.9	40,930
Skewness	3.35	5.1	12.1	2.2	1.4
Kurtosis	26.2	35.1	151.5	7.60	5.0
Observations	2,400				

Logarithmic transformations

As outlined in Table 4 variables are highly rightward skewed, taking the natural logarithms makes the variables less sensitive to any outliers and reduces the skewness of the data (Wooldridge, 2009, p. 191). Figure 3 shows the data before and after natural logarithms are

employed. The interpretation of the model now represents the elasticity of the coefficients, where a 1% change in the explanatory variables leads to $\beta\%$ percentage change in GVA per employee.

Figure 3: Before logarithmic transformation (Left) after logarithmic transformation (Right)



Methodology

Theoretical framework for the preferred model specification

The preferred model specification is a two-way fixed effect model with a lag of employment density, the rationale for this model specification is as follows. Brülhart and Mathys (2008)

found that positive agglomeration externalities take time to materialise, as a result a lag of agglomeration is used to capture the delayed impact of past agglomeration on current labour productivity. Using an Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) it is determined a lag of total employment density improves the fit of the model. Fixed effects were preferred over random fixed effects as the Hausman tests rejected the null hypothesis. Moreover, year fixed effects were used to control for shocks in labour productivity over time. The preferred model specification:

$$\text{LNY}_{it} = \beta_1 \text{LN}(\text{EMP}_{it}) + \beta_2 \text{LN}(\text{EMP}_{it-1}) + \beta_3 \text{LN}(\text{F\&I}_{it}) + \beta_4 \text{LN}(\text{M}_{it}) + \beta_5 \text{LN}(\text{DEGREE}_{it}) + \mu_i + \lambda_t + \varepsilon_{it}$$

Table 5: Preferred model specification

Variable	Represents:
LNY_{it}	Natural log of GVA per employee in region 'i' at time 't'
$\beta_1 \text{LN}(\text{EMP}_{it})$	Natural log of total employment density per Km^2 in region 'i' at time 't'
$\beta_2 \text{LN}(\text{EMP}_{it-1})$	Natural log of total employment density per Km^2 in region 'i' at the previous period 't'
$\beta_3 \text{LN}(\text{F\&I}_{it})$	Natural log of financial and insurance density per Km^2 in region 'i' at time 't'
$\beta_4 \text{LN}(\text{M}_{it})$	Natural log of manufacturing employment density per Km^2 in region 'i' at time 't'
$\beta_5 \text{LN}(\text{DEGREE}_{it})$	Natural log of economically active people with a degree or equivalent in region 'i' at time 't'
μ_i	Region fixed effects by region 'i'
λ_t	Year fixed effects in time 't'
ε_{it}	The error term in region 'i' at time 't'

Robustness Testing

Stationarity and Time Fixed Effects

To ensure there is no spurious relationship being modelled the variables must be stationary meaning the variance and mean of the data is constant (Ballinari, 2021, p61). The alternative is that the data has a unit root and is non-stationary. A Levin-Lin-Chu (LLC) test is conducted the null hypothesis is that the data has a unit root. All variables have a p-value of

0.0000 bar EMP, which has a p-value of 0.067 meaning the null is not rejected (Levin et al, 2002).

When running an LLC with a time trend EMP has a p-value of 0.0000 this indicates that the data has a deterministic trend. Typically, linear time trends are used to detrend the data, but Mathys and Brühlhart's (2008) used year fixed effects as they capture non-linear shocks for each specific year as well as detrend the data. Therefore, year fixed effects are employed in the model where the year dummy for 2007 is dropped to prevent perfect collinearity. To give credence that the year fixed effects is effective in addressing the deterministic trend an LLC is conducted on the residuals of the preferred model specification and the p-value is 0.0000 meaning the residuals are stationary (Wooldridge, 2009, p. 363). Moreover, a joint F test is used and a p-value of 0.0000 is calculated, meaning the year dummies are jointly significant and improve the model fit (Wooldridge, 2009, p. 149).

Selecting a lag structure

As mentioned in the summary of the literature review the positive externalities of agglomeration may take time to materialise, as a result several different models with different lag structures are estimated. Table 6 shows the different models alongside their AIC and BIC values, a low AIC and BIC value suggests a better goodness of fit. To fairly compare AIC and BIC between the models' observations were restricted to the smallest sample size that being 2,080 observations (Gujarati and Porter, 2009, p 494). Model 3 had the lowest AIC and BIC values indicating a better fit compared to the other models. Therefore, a lag of total employment density will be used in the model specification to control for the impact of past urbanisation on GVA per employee.

Table 6: Selecting a Lag Structure

Different Models	Lag Structure	AIC	BIC
Model 1	No Lags	-2218.855	-2122.973
Model 2	LAG_EMP LAG_F&I LAG_M	-2222.903	-2110.100
Model 3	LAG_EMP	-2226.17	-2124.648
Model 4	LAG_EMP	-2218.176	-2088.453

	LAG_F&I LAG_M 2LAG_EMP 2LAG_F&I 2LAG_M		
Model 5	LAG_EMP 2LAG_EMP	-2224.619	-2117.456
Observations	2,080		

Including a lag of an explanatory variable does introduce high levels of multicollinearity, which as previously stated can inflate standard errors. The inclusion of LAG_EMP caused the VIF for EMP to increase, but multicollinearity remains below 10 for the other independent variables. Kluge and Lehmann (2012) put forward the argument that the omission of explanatory variables due to multicollinearity can in of itself create omitted variables bias, which can be a greater source of bias then multicollinearity. As a result, the high level of multicollinearity between EMP and LAG_EMP is tolerated as it improves the model specification fit, controls for omitted variable bias and aligns with pre-existing literature.

Estimating a fixed effect model

Fixed effects are frequently used in the literature as it controls for time invariant characteristics across locations such as inherent productivity advantages or differences in geographies. When running a fixed effect model (FEM) Stata produces an F-test to compare a pooled OLS to a FEM. The null is a pooled-OLS is preferred, the p-value is 0.0000 meaning the null hypothesis can be rejected and fixed effects are preferred (Wooldrige, 2009, p. 143.).

To determine whether a FEM or random effect model (REM) is the most appropriate a Hausman test is conducted. The null hypothesis of a Hausman test is that the error terms are not correlated with the independent variables meaning a REM is more appropriate. The alternative is that error terms are correlated with the independent variables and a FEM should be used. The Hausman test gives a p-value of 0.0000 meaning a FEM is preferred over a REM (Wooldrige, 2009, p. 493).

Heteroskedasticity and Autocorrelation

For the OLS estimator to remain the best linear unbiased estimator the Gauss-Markov theorem assumptions must be met. This includes the assumption of homoskedasticity and absence of autocorrelation. Using a modified Wald test for groupwise heteroskedasticity the null hypothesis that the variance of the residuals is constant across the panels is rejected meaning heteroskedasticity is present (Baum, 2006). The Wooldridge test for autocorrelation returned a p-value of 0.0000, meaning the null hypothesis that there is no first order autocorrelation is rejected (Drukker, 2003). To ensure the inference of the model's results are still reliable in the presence heteroskedasticity and autocorrelation, Kluge and Lehmann (2012) used both Driscoll-Kraay and clustered standard errors. Driscoll-Kraay standard errors are used when cross-sectional dependence is identified if it's not present clustered standard errors are preferred. A Pesaran test is conducted and a p-value of 0.404 was calculated meaning the null hypothesis that there is cross sectional independence cannot be rejected. Therefore, regional clustered standard errors are employed in the preferred model (Pesaran, 2004).

Testing for endogeneity

A reoccurring source of bias in the literature is the presence of simultaneous causality and omitted variables, the presence of both violates the Gauss-Markov assumption of exogenous explanatory variables and makes the results from the models inaccurate (Wooldridge, 2009, p 508). To test for endogeneity Klude and Lehamm (2012) used a 2SLS with 1-to-2-year lags of agglomeration, the same instrument selection is used in this paper. It is assumed all agglomeration explanatory variables are potentially endogenous.

All the literature reviewed with a control for education assumed it was exogenous, this is typically due to a lack of instrumental variables which could be used for both agglomeration and education. As this paper is focused on the potential presence of endogeneity in agglomeration, it is assumed DEGREE is exogenous in line with the literature. Table 7 shows the 2SLS model and diagnostic tests, lags of periods 1 and 2 were used for the endogenous regressors F&I, M, and EMP. These lags were chosen after experimenting with different lag lengths as they had the largest F-stats for the Cragg-Donald and Kleibergen-Paap rk Wald tests.

The Kleibergen-Paap rk LM statistic has a p-value of 0.0000 meaning the null hypothesis that instruments are not correlated with the endogenous regressors is rejected at a 1% level (Kleibergen and Paap, 2006). This suggests a strong level of confidence that the instruments are correlated with the endogenous regressors. The Hansen J test has a p-value of 0.80 meaning the null hypothesis that the instruments are uncorrelated with the error term and are valid cannot be rejected. The Durbin-Wu-Hausman test p-value is 0.65 meaning the null hypothesis that the assumed endogenous regressors is exogenous cannot be rejected (Guo et al, 2018).

The Cragg-Donald Wald F-statistic and the Kleibergen-Papp rk Wald F is 7.9 and 7.3 respectively. The rule of thumb for determining a strong instrument is an F statistic of 10 or greater (Keane and Neal, 2026). As the instruments are below 10 you can assume they are weak. A more accurate measure of instrument weakness comes from the Stock-Yogo values, as the instruments are less than the 5% IV relative biases of 9.53 but greater than the 10% IV relative bias of 6.61 we can more accurately conclude they are moderately weak (Stock and Yogo, 2002). Weak instruments can lead to unreliable tests results which in turn causes type II errors where the null hypothesis is incorrectly not rejected. Therefore, this is not a confident rejection of endogeneity being present in the model (Andrews et al, 2019).

table 7: Instrumental variables and 2SLS diagnostic tests

Endogenous regressors	Instrumental variables
EMP	LAG1_EMP LAG2_EMP
F&I	LAG1_F&I LAG2_F&I
M	LAG1_M LAG2_M
Diagnostic Tests: P-values	
Kleibergen-Paap rk LM Statistic	0.0000
Hansen J Statistic	0.80
Durbin-Wu-Hausman Statistic	0.65
Instrument Strength: F-statistics	
Cragg-Donald Wald F statistic	7.9
Kleibergen-Paap rk Wald F statistic	7.3
Stock-Yogo Critical Values	
5% maximal IV relative bias	9.53
10% maximal IV relative bias	6.61
Estimation Method	
Fixed Effects	Yes
Year Dummies	Yes
Clustered Standard Errors	Yes
Observations	2,080

Discussion of Results

Preferred Model Specification

Table 8: Preferred model results

Preferred Model Specification			
Variables	Coefficient	P-Value	Standard Errors
M	0.005	0.744	0.013
F&I	0.023***	0.008	0.008
EMP	-0.69***	0.000	0.052
LAG_EMP	0.22***	0.000	0.059
DEGREE	-0.016	0.325	0.016
Within R Square	0.48		
Between R Square	0.15		
Overall R square	0.13		
Observations	2,240		
Panel units	160		
Estimation Method			
Fixed effects	Yes		
Year fixed effects	Yes		
Clustered standard errors	Yes		
*, **, *** represents significance levels of 10%,5% and 1% respectively.			

Results for urbanisation

Table 8 shows the coefficients and p-values for the preferred specification, the within R-square is 48% meaning nearly half of the variation within regions is explained by the model similar to Ciccone and Hall (1996) who found a R-square of 55%. A 1% increase in EMP leads to a 0.69% decrease in GVA per employee with results at a 1% level of significance, meaning immediate increases in urbanisation decreases labour productivity. However, a 1% increase in LAG_EMP has a 0.22% increase in GVA per employee at a 1% level of significance, meaning past urbanisation has a positive and statistically significant impact on labour productivity suggesting Jacobean externalities take at least a year to materialise. The overall net impact of urbanisation on labour productivity is a decrease of 0.47% showing that congestion externalities are larger than Jacobean externalities in the long-term. These results

align less with Ciccone's work and more with Mathys and Brülhart (2008) whose TWFE model showed total employment density had a net negative impact on labour productivity of -0.57%.

Results for localisation

M has an elasticity of 0.005% suggesting a positive impact of localisation on manufacturing. However, this variable is not significant with a p-value of 0.744, meaning there is only weak evidence for localisation within manufacturing increasing labour productivity. Ahrend et al (2017) found insignificant results for low-technology manufacturing, it is possible that the manufacturing 1-digit SIC code in this dataset may disproportionately consist of low technology intensive manufacturing. Another explanation is that the inclusion of the fixed effects maybe be amplifying noise in M, as the fixed effects remove the between variation leaving only the smaller within variation (Wooldridge, 2009, pp. 481-483). This could explain why in the pooled OLS model in

Table 9, manufacturing had a negative impact on GVA per employee and was significant to a 1% level, as the greater variation was potentially leading to more precise results. On the other hand, a 1% increase in F&I leads to a 0.023% increase in labour productivity at a 1% level of significance. These results align with the literature that Marshallian externalities generated by increasing localisation in financial and insurance services are effective in increasing labour productivity.

Results for the amount of graduates

A 1% increase in DEGREE leads to a 0.016% decrease in Labour productivity, this finding is counterintuitive as the literature suggests a highly educated workforce is associated with greater labour productivity. However, the variable is not significant with a p-value of 0.325. Some literature such as Cingano & Schivardi (2004) as well as Ahrend, et al (2017) found insignificant results for their human capital control. The latter of which found that the share of graduates in the United States had an insignificant and negative elasticity of -0.02 despite being positive and significant for other OCED countries. Despite this the education control remained in their models as it can still reduce any omitted variable bias, therefore the control remains in the preferred model specification.

Table 9 shows the pooled OLS model where DEGREE is positive and significant to a 5% level, suggesting when fixed effects are employed and the between variation is removed, the remaining within region variation maybe too small. Similar to M this could result in a weak signal and an imprecise estimate in the preferred model specification.

Table 9: Pooled OLS and 2SLS model

Variables	Pooled OLS			2SLS		
	Coefficient	P-Value	Standard Errors	Coefficient	P-Value	Standard Errors
M	-0.066***	0.001	0.020	-0.0025	0.941	0.034
F&I	0.046***	0.009	0.017	0.034*	0.072	0.020
EMP	-0.096	0.499	0.140	-0.38	0.293	0.360
LAG_EMP	0.14	0.328	0.140	0.008	0.968	0.210
DEGREE	0.046**	0.047	0.023	-0.02	0.252	0.016
Observations	2,240			2,080		
Panel Units	160					
Estimation Method						
Fixed effects	No			Yes		
Year fixed effects	Yes			Yes		
Clustered standard errors	Yes			Yes		

The pooled OLS in

Table 9 shows urbanisation having a net positive elasticity of 0.04% on labour productivity however, EMP and LAG_EMP are insignificant with a p-value of 0.499 and 0.328 respectively.

In the 2SLS the only significant result is F&I to a 10% significant level with a 1% increase in F&I having an 0.034% increase in labour productivity, thereby providing further evidence of Marshallian externalities in finance and insurance activities promoting labour productivity. Moreover, the lack of significant results could be a function of the weak instruments used, or that an OLS estimator is more effective than a 2SLS.

Sub-sample analysis

A sub-sample analysis is taken to determine how agglomeration affects labour productivity in both metropolitan and non-metropolitan locations first used by Anastassova (2006). This paper uses the ‘City and Town Classification’ that outlines which LAD constitutes a core city and is thus a metropolitan region, 42 ITL3 are designated as metropolitan and 118 as non-metropolitan (Baker, 2018). Table 10 shows the results between the metropolitan and non-metropolitan samples.

Table 10: Metropolitan and Non-Metropolitan sample

Variables	Metropolitan			Non-Metropolitan		
	Coefficient	P-Value	Standard errors	Coefficient	P-Value	Standard errors
M	0.02	0.487	0.029	-0.0031	0.844	0.015
F&I	0.027**	0.029	0.012	0.021*	0.057	0.011
EMP	-0.69***	0.000	0.114	-0.70***	0.000	0.057
LAG_EMP	0.29**	0.018	0.120	0.18***	0.008	0.068
DEGREE	-0.041	0.281	0.037	-0.011	0.498	0.017
Within R-squared	0.55			0.46		
Between R-square	0.17			0.16		
Observations	588			1,652		
Panel units	42			118		
Estimation Method						
Fixed effects	Yes					
Year fixed effects	Yes					
Clustered standard errors	Yes					

Results for urbanisation

The results for the metropolitan sample show that increasing EMP and LAG_EMP by 1% leads to a net negative impact of 0.4%. In the non-metropolitan sample, the net negative impact is 0.52%. This does suggest that urbanisation externalities have a stronger diseconomies effect on non-metropolitan areas, Anastassova (2006) found that urbanisation in their metropolitan sample lead to larger increase in labour productivity. This could suggest that urbanisation is more effective in metropolitan regions potentially due to metropolitan

regions having a larger absolute size of agglomeration, which allows for size dependent agglomeration externalities like better labour pool match making to occur. However, the results for LAG_EMP for the metropolitan sample is significant to a 5% level, whilst LAG_EMP in the non-metropolitan sample is significant to a 1% level suggesting a lower predictive power in the metropolitan sample when compared to the non-metropolitan sample.

Results for localisation and amount of graduates

In the metropolitan sample increasing F&I by 1% led to an increase in labour productivity of 0.027% at a 5% significant level and in the non-metropolitan sample it is an increase of 0.021% at a 10% significant level. Thereby enforcing the notion that Marshallian externalities are present for localisation in the financial and insurance services for metropolitan regions with a marginally weaker correlation within non-metropolitan regions.

Increasing M by 1% led to a 0.02% increase in labour productivity for metropolitan regions and a decrease of -0.0031% in non-metropolitan regions. This does suggest a non-linear relationship for M, however, both results are not significant with p-values of 0.487 for the metropolitan sample and a p-value of 0.844 for the non-metropolitan sample. These results maybe a function of the imprecise estimation. Moreover, in the metropolitan sample the 1% increase in DEGREE led to a 0.041% decrease in labour productivity and a -0.011% in the non-metropolitan sample. This does suggest the diseconomies effect is stronger in the metropolitan sample, but as both results are insignificant it is likely a function of the lack of variation in the data rather than a reflection of a true diseconomies of DEGREE.

Policy recommendations

The research in this paper shows that increasing localisation for financial and insurance services can increase labour productivity in both metropolitan and non-metropolitan regions in Great Britain. Policy levers which expedite the planning process for new developments and make planning regulations more flexible could help incentivise greater localisation for financial and insurance services. Examples of such policies include establishing Urban Development Corporations (UDCs) which allow the central government to streamline redevelopment and build larger office floor spaces to accommodate for trading floors for

financial institutions, thereby incentivising greater localisation for financial and insurance services (Smith, 1991).

Moreover, policy levers which address the congestion externalities of urbanisation can be used. Such as greater investment into transportation which can decrease traffic and overcrowding on public transport, as well as greater protection of green spaces and policies aimed at reducing air pollution (Zheng, 2001). Furthermore, research from Newland et al (2025) found that increasing the housing supply can increase labour productivity particularly in highly agglomerated areas in London and the Southeast of England. They argue policies which keep house prices affordable can stop productive employees being deterred from working in highly urbanised areas because they are forced to pick between affordable housing and investments into their productivity. As a result, increasing the supply of affordable housing and affordable rental properties by increasing construction of new homes or including rental controls could help negate any diseconomies effect of agglomeration. Reducing these congestion externalities, can decrease the immediate diseconomies effect of urbanisation and even increase the Jacobean externalities of past urbanisation on labour productivity.

Limitations of the research and recommendations for future research

A limitation in this research is the ambiguity of whether endogeneity is present in the model. The instrumental variables used in the 2SLS were weak meaning the results from the Durbin-Wu-Hausman and Hansen J tests are potentially unreliable. Additionally, much of the literature assumes controls for education are exogenous, this is typically due to constraints on identifying instruments for both agglomeration and education. Future research should test for endogeneity in their human capital controls and seek to identify stronger instruments, or they should consider using a system-GMM model similar to Mathys and Brühlhart's (2008). Moreover, the Wooldridge test for autocorrelation and modified Wald test detected both autocorrelation and heteroscedasticity respectively. Whilst clustered standard errors were used to adjust the inference of the model, different model specifications like a generalised least squares model could be used to correct for both autocorrelation and heteroscedasticity. Additionally, future research could use a Herfindahl-Hirschman Index (HHI) to measure urbanisation as it can more accurately capture diversified economies by measuring market concentration between firms (Brezina et al, 2014). In addition, detailed firm level datasets

could be utilised as they provide a greater range of controls across firms and or employee characteristics. These greater range of controls could provide more accurate measures of high and low technology intensive manufacturing as well as employee education and skillset, they may also be better suited for capturing any omitted variables bias.

Conclusion

The research in this paper set out to answer whether agglomeration led to positive and significant increases in regional labour productivity for Great Britain. The research used a panel dataset from 2007 to 2021 across 160 regions in Great Britain. It contributes to the small amount of literature which uses a dynamic model to control for the delayed impact of agglomeration on labour productivity and revisits this topic using recent data for Great Britain.

The findings in this paper are that increased urbanisation has a net negative impact on labour productivity, with positive Jacobean externalities from urbanisation taking a year to materialise. This result aligns with literature which suggests that the diversified economies created by increased urbanisation have a diseconomies effect on labour productivity. On the other hand, Marshallian externalities from increased localisation in financial and insurance services have a statistically significant and positive impact on labour productivity even when controlling for the number of graduates in a region. Increasing localisation for manufacturing was shown to increase labour productivity, however, these results were insignificant making them potentially unreliable. The sub-sample analysis showed that urbanisation has a net negative impact on labour productivity across metropolitan and non-metropolitan regions. Whilst increasing localisation for financial and insurance services had a positive and significant impact on labour productivity for both metropolitan and non-metropolitan regions. Policymakers could promote further localisation of financial and insurance services by altering planning regulations to accommodate and incentivise the construction of greater density commercial space for financial and insurance firms. Additionally, policymakers could consider ways to mitigate congestion externalities particularly those that keep land and house prices low, improve transportation and or reducing the rate of environmental degradation. This research has demonstrated agglomeration can have both a positive and negative impact on labour productivity depending on the type of agglomeration occurring and the industries being controlled for. Further research could focus on identifying stronger instrumental

variables or could use a GMM model to identify and control for potential endogeneity. Different measures of urbanisation such as a HHI can be used to measure the impact of urbanisation more accurately. Alongside using detailed firm level data to control for a broader array of firm and employee characteristics.

Bibliography

- Anastassova, L. (2006). '*Productivity differences and agglomeration across districts of Great Britain*'. [online] Prague: Center for Economic Research and Graduate Education (CERGE). p.289. Available at: <https://www.cerge-ei.cz/pdf/wp/Wp289.pdf>
- Andrews, I., Stock, J.H. and Sun, L. (2019c). '*Weak Instruments in Instrumental Variables Regression: Theory and Practice*'. Annual Review of Economics, 11(1), pp.727–753. <https://doi.org/10.1146/annurev-economics-080218-025643>.
- Baker, C. (2018). City & Town Classification of Constituencies & Local Authorities. *commonslibrary.parliament.uk*. House of Commons Library. [online] Available at: <https://commonslibrary.parliament.uk/research-briefings/cbp-8322/>.
- Ballinari, D. (2021) '*INTRODUCTION TO TIME SERIES ECONOMETRICS Fall Semester 2021*', pp. 8–158. Available at: https://dballinari.github.io/assets/script_itse.pdf.
- Baum, C.F. (2006) 'Stata Tip 38: '*Testing for Groupwise Heteroskedasticity*', The Stata Journal: Promoting communications on statistics and Stata, 6(4), pp. 590–592. Available at: <https://doi.org/10.1177/1536867x0600600412>.
- Beaudry, C. and Schiffauerova, A. (2009c) '*Who's right, Marshall or Jacobs? The localization versus urbanization debate*', Research Policy, 38(2), pp. 318–337. Available at: <https://doi.org/10.1016/j.respol.2008.11.010>.
- Brezina, I., Pekár, J., Čičková, Z. and Reiff, M. (2014b). '*Herfindahl–Hirschman index level of concentration values modification and analysis of their change*'. Central European Journal of Operations Research, 24(1), pp.49–72. Available at: <https://link.springer.com/article/10.1007/s10100-014-0350-y>
- Brühlhart, M. and Mathys, N.A. (2008) '*Sectoral agglomeration economies in a panel of European regions*', Regional Science and Urban Economics, 38(4), pp. 348–362. Available at: <https://doi.org/10.1016/j.regsciurbeco.2008.03.003>.
- Ciccone, A. (2002). '*Agglomeration effects in Europe*'. European Economic Review, 46(2), pp.213–227. Available at: [https://doi.org/10.1016/S0014-2921\(00\)00099-4](https://doi.org/10.1016/S0014-2921(00)00099-4).
- Ciccone, A. and Hall, R.E. (1996). '*Productivity and the Density of Economic Activity*'. The American Economic Review, Vol. 86, No. 1. [online] 86(1), pp.54–70. Available at: <https://www.jstor.org/stable/2118255>.
- Cingano, F. and Schivardi, F. (2004) '*Identifying the Sources of Local Productivity Growth*', Journal of the European Economic Association, 2(4), pp. 720–742. Available at: <https://doi.org/10.2307/40004879>.
- Corradini, C., Morris, D. and Enrico Vanino (2025) '*Marshallian agglomeration, labour pooling and skills matching*', Cambridge Journal of Economics. Available at: <https://doi.org/10.1093/cje/beaf010>.

- Cumbria Intelligence Observatory (2020). *Cumbria Observatory – Unitary Local Authority*. [online] Cumbriaobservatory.org.uk. Available at: <https://www.cumbriaobservatory.org.uk/unitary-local-authority/>.
- Cyngor Sir Ceredigion County Council (2020). *Growing Mid Wales - Ceredigion County Council*. [online] Ceredigion.gov.uk. Available at: <https://www.ceredigion.gov.uk/your-council/partnerships/growing-mid-wales/>
- Data.gov.uk. (2024). *Local Authority Districts (December 2014) Boundaries GB BFE - data.gov.uk*. [online] Available at: <https://www.data.gov.uk/dataset/6e2333b8-d8cb-464f-b8b6-ff9f032f7100/local-authority-districts-december-2014-boundaries-gb-bfe>.
- Drukker, D.M. (2003) 'Testing for Serial Correlation in Linear Panel-data Models', 'The Stata Journal: Promoting communications on statistics and Stata', 3(2), pp. 168–177. Available at: <https://doi.org/10.1177/1536867x0300300206>.
- East and North Hertfordshire HCP (2024). *East and North Hertfordshire HCP*. [online] Herts and West Essex ICS. Available at: <https://www.hertsandwestessex.ics.nhs.uk/about/hcp/east-north-hertfordshire/>.
- East Ayrshire Council (2026). *About the Council · East Ayrshire Council*. [online] East-ayrshire.gov.uk. Available at: <https://www.east-ayrshire.gov.uk/CouncilAndGovernment/About-the-Council/About-the-Council.aspx>
- East Dunbartonshire Council (2020). *East Dunbartonshire Area Profile - East Dunbartonshire Council*. [online] Eastdunbarton.gov.uk. Available at: <https://www.eastdunbarton.gov.uk/about/our-council/about-the-council/east-dunbartonshire-area-profile/>.
- Gujarati, D.N. and Porter, D.C. (2009) 'Basic Econometrics'. The McGraw-Hill Series Economics. New York: McGraw-Hill Irwin. Fifth Edition. Available at: https://cbpbu.ac.in/userfiles/file/2020/STUDY_MAT/ECO/1.pdf
- Guo, Z., Kang, H.-S., T. Tony, C. and Small, D.S. (2018b). 'Testing endogeneity with high dimensional covariates'. *Journal of Econometrics*, [online] 207(1), pp.175–187. Available at: <https://doi.org/10.1016/j.jeconom.2018.07.002>.
- Higher Education Staff Statistics (n.d.). 'Standard Industrial Classification: SIC2007'. [online] HESA. Available at: <https://www.hesa.ac.uk/collection/coding-manual-tools/sicsocdata/sic-2007>.
- Jacobs, J. (1969) 'The economy of cities'. Published by Random House, Inc. eBook ISBN 9780525432869. V4.1.a.
- Keane, M. and Neal, T. (2026). 'A Practical Guide to Weak Instruments'. *The Annual Review of Economics*, Annu. Rev. Econ. 2024. Vol.16:185–212. Available at: <https://www.annualreviews.org/content/journals/10.1146/annurev-economics-092123-111021#:~:text=https%3A//doi.org/10.1146/annurev%2Deconomics%2D092123%2D111021>
- Kleibergen, Frank & Paap, Richard, (2006). 'Generalized reduced rank tests using the singular value decomposition', *Journal of Econometrics*, Elsevier, vol. 133(1), pages 97-126, July. Available at: <https://ideas.repec.org/a/eee/econom/v133y2006i1p97-126.html>
- Kluge, Jan and Lehmann, Robert (2012): 'Marshall or Jacobs? Answers to an unsuitable question from an interaction model', ifo Working Paper, No. 124, ifo Institute - Leibniz

- Institute for Economic Research at the University of Munich, Munich. Available at: <https://hdl.handle.net/10419/73827>
- LAD (2020) to LAU1 to ITL3 to ITL2 to ITL1 (January 2021) Lookup in the UK (V2) (2020) Statistics.gov.uk. Available at: <https://geoportal.statistics.gov.uk/documents/cdb629f13c8f4ebc86f30e8fe3cddda4/about>.
- Levin, A., Lin, C.-F. and James Chu, C.-S. (2002) 'Unit root tests in panel data: asymptotic and finite-sample properties', Journal of Econometrics, 108(1), pp. 1–24. Available at: [https://doi.org/10.1016/s0304-4076\(01\)00098-7](https://doi.org/10.1016/s0304-4076(01)00098-7).
- Local Authority District (2018) to NUTS3 to NUTS2 to NUTS1 (January 2018) Lookup in the UK (2018) Statistics.gov.uk. Available at: https://geoportal.statistics.gov.uk/datasets/0de287b886954f54b4c2ffcf514079_0/about.
- Marshall, A. (1890). Online Library of Liberty. 'Principles of Economics', [online] 8th Edition, pp.6–627. Available at: https://oll-resources.s3.us-east-2.amazonaws.com/oll3/store/titles/2088/Fetter_1375-01_EBk_v6.0.pdf.
- Melo, P.C., Graham, D.J. and Noland, R.B. (2009b) 'A meta-analysis of estimates of urban agglomeration economies', Regional Science and Urban Economics, 39(3), pp. 332–342. Available at: <https://doi.org/10.1016/j.regsciurbeco.2008.12.002>.
- Moray Council (2020). Wards and Constituencies - Moray Council. [online] Moray Council. Available at: http://www.moray.gov.uk/moray_standard/page_52901.html
- National Institute of Economic and Social Research, NIESR (2022). 'Productivity in the UK: Evidence Review First report of the UK Productivity Commission'. [online] Available at: <https://www.niesr.ac.uk/wp-content/uploads/2022/06/Productivity-in-the-UK-Evidence-Review.pdf>
- Newland, L., Bagdatoglou, D.G., Sabri, H. and McKnight, L. (2025). 'Housing affordability and productivity'. [online] GOV.UK. Available at: <https://www.gov.uk/government/publications/housing-affordability-and-productivity/housing-affordability-and-productivity-accessible-version>
- Nomisweb.co.uk. (2025). 'labour force survey - national and regional - headline indicators - Nomis - Official Census and Labour Market Statistics.' [online] Available at: <https://www.nomisweb.co.uk/datasets/nrhix>
- North Yorkshire Council (2023). Welcome to North Yorkshire Council. [online] www.northyorks.gov.uk. Available at: <https://www.northyorks.gov.uk/welcome-north-yorkshire-council>
- Office for National Statistics (2012). 'Annual population survey (APS) QMI' - Office for National Statistics. [online] Ons.gov.uk. Available at: <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/annualpopulationsurveyapsqmi>.
- Office for National Statistics (2013). 'User information guide'. [online] Available at: https://doc.ukdataservice.ac.uk/doc/7463/mrdoc/pdf/7463_bres_user_guide.pdf [Accessed 27 Mar. 2026].
- Office for National Statistics (2015). 'What is the productivity puzzle?' [online] Ons.gov.uk. Available at: <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/whatistheproductivitypuzzle/2015-07-07>.

- Office for National Statistics (2024). 'Business Register and Employment Survey : open access - Nomis - Official Census and Labour Market Statistics'. [online] [www.nomisweb.co.uk](https://www.nomisweb.co.uk/datasets/newbres6pub). Available at: <https://www.nomisweb.co.uk/datasets/newbres6pub>.
- Office for National Statistics (2025) 'Nominal and real regional gross value added (balanced) by industry' *Ons.gov.uk*. Office for National Statistics. Available at: <https://www.ons.gov.uk/economy/grossvalueaddedgva/datasets/nominalandrealregionalgrossvalueaddedbalancedbyindustry>.
- Ons.gov.uk. (2024). 'Business Register Employment Survey (BRES) QMI - Office for National Statistics'. [online] Available at: <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/businessregisteremploymentsurveybresqmi>.
- Pesaran, M.H. (2004). 'General Diagnostic Tests for Cross Section Dependence in Panels'. *SSRN Electronic Journal*, [online] No. 1240(1240). doi: <https://dx.doi.org/10.2139/ssrn.572504>
- Rudiger Ahrend, Alexander Lembecke and Abel Schumann (2017). " The Role of Urban Agglomerations for Economic and Productivity Growth" *International Productivity Monitor*, Centre for the Study of Living Standards, vol. 32, pages 161-179, Spring. Available at: <https://ideas.repec.org/a/sls/ipmsls/v32y20179.html>.
- Shalizi, C.R. (2025) 'The Truth about Linear Regression', p. 7-407. Available at: <https://www.stat.cmu.edu/~cshalizi/TALR/TALR.pdf>.
- Smith, A. (1991). 'Political Transformation, Urban Policy and the State in London's Docklands': *GeoJournal*, Vol. 24, No. 3, *Old Trends and New Impulses in Europe's Urban Affairs*, pp. 237-246 doi: <https://www.jstor.org/stable/41145196>.
- Stock, J.H. and Yogo, M. (2002). 'Testing for Weak Instruments in Linear IV Regression'. [online] *Ssrn.com*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=346941.
- Suffolk County Council (n.d.). *Borough, district, parish and town councils*. [online] Suffolk County Council. Available at: <https://www.suffolk.gov.uk/council-and-democracy/borough-district-parish-and-town-councils>.
- SW Herts Joint Strategic Plan. (n.d.). *The South West Herts Area*. [online] Available at: <https://www.swhertsplan.com/the-south-west-herts-area>
- University of Southampton (2024): 'Skewness and Kurtosis: Maths and Stats'. [online] *Soton.ac.uk*. Available at: <https://library.soton.ac.uk/skewness-and-kurtosis>.
- Van der Panne, G. (2004) 'Agglomeration externalities: Marshall versus Jacobs, *Journal of Evolutionary Economics*', 14(5), pp. 593–604. Available at: <https://doi.org/10.1007/s00191-004-0232-x>.
- Wales' Local Authorities (2009). *Local Authorities :: Level The Playing Field*. [online] *Leveltheplayingfield.wales*. Available at: <https://mail.leveltheplayingfield.wales/local-authorities>
- Wooldridge, J. (2009). 'Introductory Econometrics: A Modern Approach'. 4th edition. [online] Available at: https://commecsinsitute.edu.pk/wp-content/uploads/2024/08/IntroductoryEconometrics_AModernApproach_FourthEdition_Jeffrey_Wooldridge.pdf.
- www.westofengland-ca.gov.uk*. (2026). *Who we are - West of England Combined Authority*. [online] Available at: <https://www.westofengland-ca.gov.uk/about-us/who-we-are/>.

Zheng, X.-P. (2001). '*Determinants of agglomeration economies and diseconomies: Socio-Economic Planning Sciences*', 35(2), pp.131–144. Available at: doi: [https://doi.org/10.1016/s0038-0121\(00\)00008-2](https://doi.org/10.1016/s0038-0121(00)00008-2).