Tracing agglomeration effects at regional level

A statistical analysis of international regional data

Gerald Holtham

Cardiff School of Management, Cardiff Metropolitan University

Robert Huggins

School of Geography and Planning, Cardiff University

Anwar Hossain

Department of Geography and Environment, University of Dhaka, Bangladesh

Why regions, cities and localities?

- Regions, cities and localities are spatial units that compete to attract investment
- They are levels at which knowledge is circulated and transferred, resulting in agglomerations, or clusters, of industrial and service sector enterprises
- We have selected regions of differing geographical size but aiming for broadly comparable degrees of scope for economic policy
- This seems the natural locus for an analysis of agglomeration effects

What Makes a Competitive Economy?

- Competitiveness is increasingly being understood in terms of creativity, knowledge and social conditions, rather than just traditional factor endowments

- Many elements influence competitiveness and these should be controlled for when attempting to isolate the effect of a single variable like agglomeration

- Relevant factors include:
  - the quality of physical infrastructure,
  - public and private investment in human capital,
  - institutional capacity,
  - social capital,
  - innovation and research facilities,
  - accessibility to markets, and so on.

.

Competitiveness schema 1

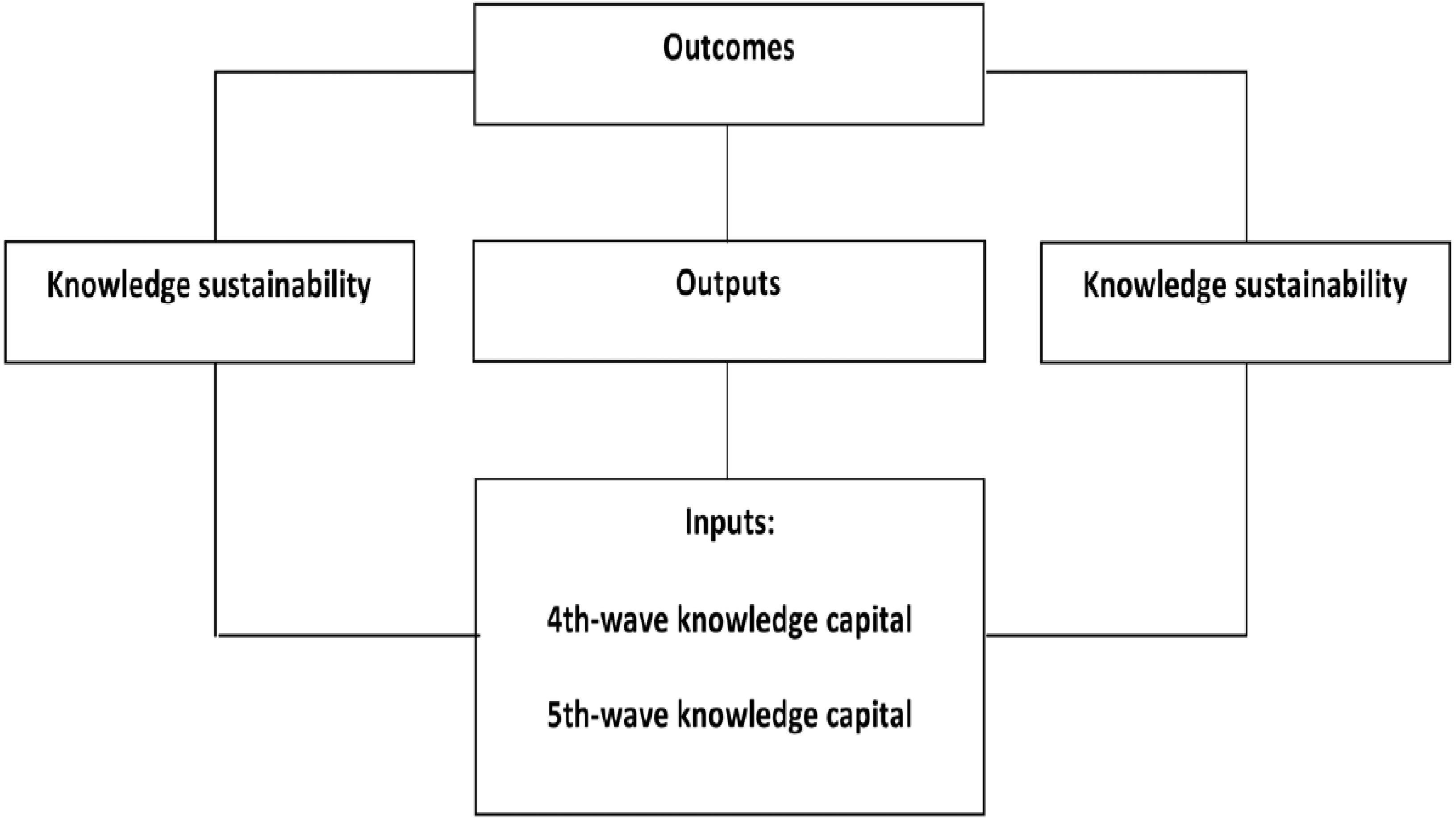
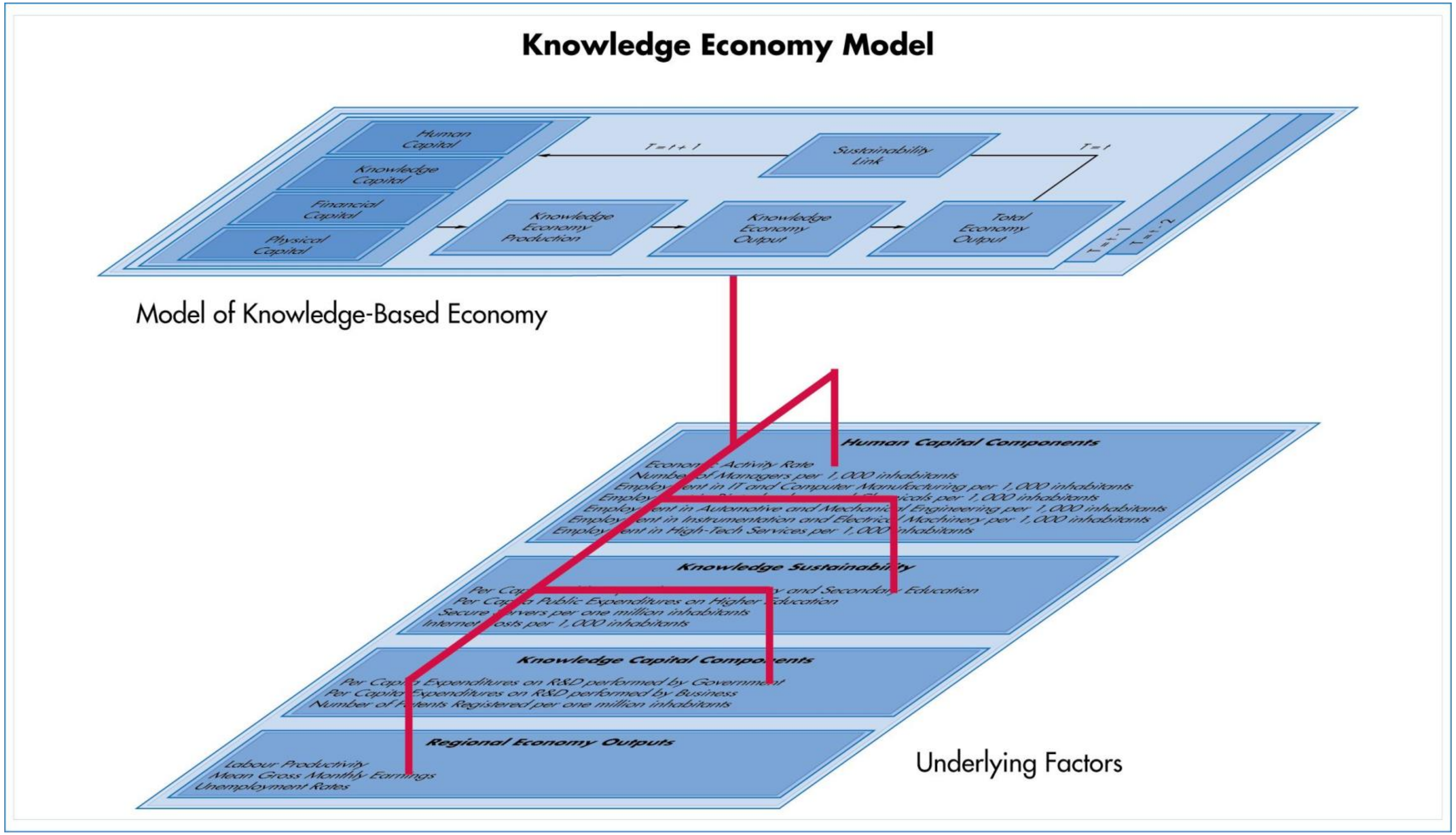


Figure 2.3 Framework underpinning the World Competitiveness Index of Regions (WCIR)

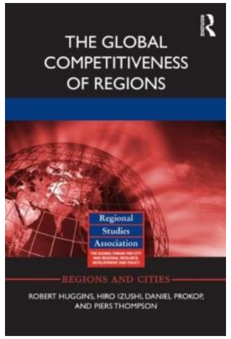
Competitiveness schema 2



WCIR data base

- Data set of 20 variables measuring or affecting competitiveness for over 400 hundred sub-state regions around the world
- Current analysis follows work published in the Welsh Economic Review that used the data set to find which variables commonly thought to indicate “competitiveness” were in fact correlated with regional success.
- Success is measured as a weighted geometric mean of GVA per head, labour productivity, monthly average wages and the economic activity rate. Correlation with GVA is 91%
- We add measures of regional agglomeration to the data set for the current analysis.

World Competitiveness Index of Regions – Original Indicators



Fifth-wave knowledge capital inputs

- employment in IT and computer manufacturing per 1,000 employees
- employment in biotechnology and chemicals per 1,000 employees
- employment in high-technology services per 1,000 employees
- per capita private equity investment.

Fourth-wave knowledge capital inputs

- employment in automotive and mechanical engineering per 1,000 employees
- employment in instrumentation and electrical machinery per 1,000 employees
- economic activity rate
- number of managers per 1,000 employees
- per capita expenditures on R&D performed by government
- per capita expenditures on R&D performed by business
- number of patents registered per one million inhabitants.

Knowledge sustainability

- per capita public expenditures on primary and secondary education
- per capita public expenditures on higher education
- secure servers per one million inhabitants
- internet hosts per 1,000 inhabitants
- broadband access per 1,000 inhabitants.

Outputs/outcomes

- labour productivity
- mean gross monthly earnings
- unemployment rates.

Agglomeration Economies

- If agglomeration economies are important, they will make workers in large cities and their environs more productive than workers in smaller cities and regions.
- There are three possible explanations for why firms continue to produce in cities even when the cost of doing business is high:

  - ➢ The city has some natural advantage, such as a port.
  - ➢ As cities grow the concentration of people and jobs lead to efficiency gains and cost savings for firms - agglomeration economies.
  - ➢ The presence of a talented and flexible labour force makes it easier for entrepreneurs to start new businesses. Activities requiring high skill may tend to cluster. This phenomenon is a case of “sorting” rather than “agglomeration” (Carlino 2011)

The Potential Over-Estimation of Agglomeration Economies (correlation is not causation!)

- There are two potentially important sources of over estimation of agglomeration economies: more productive places may attract more people, and more productive people may sort themselves into large cities.
- Large cities may draw people, especially highly skilled ones, leading to a potential overestimation of city size's effect on city productivity
- We take two agglomeration measures per region: 1. the population of the largest city in the region; 2. the overall population density of the region
- Our prior hypothesis is that any effect of regional population density on productivity may be less due to sorting than is the apparent effect of city size – but this is untested

Empirical Evidence on Sorting

- To date the evidence on the impact of sorting is mixed:
  - Gaubert (2018) finds that nearly half of the productivity advantage of large cities is due to firm sorting, the rest coming from agglomeration economies.
  - Eeckhout et al. (2014) find that large cities disproportionately attract both high- and low skilled workers, while average skills are constant across city size. This pattern of spatial sorting is consistent with extreme-skill complementarity, where the productivity of high-skilled workers and of the providers of low-skilled services are mutually enhanced.
  - Combes et al. (2012) find that although firms are more productive on average in larger cities, firm selection cannot explain spatial productivity differences.
  - Behrens et al. (2014) that elements relating to (1) more talented individuals sorting into large cities, (2) large cities selecting more productive entrepreneurs and firms, and (3)agglomeration economies combine to produce positive complementarities.

Regression results 1

- Regression analysis is work in progress. Cross section regressions are run on some 300 sub-state regions in the WCIR data base drawn from 32 countries.
- The poorest regions are excluded after cluster analysis showed that a different model was appropriate to regions at different stages of development.
- Initial OLS regressions suggested a significant effect of population density but not of city size. However these results were marred by considerable heteroscedasticity rendering significance tests invalid.
- Re-estimation by Estimated Generalised Least Squares changed the finding. City size was significant with 95 per cent confidence; the confidence level for population density fell to only 91 per cent. The adjusted coefficients (beta) suggested the effects were of similar magnitude.

Regression results 2

Dependent variable: geo mean GVA per head, labour prod, mean monthly wage, activity rate.

EGLS Regression

Adj Rsq: 0.95
F: 172
DWS 1.68

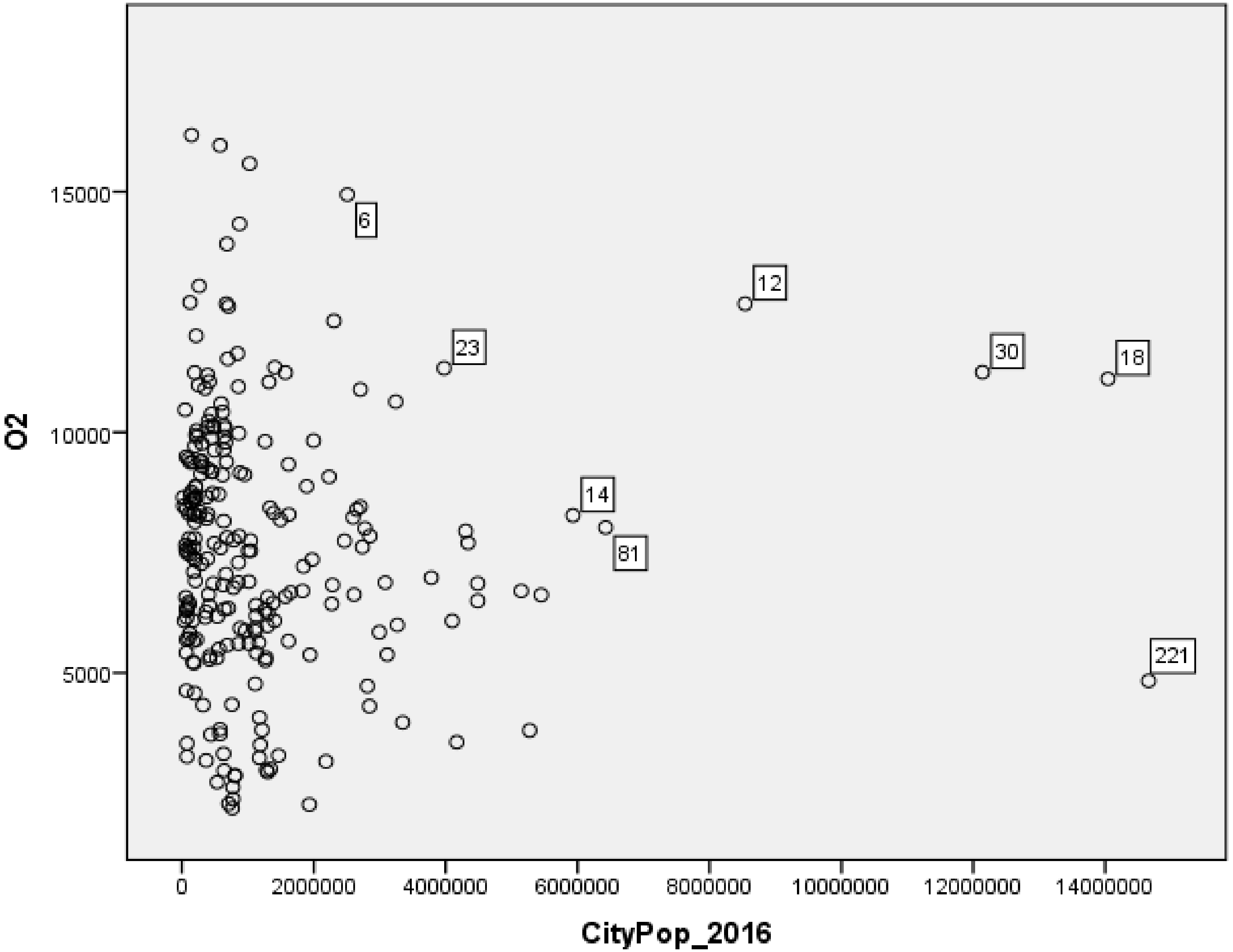
Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	489.450	96.726		5.060	0.000
Employment in Automotive and Mech Eng	5.253	2.406	0.055	2.184	0.030
Employment in High-Tech Services	19.450	3.014	0.136	6.452	0.000
Per Capita Expenditures on govt R&D	1.815	0.637	0.052	2.847	0.005
Per Capita Expenditures on business R&D	0.190	0.206	0.022	0.922	0.357
Per Capita Private Equity Investment	3.869	0.981	0.101	3.946	0.000
Higher Education Per Capita Public Expenditures	1.947	0.351	0.185	5.545	0.000
Broadband Access per 1,000 people	4.807	0.334	0.415	14.371	0.000
RegionPop density	0.272	0.156	0.031	1.744	0.083
City Pop	6.830E-05	0.000	0.034	2.004	0.046
DUS	5325.095	262.307	0.519	20.301	0.000
DLUX	8243.249	2360.609	0.049	3.492	0.001
DBEL	1149.507	612.216	0.027	1.878	0.062
DCAN	1738.865	351.884	0.094	4.942	0.000
DFRA	974.251	326.090	0.044	2.988	0.003
DCZR	-704.227	251.442	-0.054	-2.801	0.006
DSWE	-1803.765	403.696	-0.083	-4.468	0.000
DITA	2157.141	190.012	0.168	11.353	0.000
DSWI	2765.003	1135.287	0.034	2.436	0.016
DIRE	1743.484	694.558	0.036	2.510	0.013
DSPA	1374.936	290.654	0.068	4.730	0.000
DOST	1428.091	573.900	0.036	2.488	0.014
DGRE	1516.645	346.579	0.064	4.376	0.000
DHUN	-856.253	209.466	-0.064	-4.088	0.000
DPOR	1287.643	336.054	0.056	3.832	0.000
DEST	-1118.281	372.116	-0.045	-3.005	0.003
DPOL	-1428.203	117.328	-0.238	-12.173	0.000
DLITH	-1192.064	231.512	-0.080	-5.149	0.000
DLAT	-909.162	249.506	-0.057	-3.644	0.000
DNZ	1976.446	671.768	0.041	2.942	0.004

Outliers and the cluster 1

The top three regions:
Bridgeport CT, Luxemburg and St Jose CA have only a small city population.

Outliers:

6: Brussels
12: New York city
23: Los Angeles
30: Ile de France
18: London
14: Ontario
81: Madrid
221: Istanbul

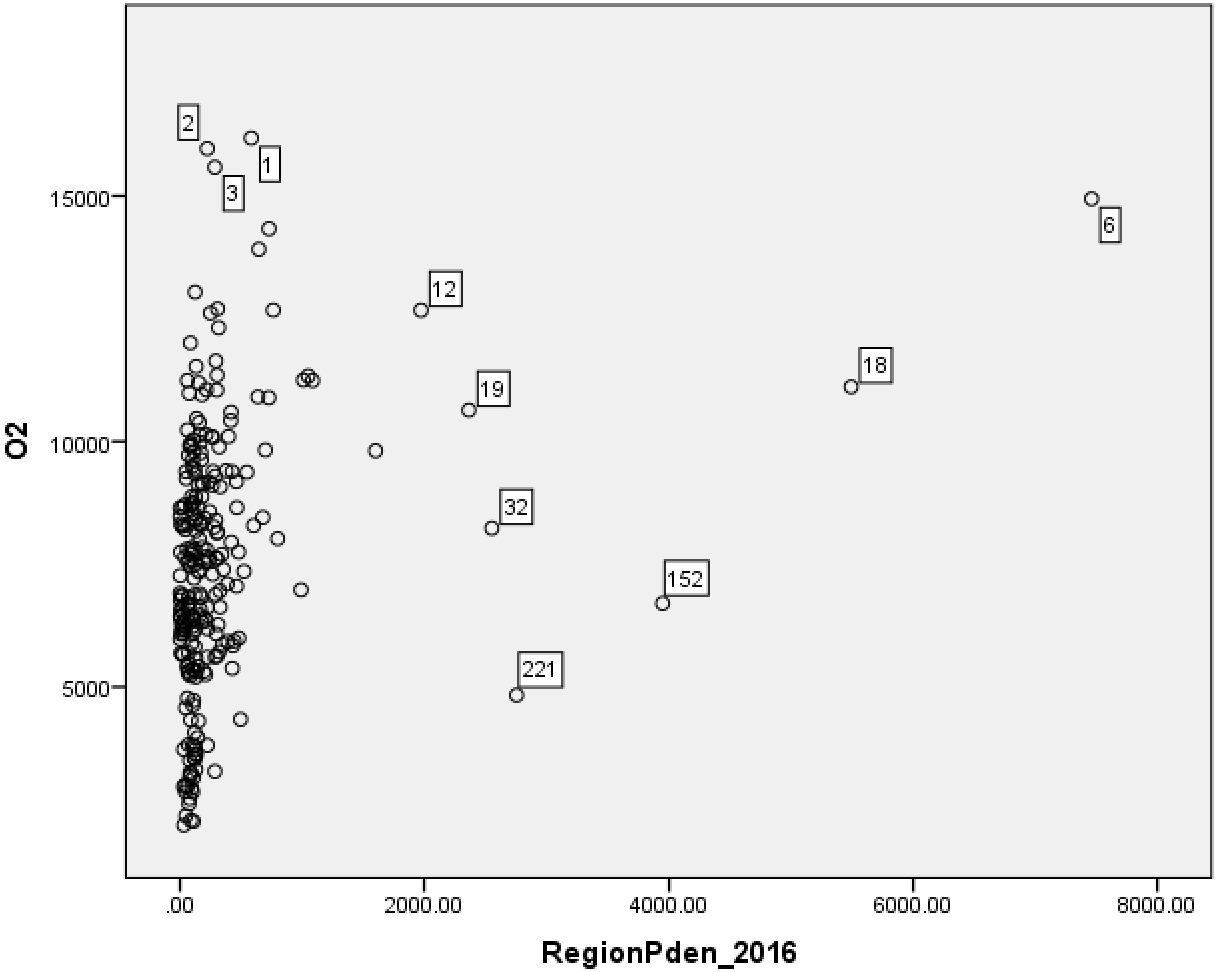


Outliers and the cluster 2

The top three regions: Bridgeport CT, Luxemburg and St Jose CA have only a small population density.

Outliers:

6: Brussels
12: New York city
18: London
19: Hamburg
32: Prague
152: Berlin
221: Istanbul



Summary

- A correlation between city size and an index of productivity, GVA per head and wages is detectable across regions. A similar correlation exists for population density but is less statistically significant. The bivariate correlations are not visually impressive
- We supposed regional population density would be less prone to sorting effects than the size of the largest city but while the results do not allow us to distinguish agglomeration from sorting, the outlying position of Brussels suggests sorting is important in the case of population densities.
- Unsurprisingly any agglomeration effect appears to be considerably weaker (lower beta) than other elements in driving productivity such as access to broadband, expenditure on higher education and the extent of employment in high-tech services
- Results are highly provisional since coefficient estimates are not stable and regression statistics indicate residual heteroscedasticity. Given data shortages at the regional level, the probability of omitted variables bias is high.

Continuing research

- Refinement of both specification and econometrics is in order
- Neighbourhood effects are inadequately represented by country dummy variables. A large city, for example, could have an influence on areas in other administrative regions that are closer than parts of its own region
- Introducing proximity or neighbourhood effects explicitly using spatial regression techniques might reduce the importance of country dummies, increase the significance of the city-size variable and should reduce residual heteroscedasticity
- Non-linearity in variables could also be important; is there an optimal density before negative effects of congestion start to weigh?
- Current regressions are pure cross-section which means some explanatory variables should have been broadly stable for a long time for correlations to be meaningful, given that their effect is long run. Even then causation could be in either direction. We are currently assembling more data to allow temporals lags to be introduced.

References

Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3), 507-553.

Carlino, G.A. (2011) Three Keys to the City: Resources, Agglomeration Economies, and Sorting, Phildelphia Fed Business Review Q3.

Combes, P. P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6), 2543-2594.

Eeckhout, J., Pinheiro, R., & Schmidheiny, K. (2014). Spatial sorting. *Journal of Political Economy*, 122(3), 554-620.

Gaubert, C. (2018). Firm sorting and agglomeration (No. w24478). National Bureau of Economic Research.