Does Regional Economic Growth Depend on Proximity to Urban Centres?

Rudiger Ahrend,
Abel Schumann

OECD

OECD REGIONAL DEVELOPMENT WORKING PAPERS

This series is designed to make available to a wider readership selected studies on regional development issues prepared for use within the OECD. Authorship is usually collective, but principal authors are named. The papers are generally available only in their original language English or French with a summary in the other if available.

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

This document has been produced with the financial assistance of the European Union. The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to either gov.contact@oecd.org or the Public Governance and Territorial Development Directorate, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

Authorised for publication by Rolf Alter, Director, Public Governance and Territorial Development Directorate, OECD.

DOES REGIONAL ECONOMIC GROWTH DEPEND ON PROXIMITY TO URBAN CENTRES?

Rüdiger Ahrend and Abel Schumann[1,2]

ABSTRACT

This paper analyses the spatial patterns of regional economic growth in Europe over the 1995 to 2010 period. It finds that regions, which contain large urban agglomerations, have been growing significantly faster than those that do not. Furthermore, proximity to large urban agglomerations has been positively correlated to economic growth. Halving travel time to a large urban agglomeration is associated with a 0.2 to 0.4 percentage points increase in annual per capita growth. More generally, the study also shows that measures of population density are positively correlated to growth. Among the different measures, by far the best predictor of growth between 1995 and 2010 is the maximum population density of a region.

Tables

Figures

DOES REGIONAL ECONOMIC GROWTH DEPEND ON PROXIMITY TO URBAN CENTRES?

1. INTRODUCTION

It is well known that urban agglomerations are centres of economic activity, which are more productive on a per capita basis than less densely populated areas. Consequently, it can be expected that regions, which contain large urban agglomerations, tend to have higher per capita incomes than those that do not. This raises the question of whether the economic performance of a region may be more generally dependent on its location relative to urban agglomerations. This paper therefore studies how patterns of human settlement are related to regional economic growth in Europe between 1995 and 2010. First, it describes general trends in regional population patterns and in regional GDP growth between 1995 and 2010. Next, it analyses the impact of urban agglomerations on the economic growth of the regions in which they are located. Subsequently, it studies the role of urban agglomerations that are located outside of a given region. In particular, it focuses on the distance from a region to the closest large urban agglomeration, distinguishing between geographical distance and distance measured in terms of travel time. Lastly, the paper relaxes the focus on urban agglomerations and examines how the spatial population distribution more generally is related to economic performance.

The following novel results are presented in the paper:

- The more sparsely populated a region, the more geographically concentrated is its population according to concentration measures such as the Gini coefficient.

- The number of strongly under-performing regions in terms of per capita GDP decreases between 1995 and 2010, whereas the number of regions with per capita GDP strongly above the median increased over the same time. This is a positive result in so far as a particular emphasis is placed on improving economic conditions in the poorest regions.

- Regions, which contain urban agglomerations with more than 500,000 inhabitants, have been growing approximately 0.2 percentage points faster on a per capita basis than regions without them.

- Travel time is a significantly better predictor of economic growth than simple geographic distance.

- A halving of travel time to large urban agglomerations is associated with an increase in yearly GDP growth per capita by approximately 0.2 to 0.4 percentage points.

- The relative centrality of a region (as calculated by its population density relative to the population density of the surrounding areas) is strongly positively correlated to economic growth per capita.

- Most regional measures of population distribution are correlated to economic per capita growth rates, but the maximum population density of a region is by far the strongest predictor of

economic growth and dominates all other measures when their effect on growth is estimated jointly.

The paper is structured in the following way. Section 2 presents the data that is used in the analysis and briefly outlines the empirical strategy. Section 3 presents the empirical results. More precisely, Section 3.1 discusses changes in the per capita income distribution. Section 3.2 focuses on the role of urban agglomerations for economic growth within regions. Section 3.3 studies the effects of distance from large urban agglomerations and focuses on the difference between geographic distance and travel time. Section 3.4 analyses the predictive power of several different measures of population density on regional economic growth. Section 4 concludes.

2. DATA AND METHODOLOGY

2.1 Data

Regional GDP and Population Data

The OECD defines sub-national territorial units at two different levels; the higher Territorial Level 2 (TL2) and the lower Territorial Level 3 (TL3). In total there are 362 TL2 regions and 1792 TL3 regions in OECD countries. As there is very little data on the TL3 level available outside of Europe, this analysis focuses on approximately 600 European TL3 regions from 18 countries for which GDP growth rates and other data is available from 1995 to 2010. Excluded from the analysis are those regions that are not part of the mainland of a country, such as exclaves and oversea territories.

The population of TL3 regions in the sample varies from 20 000 to 6.3 million with the median region having 429 000 inhabitants. Similarly, there are large differences in the average area that a region covers. The geographically smallest region has an area of 9 km² whereas the largest covers 59 000 km². Most common are areas of several thousand km². The median population density of 129 inhabitants per km² is typical for intermediate regions that contain some mid-sized urban centres but also sparsely populated areas. However, especially in more densely populated countries, such a population density occurs also in regions that are entirely rural without containing any urban centres.

Urban Agglomeration Data

This study uses the OECD definition of Functional Urban Areas (FUA) as developed in OECD (2012), which is based on urban cores and economically integrated surroundings. It identifies urban cores as contiguous built up areas and uses commuting patterns to delineate the surrounding areas that belong to the FUA. An area is considered part of a Functional Urban Area if at least 15% of employees commute into the urban core.

Travel Time and Distance Data

The study uses travel time and travel distance to urban centres above varying population thresholds as explanatory variables for economic growth. Travel time is measured as the number of minutes required to travel by car from the centroid of a TL3 region to the centre of the closest FUA above the respective population threshold. Travel distance indicates the distance in kilometres between those two points using

the fastest road connection. Travel time assumes normal road conditions without congestion. The data is based on route planning information from Google Maps that has been collected using Google Maps' Application Programming Interface (API). When centroids of regions do not lie exactly on a road, the closest point on a road has been used as start or end point of a route, respectively.

The median travel distance to the centre of the closest FUA with at least 500 000 inhabitants is 105 km. Due to a small number of very remote regions, the mean distance is larger with 127 km. Mean and median travel times are 76 and 105 minutes, respectively. Across all observations, the average travel speed as predicted by Google Maps is 78 km/h. The correlation between travel time and travel distance is 0.86. This relatively high correlation indicates a fairly homogenous quality of road infrastructure across Europe.

Travel distance is also closely correlated to geographical distance. Mean and median geographical distances to the closest FUA with at least 500 000 inhabitants are 116 km and 88 km, respectively. Thus, the median travel distance is approximately 20% larger than the median geographical distance. The correlation between travel distance and geographical distance is 0.71 and thus somewhat lower than the correlation between travel distance and travel time.

Looking at distance to very large FUAs with more than 2 million inhabitants gives a similar picture. Mean travel time is 390 minutes and median travel time is 307 minutes. Mean and median travel distance are 628 km and 506 km, respectively. The corresponding mean and median geographical distances are 432 km and 386 km.

Geographical distances are calculated as straight line distances on the globe. Figure 1 shows the distances graphically. Darker shades indicate a larger distance to the next FUA above 500 000 and 2 million inhabitants, respectively.

Figure 1. Distance to closest FUA with 500 000 inhabitants (left) and 2 million inhabitants (right)
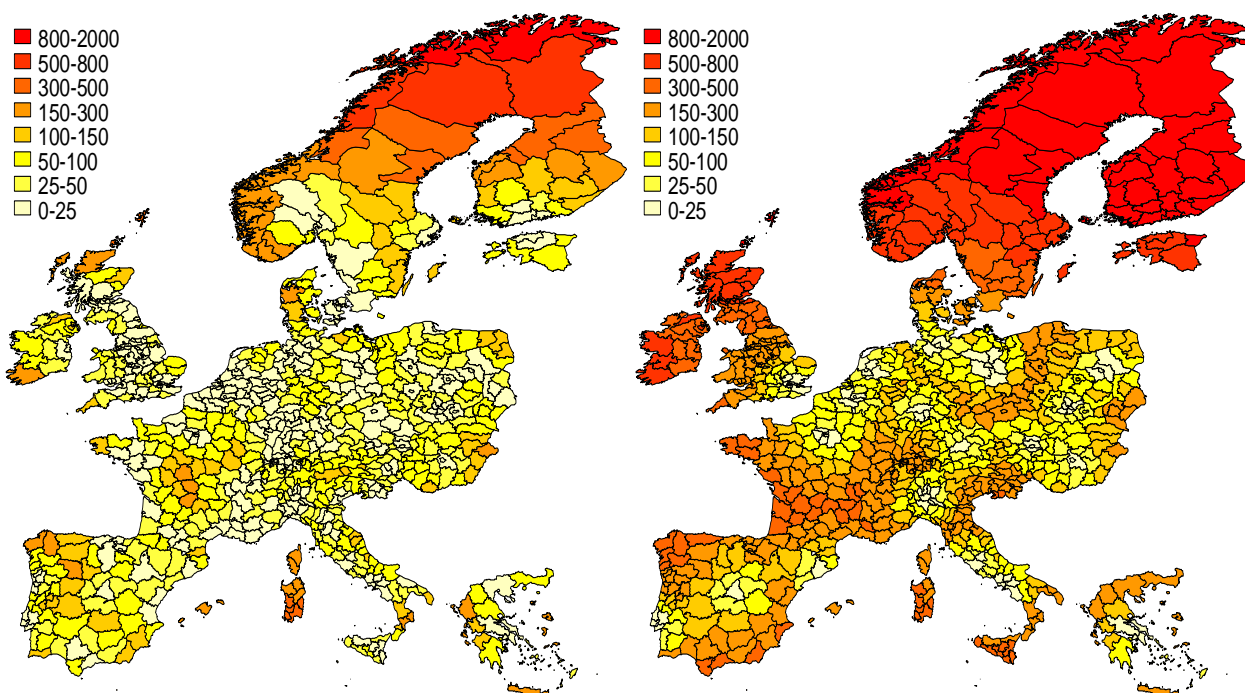
Note: Figure 1 shows the distance in kilometres to the closest functional urban area (FUA) with at least 500 000 inhabitants (left), and 2 million inhabitants (right). Darker colours indicate larger distances. With the exception of northern Europe, most regions are relatively close to FUAs with 500 000 inhabitants, but distances to large FUAs vary strongly.

Population Density

The study uses different measures of the population density distribution across space within regions. This data is primarily based on population density estimates for Europe by Gallego (2010). The estimates provide the population density for cells in a 100m grid. They are derived from land cover satellite data from the CORINE project and are available for the European Union. Their spatial resolution is much higher than that of any population data based on administrative boundaries. Therefore, they allow for a more detailed description of the population density distribution across space. In order to improve the computational tractability, the resolution of the population data has been reduced to a 1km grid.

In order to compute measures of population distribution, each of the 1km grid cells is considered a single observation and is assigned to the TL3 region into which its centroid falls. Once all grid cells for a region are extracted, it is straightforward to compute distributional measures of the population distribution across space. Among them are the population density at given percentiles and the share of land used by a given percentage of the population within a region. Using this information, it is also possible to compute the Gini coefficient of population distribution across space. A coefficient of 1 signifies perfect concentration, *i.e.* the entire population of a region living within a single grid cell. A coefficient of 0 indicates equal distribution across space.

In general, the more sparsely populated a region, the more geographically concentrated is its population as measured by the population Gini coefficient. The correlation of regional population density and its population Gini coefficient is strong at -0.65. Figure 2 shows the cumulative population distributions over space for Outer London West in the UK and Cáceres in Spain. Outer London West is a very densely populated area with 3 952 inhabitants per km². It has an unusually low Gini coefficient of 0.45 and displays a very homogenous population distribution across space. In contrast, Cáceres is very sparsely populated (21 inhabitants per km²) and has a population Gini coefficient of 0.89. It displays the unequal population distribution that is typical for sparsely populated areas. The unweighted average Gini coefficient of all regions in the sample is 0.82.

Figure 2. Cumulative population distribution of Outer London West (left) and Cáceres (right)
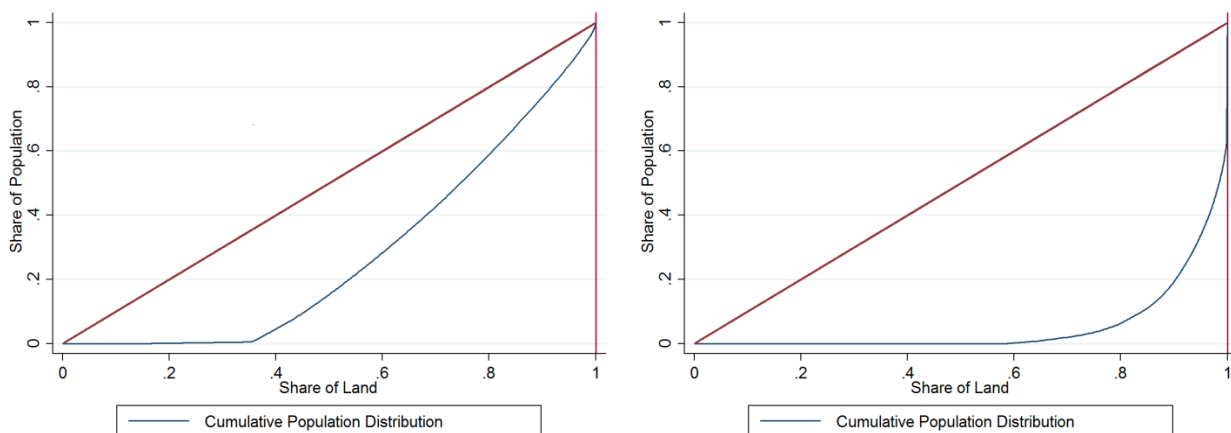
2.2 Empirical Strategy

The empirical strategy in this paper is based on cross-section regressions of the average annual regional per capita GDP growth rate between 1995 and 2010 on an outcome variable of interest. Most

specifications include initial log-per capita GDP in 1995 and a set of country dummies as control variables. The baseline regression is given by

$$\overline{\Delta}\, pcGDP_i^{95-10} = \alpha + \beta_1 x_i + \beta_2 \log\left(pcGDP_i^{95}\right) + \sum_{c=1}^{C} \gamma_c dum_i^c + \varepsilon_i,$$

where $\overline{\Delta}pcGDP_i^{95-10}$ indicates the average annual growth rate of per capita GDP between 1995 and 2010 in region i, x_i is the respective explanatory variable, $\log(pcGDP_i^{95})$ is a control variable for log per capita GDP at the beginning of the observation period, and $\sum_{c=1}^{C} dum_i^c$ is a set of dummy variables for country c. The set of country-dummies implies that, generally, within-country effects are estimated. This ensures that the estimates are not affected by country-wide developments that are unrelated to regional characteristics. It is furthermore a way of dealing with the problem of shocks that are clustered on the country-level and which could lead to a severe underestimation of the estimated standard errors. Controlling for initial GDP is required in many cases to avoid that estimates are biased by regression to the mean.

3. EMPIRICAL RESULTS

This section presents the main results of the paper. It is split into four parts. In the first part, general growth trends are described. The second part analyses the role of urban agglomerations within regions for economic growth, while the third part looks at the role of urban agglomerations outside of regions. The fourth part compares different measures of population distribution as predictors for regional economic growth.

3.1 Income Distribution across Regions over Time

Table 1 presents baseline estimates of per capita GDP growth on per capita GDP in 1995. In all estimated specifications, the coefficient on initial per capita GDP is strongly negative and highly significant. This implies that there is generally mean reversion among regional per capita GDP levels. Regions that have been over-performing or under-performing in terms of per capita GDP in 1995 generally moved closer to average per capita GDP over the subsequent 15 years. However, the estimated variance of the deviation of regional per capita GDP from country averages does not change significantly. It is virtually identical in 1995 and 2010. Expressed in the terminology of the economic growth literature, this implies that that there has been beta convergence but no sigma convergence of per capita GDP. In other words, levels of within-country regional inequality have not been declining over the observation period. These results mirror the findings by Young et al. (2008), who argue that there has been beta convergence but no sigma convergence of county per capita GDP levels in the United States.

Table 1.   Baseline estimates

	Aver. yearly per capita GDP growth	Aver. yearly GDP growth	Aver. yearly per capita GDP growth	Aver. yearly GDP growth	Aver. yearly population growth
	(1)	(2)	(3)		
Log p.c. GDP in 1995	-0.54*** (0.20)	0.40** (0.17)	-0.64*** (0.22)	0.22 (0.20)	
Log Pop. 95			0.14*** (0.04)	0.19*** (0.05)	0.12*** (0.02)
Constant	8.46*** (1.80)	-0.12 (1.56)	7.41*** (2.33)	-1.18 (2.04)	-0.22 (0.37)
Country-FE	YES	YES	YES	YES	YES
N	662	687	603	603	603

***, **, and * indicates significance at 99%-level, 95%-level, and 90%-level, respectively.

In contrast, when using GDP growth instead of per capita GDP growth to estimate the effect of initial GDP levels on subsequent growth a clear divergence trend becomes possible. When using the following specification

$$\overline{\Delta GDP}_i^{95-10} = \alpha + \beta_1 \log(pcGDP_i^{95}) + \sum_{c=1}^{C} \gamma_c dum_i^c + \varepsilon_i,$$

the coefficient on $\log(pcGDP_i^{95})$ is positive and highly significant. This indicates that regions with a higher per capita GDP level in 1995 have been growing faster in absolute terms over the following 15 years. As there has been no divergence in per capita GDP levels, the reason for the divergence in absolute GDP levels lies in population growth. Regions that had a high GDP per capita in 1995 were gaining population relative to regions with a low GDP, which resulted in above-average GDP growth.

Figure 3. Demeaned distributions of log per capita income (kernel density estimates)
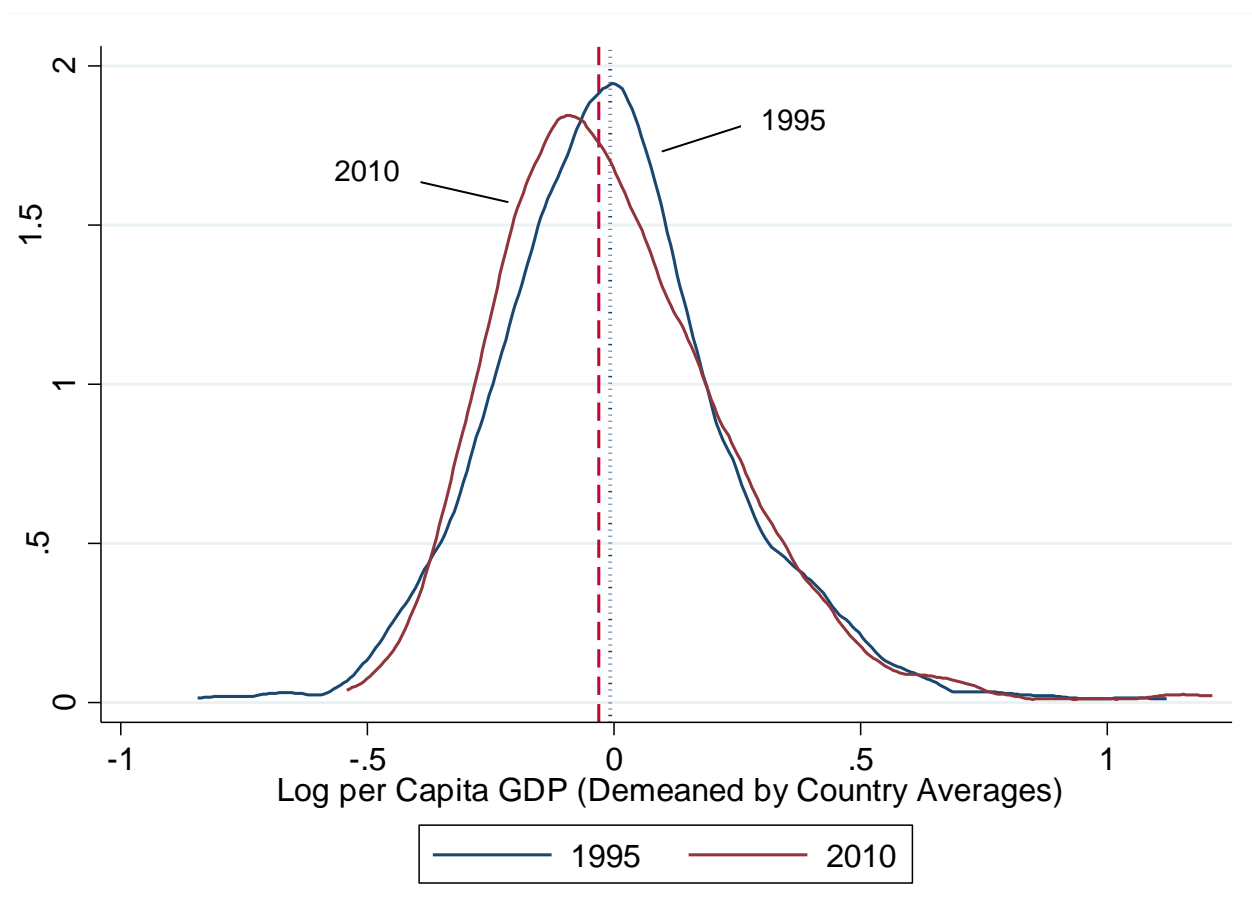
Figure 3 shows kernel density estimates of the distribution of log per-capita income in 1995 and 2010. Both distributions are normalized to mean 0. The median per capita income in 1995 is shown by the dotted line; the median per capita income in 2010 is shown by the dashed line. The figure illustrates that there is little change in the overall dispersion of within-country regional per capita incomes. Both distributions are skewed to the right, but the degree of skewness is significantly larger in 2010 (skewness coefficients for

1995 and 2010 are 0.52 and 1.01, respectively). As the distributions are normalized to mean 0, the increased skewness can be seen in Figure 3 by the leftward shift of the median in 2010 (the dashed line) relative to the median 1995 (the dotted line). In 2010, regions on the left-hand side of the 2010-median are closer to the median than those on the right-hand side of the median. In contrast, for the distribution of 1995 incomes, the average distance to the 1995-median was more similar on either side of the median. In practical terms, the increase in positive skewness implies that in 2010 fewer regions strongly underperformed than in 1995, while more regions performed strongly above average.[3]

Figure 4 shows the development of regional per capita GDP on a map. The darker the colour of a region, the higher is its per capita GDP. As in case of Figure 3, per capita GDP has been demeaned by country-averages for the respective years. While the effect is not easily observable, close inspection of the figure shows that there are fewer very lightly coloured (*i.e.* underperforming) regions in 2010 than in 1995. Note that regions left entirely blank indicate missing data.

Figure 4. Regional per capita GDP in 1995 (left) and 2010 (right) as percentage of country average
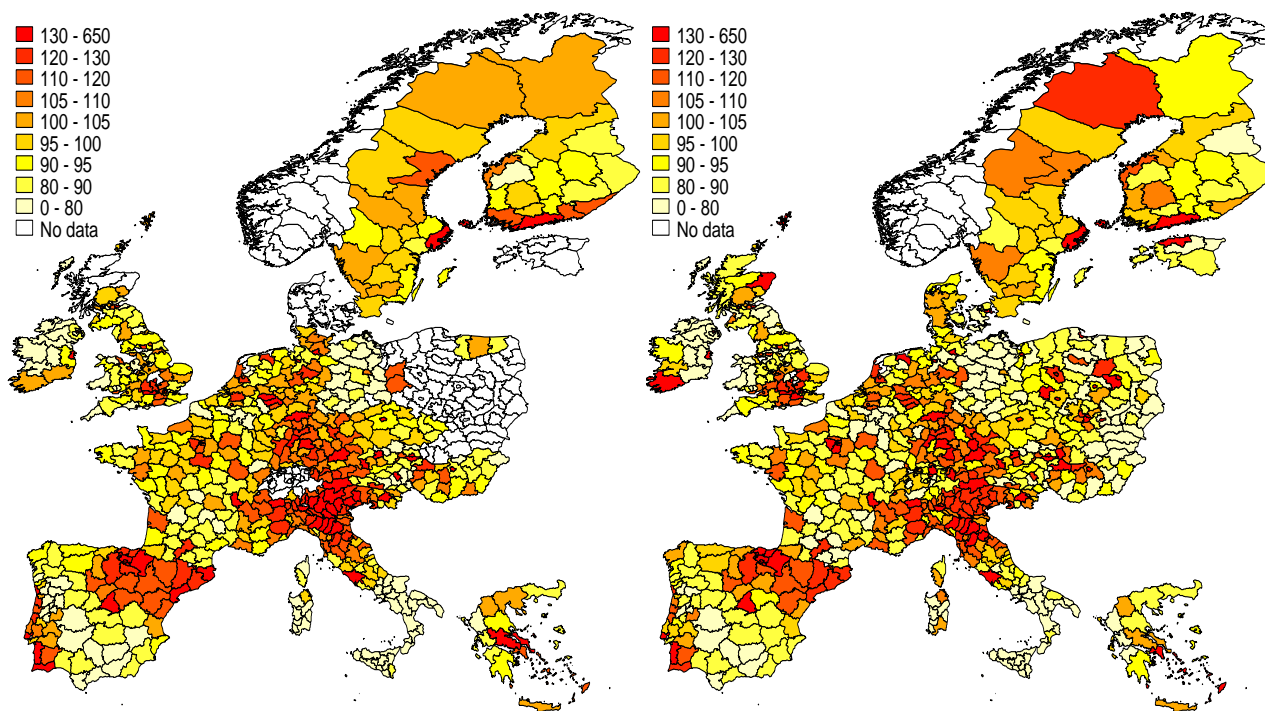
Note: Figure 4 shows regional per capita GDPs in 1995 (left) and 2010 (right) as percentage of the country averages for the respective year.

3.2 Cities within Regions

One of the reasons for the increase in skewness is a catching-up of strongly under-performing regions, which leads to a lower number of outliers at the lower end of the distribution. Another reason could be the role of large urban agglomerations. Regions that contain an urban agglomeration above 500 000 inhabitants had a per capita income that was approximately 21% larger than the respective country average

---

3.   The change in the distribution of per capita GDP is not driven by the recent financial and economic crisis. The distribution for 2007 (not shown) looks very similar to the one for 2010 and has a similar skewness coefficient of 0.95 (compared to 1.01 in 2010). For absolute GDP levels, neither dispersion nor skewness of the distribution changes between 1995 and 2010.

in 1995. Nevertheless, regions that contain such agglomerations had a much higher per capita GDP growth over the subsequent 15 years. Table 2 shows how the presence of a large urban agglomeration affects regional per capita GDP growth. The explanatory variables are specified as dummy variables that take on the value 1 if a region contains a FUA with the respective size and 0 otherwise. Regions that contain urban agglomerations above 200 000 inhabitants grew 0.16 percentage points faster than regions without such urban agglomerations. For regions with urban agglomerations above 500 000 inhabitants, the difference in yearly per capita growth rates is approximately 0.2 percentage points (Column 1) and for those with urban agglomerations above 2 million inhabitants it is more than 0.5 percentage points (Column 2).

While specifications (1) and (2) are preferred, because the inclusion of initial per capita GDP as control variable accounts for potential regression-to-the-mean, there are also some reasons why this control variable may be excluded. Per capita GDP differences in 1995 among regions with and without large urban agglomerations are large. Therefore, controlling for initial per capita GDP could cause the coefficient on the presence of urban agglomerations to be significant and positive even if the impact of urban agglomerations had not changed at all. Therefore, specifications (3) and (4) present the same estimates without controlling for initial GDP levels. Comparing specification (1) with specification (3) shows that omitting the control for initial per capita GDP levels only slightly reduces the estimated growth contribution of urban agglomerations with more than 500 000 inhabitants whereas the estimated coefficient on urban agglomeration with more than 2 million inhabitants becomes much smaller. Nevertheless, specification (3) and (4) show that large urban agglomerations have been growing much faster in absolute terms between 1995 and 2010 and not only conditional on their initial per capita GDP.

Table 2.   Economic growth and size of the largest FUA within a region

	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth
	(1)	(2)	(3)	(4)
Agglomeration >500 000	0.23*** (0.10)		0.19** (0.09)	
Agglomeration >2 000 000		0.54** (0.26)		0.29 (0.23)
Agglomeration 500 000- 2 000 000		0.28*** (0.10)		0.17* (0.09)
Agglomeration 200 000- 500 000	0.16** (0.07)	0.16** (0.07)	0.13* (0.07)	0.13* (0.07)
Per Capita GDP in 1995	-0.65** (0.24)	-0.67*** (0.24)		
Constant	9.25*** (2.60)	9.43*** (2.61)	2.20*** (0.07)	2.20*** (0.07)
Country-FE	YES	YES	YES	YES
N	603	603	603	603

***, **, and * indicates significance at 99%-level, 95%-level, and 90%-level, respectively.

3.3 Travel Time

The previous estimates have shown that regions containing large urban agglomerations have been growing faster between 1995 and 2010. This section estimates whether a correlation between economic growth and proximity to large urban agglomerations exists, too. It is in spirit similar to Veneri and Ruiz (2013) who analyse the effects of proximity between rural and urban regions, but differs from their analysis by using a larger set of distance measures and considering only large urban agglomerations.

Table 3 presents the results of a series of regressions that show the effect of travel times and geographical distance on economic growth. Column (1) shows the estimate for the following specification.

$$\overline{\Delta pcGDP}_i^{95-10} = \alpha + \beta_1 I_i^{45-90} + \beta_2 I_i^{90-180} + \beta_3 I_i^{180-300} + \beta_4 I_i^{>300}$$

$$+ \beta_5 \log(pcGDP_i^{95}) + \sum_{c=1}^{C} \gamma_c dum_i^c + \varepsilon_i,$$

It includes four dummy variables that indicate whether a region is within the given number of minutes by car from the next urban agglomeration with at least 2 million inhabitants. The base category is regions that are within less than 45 minutes of such urban agglomerations. It shows that cities that are within 45 to 90 minutes of such agglomerations have been growing approximately 0.5 percentage points slower per year than those that are within less than 45 minutes. For regions that are further away from large urban agglomerations, the negative difference in growth is even larger.

Specification (2) also estimates a relation between travel time and economic growth but uses a log-linear specification of travel time instead of a set of dummy variables to model the relation between the two variables. The estimated coefficient on travel time is small and insignificant. The reason behind this result is that the negative relationship between travel time and economic growth breaks down at more than roughly 400 minutes travel time. Specification (4) takes this into account by restricting the sample to regions that are within 8 hours (a day's drive) by car to an urban agglomeration with 2,000,000 inhabitants. It shows that for distances below that threshold a robust negative relationship between distance and economic growth exists. A 100% increase in travel time is associated with an annual per capita GDP growth rate that is approximately 0.2 percentage points lower.

Specifications (3) and (5) repeat the exercise but use logs of geographic distance to the next urban agglomeration of more than 2 000 000 inhabitants as explanatory variable. Although less pronounced, the emerging pattern is similar: while there is an overall negative relationship, a statistically significant result can only be found for regions that are not too far away from the next large urban agglomeration.

Lastly, specification (6) compares the explanatory power of travel time and distance for per capita economic growth by including both variables. When controlling for distance, the interpretation of the travel time coefficient changes somewhat. Conditional on distance, travel time becomes a measure of road connectivity. In this specification, the coefficient on log-travel time remains significant at the 90% level and is exactly twice as large as the coefficient on log-distance, which turns insignificant. Both point estimates on the two coefficients are directly comparable in their magnitude because both variables are in logarithmic terms. Therefore, the coefficients can be interpreted as the effect of a percentage change of the explanatory variable on the outcome. As both variables also have similar standard deviations, a fixed percentage change has a comparable importance for both variables. Therefore, the difference in the estimated coefficients implies that actual distance is of lower importance than travel time for economic growth.

Although travel time between regions and large urban agglomerations is strongly correlated with economic growth, the picture is less clear with respect to smaller urban agglomerations. While some specifications (not shown) suggest that closeness to small and medium sized urban agglomerations had a positive effect on economic growth between 1995 and 2010, coefficients are imprecisely estimated and typically not statistically significant.

Table 3. Distance to urban agglomerations with at least 2 million inhabitants

	Full sample			Restricted sample		
	(1)	(2)	(3)	(4)	(5)	(6)
45-90 Min	-0.50*					
	(0.29)					
90-180 Min	-0.62**					
	(0.29)					
180-300 Min	-0.79**					
	(0.31)					
>300 Min	-0.87***					
	(0.32)					
Log travel Time		-0.02		-0.216**		-0.17*
		(0.08)		(0.097)		(0.10)
Log Distance			-0.07		-0.14**	-0.08
			(0.05)		(0.06)	(0.06)
Log p.c. GDP 95	-0.81***	-0.00**	-0.80***	0.000	-0.61**	-0.64**
	(0.21)	(0.00)	(0.22)	(0.000)	(0.29)	(0.29)
Constant	11.63***	3.86***	11.17***	4.122***	9.71***	10.66***
	(2.05)	(0.505)	(2.08)	(0.868)	(3.30)	(3.47)
Cut-off time	--	--	--	0-480 Min	0-480 Min	0-480 Min
Country FE	YES	YES	YES	YES	YES	YES
N	545	545	545	385	385	385

***, **, and * indicates significance at 99%-level, 95%-level, and 90%-level, respectively.

Estimated travel times are likely to be subject to some degree of measurement error. Travel times have been computed from the point of a road that is closest to the centroid of a TL3 region to the centre of the nearest metropolitan area. Both the choice of start and end point introduce some degree of measurement error. In some cases, the centroid of a TL3 region will lie in a remote area, while in others it will fall next to a large road. If a TL3 centroid is unusually remote, travel times will be overestimated, if it is unusually close to a large road, travel times will be underestimated. As there is no reason to assume that the measurement error is correlated to the true outcome, this is so-called classical measurement error, which leads to attenuation bias. This biases coefficients towards zero (*i.e.* the bias is negative for a positive coefficient and positive for a negative coefficient). It is worth noting that as long as there is no criterion to define representative travel times for a region, any choice of starting and end point introduces similar biases. The current choice has the advantage that the introduced biases due to measurement error are

obvious and identical in their direction. They imply that the estimated coefficients on travel time will be lower bounds of the true estimates, *i.e.* that the true effects of travel time are even larger than the estimates obtained from this specification.

Several other issues potentially bias the estimates above. One potential source of bias is related to reverse causality. Travel time was measured in 2013, after the end of the observation period. If policy makers respond to expected or actual growth rates by investing in infrastructure, it is possible that the differences in travel time are at least partly caused by per capita GDP growth. If fast growing regions receive more infrastructure investments, the estimate on travel time would be negatively biased (*i.e.* its effect would be overestimated). Correspondingly, if slow growing regions receive more infrastructure investments, the effects would be underestimated. Furthermore, travel time might be correlated to other factors that affect growth rates over the observation period. Such factors would introduce omitted variable bias in the estimates. Especially in sparsely populated regions, for example, the actual investment into road construction can have a sizable impact on per capita GDP. It might therefore be the case that part of the higher growth in better connected regions comes from the investment that made them better connected in the first place. In this case, the coefficient on travel time would be positively biased.

In order to provide evidence on the importance of the endogeneity problems mentioned above, travel time is instrumented by geographical distance.[4] Table 4 presents IV estimates of specifications (2) and (3) of Table 3. The point estimates on travel time increase considerably in magnitude relative to the OLS specifications.[5] This could indicate that the OLS specification suffers from a positive bias, which in the given case would mean that the estimated coefficients are biased towards zero. Such a bias could occur for example if under-performing regions receive more infrastructure investments or if there are omitted variables such as the quality of governance that are positively correlated to growth rates and the amount of infrastructure investment that a region receives. It is furthermore likely that distance as instrument helps to alleviate the attenuation bias introduced by measurement error. As only the difference in travel time predicted by distance is used in the second-stage regression, any measurement error due to the road conditions close to a regions' centroid is removed.

---

4.   The exogeneity assumption for geographical distance being a valid instrument requires that any correlation between geographical distance and regional growth is only due to the effect of geographical distance on travel time. As always, this assumption is statistically untestable and should be carefully assessed. It can be ruled out that geographical distance is affected by reverse causality because distance is time invariant. Therefore, one of the primary sources of endogeneity can be eliminated. However, omitted variables remain potentially a problem even though there are no obvious reasons why other variables should be correlated to both economic growth and distance to large agglomerations once travel time is taken into account.

5.   The F-statistic of the first-stage regression of travel time on distance is 184.91 and therefore well above any commonly required threshold for relevance.

Table 4.   Distance to urban agglomerations with at least 2 million inhabitants (IV estimates)

	(1)	(2)
Log travel time (instrumented with log distance)	-0.17	-0.41**
	(0.16)	(0.20)
Log per capita GDP in 1995	-0.00***	-0.00*
	(0.00)	(0.00)
Constant	3.09***	4.31***
	(1.13)	(1.33)
Cut-off Time	--	0-480 min
Country FE	YES	YES
N	545	385

***, **, and * indicates significance at 99%-level, 95%-level, and 90%-level, respectively.

3.4 Population Density

3.4.1 Population Density and Economic Growth

The estimates presented above were related to the locations of urban agglomerations. This section relaxes the focus on urban agglomerations and looks at more general measures of population distribution across space. In particular, it compares how different measures of population density are correlated to regional economic growth.

Population density of a region is strongly correlated to per capita GDP growth between 1995 and 2010, no matter what measure of population density is used. Figure 5 shows population density in 2010 on the right-hand side and average per capita growth rates demeaned by country on the left-hand side. Higher population density and stronger per capita GDP growth are indicated by darker colours. These figures give the impression that more densely populated areas have typically had an above-average growth performance.

Figure 5. Population density (left) and per capita GDP growth rates relative to country averages (right)
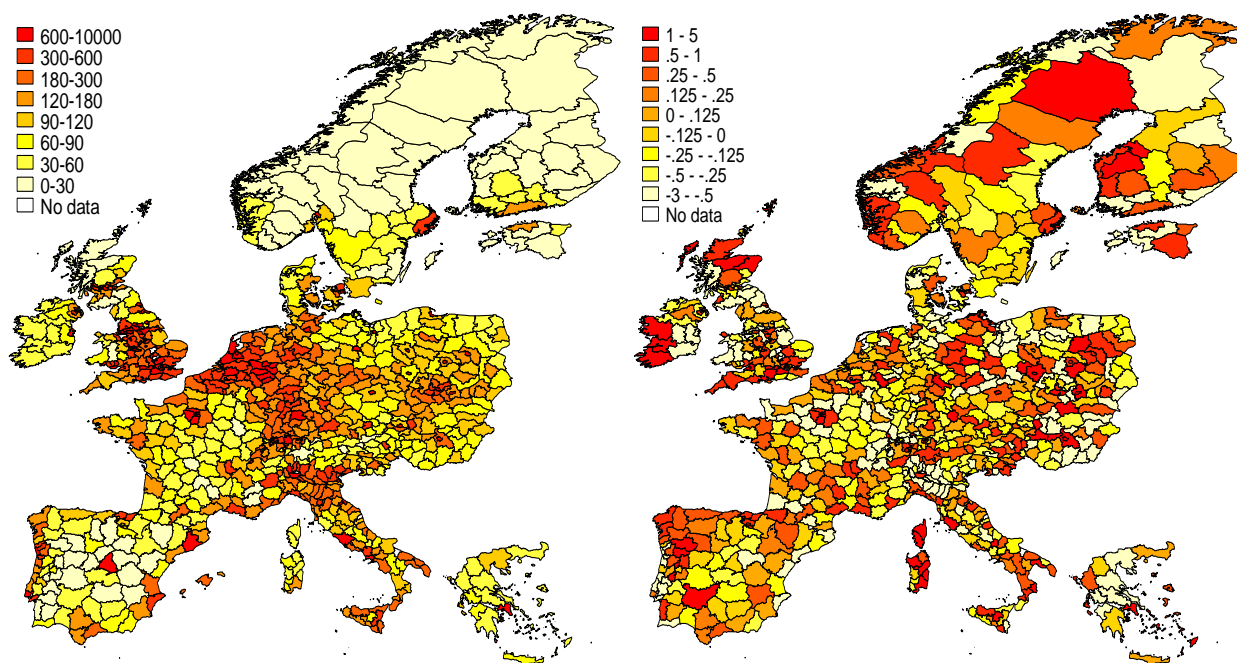
Note: Figure 5 shows population regional population densities in inhabitants per km² on the left-hand side and average regional growth rate between 1995 and 2010 (or latest available data) on the right-hand side.

Econometric analysis confirms this visual impression. Table 5 presents estimates of different measures of population density on regional per capita growth. Remarkably, by far the strongest predictor of per capita GDP growth is the maximum population density of a region. All measures of population density use logarithmic scales. Therefore, the point estimate of 0.32 for maximum population density in specification (1) implies that a 1% increase in the maximum population density in a region is associated with a 0.0032 percentage point increase in the average yearly growth rate of the region. Over the entire 15 years period, this implies that a 100% increase in the maximum population density corresponds to an approximately 5 percentage points higher per capita GDP growth.

Table 5.   Population density and economic growth

	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log maximum population dens.	0.32*** (0.08)						0.29*** (0.08)
Log 99 pctile population dens.		0.06* (0.03)					-0.01 (0.05)
Log 90 pctile population dens.			0.10*** (0.02)				0.09* (0.05)
Log mean population dens.				0.11** (0.04)			-0.06 (0.12)
Log median population dens.					0.11*** (0.03)		0.05 (0.03)
Log population						0.11* (0.04)	0.06 (0.05)
Log per capita GDP in 1995	-0.80*** (0.23)	-0.62** (0.22)	-0.77*** (0.21)	-0.75** (0.23)	-0.64** (0.21)	-0.64** (0.22)	-0.95*** (0.25)
Constant	8.20*** (2.29)	8.53*** (2.37)	10.10*** (1.92)	9.87*** (2.01)	8.94*** 2.268	7.87*** (1.95)	9.22*** (2.58)
Country FE	YES	YES	YES	YES	YES	YES	YES
N	603	603	601	603	570	618	570

***, **, and * indicates significance at 99%-level, 95%-level, and 90%-level, respectively.

Specifications (2) – (5) use other measures of the distribution of population density across space. Specification (2) uses the population density at the 99th percentile of the distribution, specification (3) uses the 90th percentile, specification (4) the median population distribution and specification (5) the average population density. Several aspects are noticeable. All estimates are significant at reasonable significance levels, but of an order of magnitude smaller. More importantly, none of the coefficients is robust to the inclusion of maximum population density as covariate: all coefficients except for maximum population density become smaller and insignificant. This pattern holds no matter if all population density measures are included in the estimation (as in specification (7)) or if each of the variables is included individually together with maximum population density (estimates not shown). Furthermore, the coefficient on maximum population density remains virtually unchanged even if dummies for the existence of urban agglomerations of different sizes are included in the estimation (estimates not shown).

Table 6.   Population density and economic growth (excluding the top quintile in terms of maximum population density)

	(1)	(2)
Maximum population dens.	0.20**	0.22**
	(0.10)	(0.09)
99 percentile population dens.		-0.03
		(0.06)
90 percentile population dens.		0.07
		(0.05)
Log average population dens.		0.01
		(0.04)
Log median population dens.		-0.03
		(0.12)
Log population		0.10
		(0.06)
Log per capita GDP in 1995	-1.15***	-1.30***
	(0.22)	(0.24)
Constant	12.07***	13.33***
	(1.89)	(2.55)
Country FE	YES	YES
N	542	513

***, **, and * indicates significance at 99%-level, 95%-level, and 90%-level, respectively.

The results above are not primarily driven by the regions that have particularly high population densities, although regions with the highest maximum population density affect the results most strongly. The pattern that is presented in Table 5 holds across the entire distribution of the maximum population density distribution. Table 6 shows specifications (1) and (7) of Table 5, but excludes the regions in the top decile of the maximum population density. The magnitude of the coefficients decreases by approximately 30%, but they remain both statistically and economically significant. Excluding further deciles reduces the precision, but leaves the point estimates in roughly the same range.

The maximum population density measure provides information only about a very small area in a region. Typically, it is the centre of the biggest city of a region that is its most densely populated part. It is unlikely that this small area is entirely responsible for the difference in regional growth that is predicted by its population density. Two related explanations for the observed effect exist:

- It is possible that the maximum population density is correlated to other characteristics of a region that determine growth without actually having a causal influence on those factors. For example, the population distribution across space could be correlated to regions' economic structure. Under this interpretation, the maximum population could be a predictor of economic growth without having any explanatory power of its own. It is, for example, likely that specialisations in the service sector or knowledge intensive industries allow for higher population

density than a focus on manufacturing, which is generally more land intensive than service provision. Therefore, maximum population density could be a good predictor of the sectoral composition of an urban agglomeration. As the service sector has been growing much faster than manufacturing over the past 15 year, the observed correlation between maximum population density and growth could be a result of the differences in the sectoral composition.

- An alternative explanation to this hypothesis is that the maximum population density of a region has effects on economic or demographic characteristics in a region, which in turn affect the growth rate. In this context, the maximum population density could provide information on how "urban" the population centre of a region is. It is widely acknowledged that – on average - larger agglomerations increase the productivity of their inhabitants, even when controlling for skill levels of the population or sectoral composition of production (see *e.g.* Combes, Duranton and Gobillon, 2011, or Ahrend, Farchy, Kaplanis and Lembcke, 2014). Maximum population density may hence be a good proxy for agglomeration benefits, and possibly a better one than the population or population density of a city. Such agglomeration benefits may then spill over to the rest of the region.

However, care should be exercised when generalising the results on population density beyond the analysed countries as the urban structure of European cities often differs from urban structures in non-European OECD countries. In particular, European urban cores tend to be old, and large parts of cities often developed several centuries ago. It is possible that the reasons behind the predictive power of the maximum population density are related to the particular urban structure of European cities, and it is therefore uncertain whether similar results would be found for countries such as the United States.

3.4.2 Centrality of regions

The previous results have shown that population patterns, and in particular those related to large urban agglomerations, are correlated to regional economic growth up to sizable distances. Furthermore, population density within regions is correlated to their economic growth. Therefore, an obvious next step is to analyse whether population density around regions is correlated to economic growth. Table 6 shows that this is not the case although specification (1) appears to provide some superficial support for it by finding a significant positive correlation between regional economic growth and the population density in a 10km band around a region. However, specification (4) shows that after controlling for population density within the region, any significant correlation disappears. It appears likely that the positive correlation in specification (1) stems only from the fact that population in a 10km band around a region is closely correlated to the population density within the region.

When using the population density in a 50 km and 100 km band around a region, respectively, specifications (2) and (3) do not find strong correlations. However, after controlling for population density within the region, population density in 50 km and 100 km is strongly negatively correlated to regional economic growth between 1995 and 2010. This implies that – holding population density within the region constant – a region has been growing more strongly if it had a more sparsely populated surrounding.

Table 7. Population density around regions

	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth	Aver. yearly per capita GDP growth
	(1)	(2)	(3)	(4)	(5)	(6)
Log pop. density in 10 km band	0.08** (0.03)			-0.03 (0.05)		
Log pop density in 50 km band		-0.03 (0.05)			-0.18*** (0.06)	
Log pop density in 100 km band			-0.111* (0.065)			-0.21*** (0.07)
Pop. density				0.31*** (0.09)	0.18*** (0.04)	0.16*** (0.04)
Per capita GDP in 1995	-0.61*** (0.21)	-0.49** (0.22)	-0.45 (0.22)	-0.79*** (0.23)	-0.75*** (0.22)	-0.75*** (0.22)
Constant	8.63*** (2.29)	7.82** (2.36)	7.88*** (2.36)	10.36*** (2.49)	10.63*** (2.42)	10.78*** (2.40)
Country-FE	YES	YES	YES	YES	YES	YES
N	603	603	603	603	603	603

***, **, and * indicates significance at 99.9%-level, 99%-level, and 95%-level, respectively.

A possible explanation for this result could imply that densely populated regions attract economic activity at the cost of surrounding areas. Firms, for example, might move from sparsely into more densely populated areas in order to increase their market potential. This could lead to a spatial concentration of economic activity in centrally located regions. In this context, the centrality of a region would be defined by its own population density relative to that of the surrounding regions. Conditional on the population density of a region, its centrality would increase with decreasing population densities in surrounding regions. If this explanation is correct, it could imply that on the regional level a pattern is emerging that is similar to the one described by Christaller (1933) for the urban level. Regions are ordered in a hierarchical structure with centres of economic activity surrounded by less active regions.

4. CONCLUSION

This paper has presented evidence on the characteristics of regional economic growth over the period 1995 – 2010. Several patterns have emerged. For a start, the distribution of regional per capita incomes has changed. While strongly over-performing and strongly under-performing regions where roughly evenly distributed in 1995, the share of strongly under-performing regions decreased over the 1995 – 2010 period, while the share of strongly over-performing regions increased. Although this has not reduced the overall variance of cross-regional per capita incomes, it is a positive result in so far as a particular emphasis is placed on improving economic conditions in the poorest regions.

The variance of regional per capita incomes has remained stable over the study period, but the variance of absolute GDP levels has increased due to an ongoing concentration process of economic activity in regions that had already above-average GDP levels in 1995. The concentration process has been primarily due to an inflow of population into regions with above-average per capita GDP levels (which also tend to be the largest regions in terms of population and hence absolute GDP).

Subsequently, the paper analyses how regional growth is correlated to the presence of large urban agglomerations. It shows that regions containing such agglomerations have been growing much faster on a per capita basis than those that do not. The larger the urban agglomeration within a region, the faster the growth the region experienced.

When analysing how economic growth is related to large urban agglomerations outside of a region, two insights are gained. First, a clear relationship between distance to large urban agglomerations and economic growth exists. Second, travel time is a stronger and more precise predictor of per capita growth than geographical distance. A 1% decrease in travel time is associated with an increase in GDP that is approximately 50% larger than the increase in GDP associated with a 1% decrease in distance.

The paper also shows that population density is positively correlated with regional economic growth. Among all measures of population density, maximum population density is by far the strongest predictor of economic growth over the 1995 to 2010 period. The point estimate on its coefficient is much larger than other measures such as the average population density of a region. Furthermore, it is the only variable that is robust to the inclusion of other measures of population density as control variables. While population density within a region is positively correlated with economic growth, population density around a region is negatively correlated with growth. Thus, more central regions as measured by their own population density relative to the population density of surrounding areas have been growing more strongly than less central regions.

REFERENCES

Ahrend, R., E. Farchy, I. Kaplanis, and A. C. Lembcke "What Makes Cities More Productive? Evidence on the Role of Urban Governance from Five OECD Countries", *OECD Regional Development Working Papers* 2014/05, OECD Publishing, Paris, http://dx.doi.org/10.1787/5jz432cf2d8p-en.

Brezzi, M., L. Dijkstra and V. Ruiz (2011), "OECD Extended Regional Typology: The Economic Performance of Remote Rural Regions", *OECD Regional Development Working Papers* 2011/06, OECD Publishing, Paris, http://dx.doi.org/10.1787/5kg6z83tw7f4-en.

Carlino, G. A., and E. S. Mills (1987), The Determinants of County Growth, *Journal of Regional Science*, *27*(1), 39-54.

Christaller, W. (1933), Die zentralen Orte in Süddeutschland, *Fischer*, Jena.

Combes P.-P., G. Duranton and L. Gobillon, 2011, "The Identification of Agglomeration Economies", *Journal of Economic Geography*, Vol. 11, pp. 253-266.

Gallego, F. J. (2010), "A Population Density Grid of the European Union."*Population and Environment* 31.6, 460-473.

Garcilazo, E. (2013), "Growth Trends and Characteristics of OECD Rural Regions", *OECD Regional Development Working Papers*, No. 2013/10, OECD Publishing, Paris, http://dx.doi.org/10.1787/5k4522x3qk9q-en.

OECD (2012), *Redefining "Urban": A New Way to Measure Metropolitan Areas,* OECD Publishing, Paris http://dx.doi.org/10.1787/9789264174108-en.

Veneri, P. and V. Ruiz (2013), "Urban-to-Rural Population Growth Linkages: Evidence from OECD TL3 Regions", *OECD Regional Development Working Papers*, No. 2013/03, OECD Publishing, Paris, http://dx.doi.org/10.1787/5k49lcrq88g7-en.

Young, A. T., M. J. Higgins and D. Levy (2008), "Sigma Convergence versus Beta Convergence: Evidence from US County-Level Data", *Journal of Money, Credit and Banking* 40.5 (2008): 1083-1093.