Economic growth and geographic proximity

C.L.F. Attfield, Edmund S. Cannon, D. Demery, Nigel W. Duck* 

Department of Economics, University of Bristol, Alfred Marshall Building, 8 Woodland Road, Bristol, BS8 1TN, UK

Received 25 May 1999; accepted 16 December 1999

Abstract

Using data on 92 countries, 90 European regions, and 48 US states we present results suggesting that physical distance between economies has little role to play in explaining the spatial correlation of growth rates. © 2000 Elsevier Science S.A. All rights reserved.

Keywords: Growth

JEL classification: O40; R11; F43

1. Introduction

One of the stylised facts of economic growth is that faster growing economies tend to be geographically clustered, e.g. in Europe or S.E. Asia. This might be because geographic proximity per se (i.e. the physical distance between economies) influences the extent to which economies’ growth rates are correlated. This would arise if, for example, transport costs are, or have been important, since such costs clearly depend upon distance. However, clustering of growth rates would also arise if factors which might affect growth, e.g. climate, political and legal traditions, and religious beliefs, were themselves spatially correlated.

In this paper we test whether it is geographic proximity per se which matters. We consider three data sets. First we use regional data from eight European countries from 1950 to 1990. Since regions from the same country tend to be relatively close together, we can test the proximity hypothesis by comparing the importance of distance versus the importance of national zero/one dummy variables, where the latter have a clear socio-economic interpretation. The other two data sets are observations

*Corresponding author. Tel.: +44-117-928-8406; fax: +44-117-928-8577.
E-mail address: n.w.duck@bris.ac.uk (N.W. Duck)
on (a) 92 countries from 1960 to 1990, and (b) 48 US states from 1900 to 1990.\footnote{The growth data are all average per capita (per worker for the country data) growth of income over 1900 to 1990 for the US, 1950 to 1990 for the European regions, and 1960 to 1990 for the countries of the world. The source of data for the US states was Barro and Sala-i-Martin (1995) from which we also obtained the Census regional classification on which the dummy variables were defined. The European regions data were based on Barro and Sala-i-Martin and the Penn World Tables. The country data were obtained from the Penn World Tables. The longitude and latitude series on which the great circle distances are based was taken from Philip’s World Atlas (Reference Edition) (1996, Chancellor Press, London). We used the latitude and longitude of state capitals for the US; the values given in the Atlas for that region for European regions or a city/town close to its geographic centre; the values for the country’s capital for the world data, except the US where we used Chicago and Pakistan (Karachi), and Switzerland (Geneva). The data are available from the authors on request.} In these two cases it is less easy than for the European data to characterise an alternative hypothesis to geographical proximity. For (a) we use continent zero/one dummy variables, and for (b) we use zero/one dummy variables based on a conventional regional grouping of those states. Although these dummies are, to some extent, arbitrary, they do have the effect of grouping data points known to have some similarities. Furthermore, if the proximity hypothesis is correct it should be robust to alternative specifications, however arbitrary.

From all three data sets the same conclusion emerges: in the absence of wider-area dummies physical distance does help to explain the pattern of economic growth but it has almost no explanatory power when those dummies are included. We conclude that distance per se is not important. The paper is in one main section where we explain our methodology, which is based on Moreno and Trehan (1997) and Anselin (1988), and present our results.

2. Methodology and results

We derive our results through estimation of the following relationship:

\[ (I - \rho W)G = X\beta + \varepsilon \]  

where \( G \) represents an \( n \times 1 \) vector of the per worker or per capita growth rates of \( n \) economies over the data period; \( W \) represents an \( n \times n \) matrix of distance weights where each weight is greater the closer the particular economies are to each other; \( X \) is an \( n \times k \) matrix of other possible influences on growth rates; \( \beta \) is a \( k \times 1 \) vector of coefficients; \( \varepsilon \) is an \( n \times 1 \) vector of error terms assumed to be distributed \( N(0, \sigma^2) \). The parameter \( \rho \) is of especial interest: the hypothesis we are investigating is that it is significantly positive even when the wider-area dummy variables are included in the \( X \) matrix.

For each of the data sets the distance weights are based on the great circle distances between each country, region or state. The definition of a representative element of \( W \), \( w_{ij} \), is:

\[ w_{ij} = \frac{1/d_{ij}}{n \sum_{j=1}^{n} 1/d_{ij}} \]  

where \( d_{ij} \) is the great circle distance between economies \( i \) and \( j \). This definition of the weights, which
forces the weights to sum to one, amounts to assuming that it is *relative* rather than *absolute* distance which matters.

Because the growth rate of each economic unit can be affected by the growth rate of the others, the likelihood function for the relationship in Eq. (1) contains a Jacobian term, giving a likelihood of the form:

\[
L = -(n/2)\ln \pi - (n/2)\ln \sigma^2 + \ln |1 - \rho W| - \varepsilon/2\sigma^2.
\]

We obtained maximum likelihood estimates using the OPTMUM routine in GAUSS (1994). In Table 1 we present, for each data set, the estimated value of \( \hat{\rho} \) and its associated \( P \)-value when the \( X \) matrix contains (i) only a constant; (ii) a constant and the initial level of income (to capture so-called convergence); and (iii) a constant, the initial level of income and the appropriate set of wider-area dummies.

The first striking feature of Table 1 is that \( \hat{\rho} \) appears to be highly significant when no other variable is included in the model and to be very similar in all three data sets. The estimates suggest that a one

| Table 1* |
|-----------------|-----------------|-----------------|
| **Data set:**   | **World**       | **Europe**      | **US**          |
| Dependent variable: | Growth per worker | Growth per capita | Growth per capita |
| Model: \([I - \rho W]G = \text{constant} + \varepsilon\) | \(\hat{\rho}\) = 0.872 (0.000) | \(\hat{\rho}\) = 0.940 (0.000) | \(\hat{\rho}\) = 0.903 (0.000) |
| Model: \([I - \rho W]G = \text{constant} + \beta_0 \ln y_{\text{init}} + \varepsilon\) | \(\hat{\rho}\) = 0.875 (0.000) | \(\hat{\rho}\) = 0.729 (0.000) | \(\hat{\rho}\) = 0.566 (0.012) |
| Model: \([I - \rho W]G = \text{constant} + \beta_0 \ln y_{\text{init}} + \text{dummies} + \varepsilon\) | \(\hat{\rho}\) = 0.488 (0.126) | \(\hat{\rho}\) = -0.100 (0.772) | \(\hat{\rho}\) = -0.088 (0.789) |
| Dummies \( P \)-value | 0.0002 | 0.0000 | 0.0037 |

*In brackets are the \( P \)-values of the null hypothesis that the coefficient directly above is zero. \( \ln y_{\text{init}} \) is the log of the economic unit’s level of income at the start of the period. There are five continent dummies included in the results for countries; seven country dummies for the European regions; and three regional dummies for the US states. The dummies \( P \)-value is in each case the probability of finding the computed likelihood ratio test statistic of the null hypothesis that the wider-area dummies are jointly insignificant if that null were true.

Experiments with different weighting procedures, e.g. using the inverse of the distance raised to a power or dropping the assumption of the weights summing to one, made little substantial difference to our results.

The coefficient on the initial level of income variable is often used to estimate the speed of convergence (see Barro and Sala-i-Martin (1995, p. 387)). However, as Friedman (1992) emphasises, this interpretation may be invalid: measurement error in output might by itself account for the negative relationship between the log of initial output and the difference between the log of end and log of initial output.

The \( P \)-values reported on coefficients using the maximum likelihood estimation procedure mentioned above are based on numerically estimated standard errors. We also computed likelihood ratio tests of the null hypothesis that \( \rho \) equals zero and obtained essentially the same substantive conclusions.
percentage point increase in the distance-weighted growth rate in the rest of the world (or of Europe or of the US) is associated with almost a one percentage point increase in the growth rate of the country (or European region or US state). A second feature is that the estimated value of $\rho$ remains highly significant when the initial level of income is included in the $X$ matrix. However, when the wider-area dummies are included this significance disappears in all three cases. Furthermore, a likelihood ratio test of the null hypothesis that the wider-area dummies are jointly insignificant produced the $P$-values shown in the last row of the table and suggest a strong rejection of the null.

3. Conclusions

Our conclusion can be simply stated. The results from three different data sets suggest that geographic proximity per se does not have a key role to play in explaining the tendency of economic growth rates to cluster.

References

GAUSS, Aptech Systems Inc, Maple Valley, WA.