Dissertation

Regional Economic Growth Across Space and Time

Ausgeführt zum Zwecke der Erlangung eines akademischen Grades
des Doktors der Volkswirtschaften
unter der Leitung von

Prof. Dr. Michael Funke
Institut für Wirtschaftssysteme, Wirtschafts- und Theoriegeschichte
Arbeitsbereich Makroökonomik und Quantitative Wirtschaftspolitik

eingereicht an der Universität Hamburg
Fachbereich Wirtschaftswissenschaften
von

Declan Curran BA MA Econ

Hamburg 2007

Contents

Acknowledgements

I wish to express my gratitude to my PhD thesis supervisor, Prof. Dr. Michael Funke, who was pivotal in shaping and directing this project. On a personal note, I wish to thank my family for their encouragement throughout my studies and, of course, my fiancé, Véronique, for all her help and support. Last but not least, I would like to thank my colleagues in the University of Hamburg for ensuring that my stay in Hamburg was a memorable one.

Introduction

The issue of economic convergence both within or across countries has proved to be an intuitively appealing one for economists and policymakers alike. An understanding of the entire distribution of, say, income or employment within or across countries, as well as of how that distribution changes over time, allows one to establish the relative economic performance of regions and assess whether action needs to be taken stimulate economic activity in under-performing regions. For example, the tendency within countries for interregional disparities in income levels to either narrow or widen over time has serious implications for the implementation and effectiveness of aid programs and government policies geared towards regional development. Equally as important is the question of whether or not the regions that are relatively poor today are the same regions that were poor, say, 100 years ago. If this is indeed the case and some regions are persistently impoverished, then proactive measures may need to be taken by policymakers to propel regions out of this poverty trap. In this way, establishing the stylised facts of economic growth that currently prevail "on the ground" provides crucial information for policy decisions.

The phenomenal surge of interest witnessed over the last two decades in the forces that lead to economic convergence shows no sign of abating. While early contributions such as Ramsey's (1928) treatment of household optimisation over time and Harrod (1939) and Domar's (1946) efforts to integrate Keynesian analysis with elements of economic growth may have marked the inception of modern growth theory, it was the neoclassical growth model developed in the 1950s and 1960s [Solow (1956), Swan (1956), Cass (1965), Koopmans (1965)] that represented a high water mark for the topic until the late 1980s when the emergence of endogenous growth models [Romer (1986), Lucas (1988), Rebelo (1991), Aghion and Howitt (1992)] breathed new life into this field of research. The central features of the neoclassical growth model are well- known: a neoclassical production function that assumes constant returns to scale, diminishing returns to each input, and a positive smooth elasticity between the inputs is combined with a constant-savings-rate rule to generate a general equilibrium of the economy. One prediction emanating from this model has displayed considerable explanatory power across countries and regions: conditional convergence. This hypothesis posits that economies or regions with lower initial levels of real per capita GDP, relative to the long-run or steady state position, tend to experience faster growth. This stems from the assumption of diminishing returns to capital: if an economy has less capital per worker than the long-run or steady state capital per worker, any additional capital per worker will yield a relatively higher rate of return and higher growth rates than those economies possessing large quantities of capital per worker. This convergence is conditional in nature because an economy's steady state level of capital and output per worker depend that economy's savings rate, population, and position of the economy – which may vary across economies. Additional sources of variation across economies (such as initial human capital stocks, infrastructure disparities, and government investment) have been included in more recent empirical work. The main deficiency of the neoclassical model is also well-known: the assumption of diminishing capital per worker implies that in the absence of technological advances, per capita growth would eventually cease – a modelling limitation which was addressed by the inclusion of an exogenous rate of technological progress. This unsatisfactory outcome that has been addressed by the class of endogenous growth models that have emerged since the late 1980s. In these models, the long-run growth rate are determined within the model, be it through the inclusion of knowledge spillovers and human capital, research and development and imperfect competition, or the explicit modelling of technology diffusion.

As Barro and Sala-I-Martin (1995) note, one clear distinction between the growth theory of the 1960s and that of the 1980s and 1990s is that recent research pays close attention to empirical implications and the relation between the theory and the data. One aspect of this empirical scrutiny is the testing the neoclassical model's prediction of conditional convergence, while another is the testing more recent endogenous growth hypotheses pertaining to the role of human capital, research and development activity, infrastructure disparities, and technology diffusion in the growth process. One of the earliest empirical approaches to the question of convergence across economies has been that of β-convergence regression analysis, as developed by Baumol (1986), Barro and Sala-I-Martin (1992), and Mankiw et al. (1992), where average per capita GDP over a given period is regressed on the initial level of GDP per capita and (in the conditional convergence case) and a set of explanatory variables. A negative estimate for the coefficient of initial GDP per capita (β) indicates that growth rates of per capita income over the time period in question is negatively correlated with initial incomes – a finding which is interpreted as a support for the hypothesis of convergence. In recent times, an wide variety of empirical approaches have been developed to empirically test for convergence: from simple plots of measures of dispersion over time to intra-distributional dynamics using Markov chains applied to *GDP* per capita. What is more, numerous studies have revealed persistent differences in per capita income among regions. Evidence shows that some regions managed to sustain high per capita income over a long time span while other regions seemed to be trapped in a low income growth path. These persistent differences are strikingly at odds with the standard neoclassical growth model, which predicts that poorer countries usually develop faster than richer ones and that there is a tendency toward convergence in levels of *GDP* per capita. While β-convergence analysis has retained its popularity as a test for empirical convergence, not least because it can easily be augmented to include newly developed spatial econometric analytical tools, it is now possible to complement this approach with a whole host of parametric and non-parametric econometric techniques. In this way, a more refined impression of the growth process evolving over time within one economy or across economies can be obtained.

The focus of this study is the interregional growth process over time within countries. A number of studies into regional convergence over time have been carried out for US states [for example, Barro and Sala-I-Martin (1991) and Rey and Montouri (1999)], Canadian provinces [Coulombe and Lee (1995)], Columbian departments [Cardenas and Ponton (1995)], Mexican states [Mallick and

Carayannis (1994)], British counties and sub-regions [Chatterji and Dewhurst (1996) and Henley (2006], and European regions [Corrado, Martin and Weeks (2004)]. As outlined presently, this study broadens the scope of the existing literature in a number of ways. This is achieved through considering case-studies and datasets which have only recently come to light, by analysing the empirics of convergence in terms of the sectoral composition of regions, by addressing outstanding problems such as commuter flows and spatial dispersion, and by utilising a wide spectrum of complementary econometric techniques.

In the first paper, entitled *"Divergence, Convergence, or Something In-between? Sectoral Trends and British Regional Economic Growth"* an analysis of British regional economic development is undertaken which focuses on NUTS 3 Gross Value Added per capita data spanning from 1995-2004 for the primary, secondary, and services sectors. The aim of this paper is to look beneath the surface of aggregate British convergence-divergence trends. A range of techniques well-known to those familiar with the existing economic growth literature such as cross-sectional "growth equations" for absolute and conditional convergence are employed, as well as newly emerging nonparametric techniques. Distinguishing aggregate British GVA per capita from its secondary and services components is of course intuitively appealing: if one were to find inconclusive evidence of aggregate GDP per capita convergence-divergence over time it may well be the case that this is concealing the off-setting effects of, say, strong services sector divergence and manufacturing convergence or vice versa. While the time-span (1995-2004) considered in this paper is dictated by data availability, this decade is nonetheless of great importance as it represents a period of great change in the secondary and services sectors, most notably the move towards outsourcing of manufacturing capabilities and the absorption by the services industry of phenomenal technological advances. The surge in services sector output, accompanied by a falling off of secondary output, identified in the paper over the 1995-2004 period justifies a more disaggregated approach to the convergence/divergence debate.

The second paper, *"Sectoral Trends and British Regional Economic Growth – A Spatial Econometric Perspective"* delves further into the issue of convergence (or lack of) across British NUTS 3 regions. This paper addresses three problems that commonly arise in empirical studies of regional growth: (i) the impact of commuter flows that inevitably permeate highly disaggregated regional data; (ii) the inconsistencies that arise between administrative districts, such as NUTS 3 sub-regions, and actual zones of economic activity. The boundaries of administrative districts are often influenced by tradition, local custom, or other arbitrary reason and may render these districts less than ideal for economic analysis; (iii) identifying and accurately controlling for spatial dependence between neighbouring geographic units. These issues are addressed by constructing a set of functional economic regions for Britain, where the 128 NUTS 3 regions are aggregated together using a method based on commuter flow data. These functional economic areas provide a means for checking the robustness of results emanating from the econometric analysis carried out on the NUTS 3 level data. This paper also strives to identify the drivers of British regional growth over the 1995-2004 period.

This is achieved by compiling a dataset which includes NUTS 3 level data on secondary level education, primary school parent-teacher ratios, capital investment, the number of enterprises, proxies for labour market conditions, as well as dummy variables capturing regional effects. This expanded dataset is incorporated into a cross-sectional regression testing conditional convergence, which has been augmented to control for spatial dependence.

Post-unification Germany is the region of interest in the third paper, which is entitled *"Drifting Together Or Falling Apart? The empirics of Regional Economic Growth in Post-Unification Germany"*. The objective of this paper is to address the question of convergence across German districts in the first decade after German unification by drawing out and emphasising some stylised facts of regional per capita income dynamics. This is accomplished by employing non-parametric techniques which focus on the evolution of the entire cross-sectional income distribution. In particular, we follow a distributional approach to convergence based on kernel density estimation and implement a number of tests to establish the statistical significance of our findings. Establishing a detailed, disaggregated picture of the regional growth process is particularly important in the face of the simplistic east-west distinction that has often been made in the discussion of regional development in post-unification Germany. Such a simplification is only correct superficially. On the surface it appears as if such a distinction exists, but in reality the situation is different. As illustrated in this paper, since German unification several prosperous counties and cities have emerged in eastern Germany and therefore the two belt hypothesis is inadequate as a guide for regional economic policies.

The final paper *"Economic Growth Across Space and Time: Subprovincial Evidence from Mainland China"* considers the persistent differences in economic performance across Chinese regions. Despite the remarkable economic growth experienced in China over the last two decades, the disparities in China's regional development are startling. Urban and rural standards of living continue to be poles apart. While urban districts reap the benefit of investment inflows and preferential government policies, rural prefectures and townships still struggle to get to grips with basic healthcare and education provision. Discussion of Chinese regional disparities has often been framed in terms of a "three-belt hypothesis" which focuses on differences between the eastern, central, and western regions. In this paper a new highly disaggregated county and city-level dataset is introduced which spans the entirety of mainland China and provides a detailed view of Chinese regional growth over the 1997-2005 period. Non-parametric kernel density estimation is employed to establish the cross-sectional GDP per capita distribution, and the distributional dynamics are investigated using the probability matrix technique and the associated stochastic kernel estimator. A set of explanatory variables is then introduced and a number of regression estimators are utilized to test for conditional β-convergence and to pinpoint influential factors for economic growth across counties and cities.

References:

Aghion P., and P. Howitt, (1992) A Model of Growth Through Creative Destruction , *Econometrica* 60 2 323-351.

Barro R. and X. Sala-I-Martin (1991) Convergence across states and regions, Brookings Papers in Economic Activity 1.

Barro R. and X. Sala-I-Martin, (1992) Convergence, *Journal of Political Economy* 100 223-251.

Barro R. and X. Sala-I-Martin, (1995) *Economic Growth*, McGraw Hill.

Baumol WJ (1986) Productivity Growth, Convergence, and Welfare: What the Long-run Data Show, *American Economic Review* 76 1072-1085.

Cardenas M. and Ponton A. (1995) Growth and convergence in Colombia: 1950-1990 *Journal of Development Economics* 47 5-37

Cass, D. (1965) Optimum Growth in an Aggregative Model of Capital Accumulation", *Review of Economic Studies* 32 233-240.

Chatterji, M. and J. H. L Dewhurst (1996) Convergence Clubs and Relative Economics Performance in Great Britain: 1977–1991, *Regional Studies* 30 31–40.

Corrado L., Martin R., and M. Weeks, (2004) Identifying And Interpreting Convergence Clusters Across Europe, Royal Economic Society Annual Conference 145, *Royal Economic Society*.

Coulombe S. and F. C Lee. (1995) Convergence across Canadian provinces, 1969 to 1991, *Canadian Journal of Economics* 28 886-98

Domar, E.D. (1946) Capital Expansion, Rate of Growth, and Employment, *Econometrica* 14 137-147.

Harrod, R. F. (1939) An Essay in Dynamic Theory, *The Economic Journal* 49 14-33.

Henley, A. (2005) On Regional Growth Convergence in Great Britain, *Regional Studies* 39 1245-1260

Koopmans, T. C. (1965) On the Concept of Optimal Growth, in *The Econometric Approach to Development Planning,* Amsterdam, North Holland.

Lucas, R. E. (1988) On the Mechanics of Development Planning, *Journal of Monetary Economics* 22 1 3-42.

Mankiw G, Romer D. and D. Weil (1992) A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics* 107 407-437.

Mallick R. and E. Carayannis (1994) Regional economic convergence in Mexico. *Growth and Change* 25 325-34.

Ramsey, F. (1928) A Mathematical Theory of Saving, *Economic Journal* 38 543-559.

Rebelo, S. (1991) Long-Run Policy Analysis and Long-Run Growth, *Journal of Political Economy* 99 3 500-521.

Rey S., and B. Montouri (1999) U.S. Regional Income Convergence: A Spatial Econometric Perspective, *Regional Studies* 33 143-156

Romer, P. M. (1986) Increasing Returns and Long-Run Growth, *Journal of Political Economy* 94 5 1002-1037.

Solow, R. (1956) A Contribution to the Theory of Economic Growth, *Quarterly Journal of Economics* 70 1 65-94.

Swan, T. W. (1956) Economic Growth and Capital Accumulation, *Economic Record* 32 334-361.

Divergence, Convergence, or Something In-between? Sectoral Trends and British Regional Economic Growth

Declan Curran

1. Introduction

In recent years the dispersion of economic growth across all UK regions has emerged as a prime objective of British policymakers. This newfound prominence afforded to the "levelling-up" of regional growth rates has been evident in a slew of documents and pronouncements emanating from the UK government and appears to pinpoint the convergence of regional economic growth rates as a necessary step in counteracting existing agglomeration tendencies and enhancing productivity across the UK as a whole.[1] Regional disparities have been synonymous with modern day British economic development - manifesting themselves in the ubiquitous "north-south divide", the bane of politicians and policymakers alike. In 2005 the gross value added (GVA) per head of population for the UK was £17,700, with London had the highest regional GVA per head of population (£24,100), and South East following with £20,400.[2] The East of England (£18,900) was the only other region to have a GVA per head of population higher than the national average.[3] Wales had the lowest GVA per head of population at £13,800.[4] Despite this dominance of the southern regions, the regeneration of many of the regional population centres, coupled with the increasing tendency of services providers to look for low cost alternatives to London, has in recent times opened the door to peripheral regions keen to refocus their commercial energies. In 2005 the North East enjoyed, along with the East Midlands and London, the strongest GVA per head growth (4.4 per cent), while the lowest growth rate (3.5 per cent) was experienced in the South East. [5]

This analysis of British regional economic development focuses on NUTS 3 (sub-regional) real GVA per capita data spanning from 1995-2004 not just for aggregate British GVA per capita, but also for

[1] See the publications HM Treasury (2001,2002), Department of Trade and Industry (2003, 2004), ODPM (2003). As noted by Monastiriotis (2006), the British Treasury's 2002 Spending Review (HM Treasury 2002) explicitly asserts a regional policy target to "make sustainable improvements in the economic performance of all English regions and over the long term reduce the persistent in gap in growth rates between the regions." Boddy et al (2005) point to the HM Treasury (2001) publication which states that "the government's central economic objective is to achieve high and stable levels of growth and employment", and goes on to argue that "real economic gain for the country as a whole will only come from a process of "levelling up"".

[2] Throughout this paper, the term "regions" denotes British NUTS 1 level disaggregation, "counties" denote British NUTS 2 disaggregation, and "sub-regions" denote British NUTS 3 level disaggregation. The term "regional economic growth" is used in a general sense to refer to the field of literature to which this paper belongs.

[3] GVA is defined as follows: Under European System of Accounts 95 (ESA95), the term GVA is used to denote estimates that were previously known as Gross Domestic Product (GDP) at basic prices. Under ESA95 the term GDP denotes GVA plus taxes (less subsidies) on products, i.e. at market prices.

[4] Data available from the Office of National Statistics (ONS) at http://www.statistics.gov.uk/CCI/nugget.asp?ID=420&Pos=&ColRank=1&Rank=374.

[5] ibid

secondary and services sectors in order to understand the driving force behind aggregate British convergence-divergence trends. Such a disaggregation is of course intuitively appealing: if one were to find inconclusive evidence of aggregate GDP per capita convergence-divergence over time it may well be the case that this is concealing the off-setting effects of, say, strong services sector divergence and manufacturing convergence or vice versa. While the time-span (1995-2004) considered in this paper is dictated by data availability, this decade is nonetheless one of singular importance as it is over this ten year period in particular that one would expect such developments as the move towards outsourcing of manufacturing capabilities and the absorption by the services industry of phenomenal technological advances to translate themselves into tangible regional economic trends: in 2004 primary, secondary (including construction), and services as defined above accounted for approximately 1%, 22% and 75% of British GVA, while the equivalent shares in 1995 were 2%, 30% and 66%, respectively.[6] This surge in services sector output, accompanied by a falling off of secondary output, justifies a more disaggregated approach to the convergence/divergence debate.

This paper is organised as follows: Section 2 provides a brief description of the data used in this study. A synopsis of British sub-regional real GDP growth over the period 1977-1995 is presented in Section 3, setting the scene for the more rigorous treatment of British sub-regional aggregate, secondary and services real GVA per capita convergence/divergence trends over the 1995-2004 period that is undertaken in Section 4 using non-parametric techniques. The distributional dynamics of real GVA per capita are then considered in Section 5 with the aid of transition probability matrices and stochastic kernel estimates. Section 6 provides parametric growth equation techniques, which serve as a useful foil with which to compare the finding of the previous section. Section 7 concludes and suggests some possible extensions of this research effort.

2. Data sources

While it is British regional economic growth over the decade 1995-2004 that this paper focuses on, a brief treatment of economic growth over the 1977-1995 period is provided, by way of background.[7] National Accounts GDP per capita data at current prices for the 62 British counties over the period 1977-1995 is available from UK Office of National Statistics. These 62 British counties correspond to existing NUTS 2 level of disaggregation. No NUTS 3 level GDP data is available for the 1977-1995 period. As no regional GDP deflator is available for the 1977-1995 period, this GDP per capita data is deflated using the national GDP deflator (in 2002 UK£). Turning to the 1995-2004 period, unadjusted (constrained to headline NUTS2) total gross value added (GVA) by NUTS3 area at current basic prices for the years 1995 to 2004 is available from the Office of National Statistics (www.statistics.gov.uk), as well as being disaggregated for 1) agriculture, hunting and forestry 2)

[6] Calculations based on National Accounts GVA data available from Office of National Statistics, as discussed in Section 2.
[7] For the purposes of this study, only Great Britain is considered, i.e. Northern Ireland is not included.

Industry, including energy and construction and 3) service activities, including Financial Intermediation Services Indirectly Measured (FISIM). These three categories are henceforth referred to as "primary", "secondary", and "services", respectively. Estimates of workplace based GVA allocate income to the region in which commuters work. Per capita estimates can then be constructed using NUTS 3 level population data available from Nomis (www.nomisweb.co.uk). Unfortunately, regional deflators such as the Retail Price Index (RPI) are only available for UK for the years 2000, 2003, and 2004, and the methodology for this index is still in a formative stage. One could just use the yearly national deflator for each NUTS 3 region. However, this is unsatisfactory as it makes no allowance whatsoever for regional price differences – particularly British secondary, services, and aggregate GVA per capita exhibit clear regional trends, as illustrated in Figures 2-4. In this study, regional deflators for each year have been constructed by weighting the 1995-99 national RPI figure by the 2000 regional RPI weights. Similarly for 2001-2002 regional RPI the 2003 regional RPI figures are used as weights. The basket used to calculate the RPI figures include both consumer goods and services such as household services, personal services, and leisure services.[8]

3. British regional growth 1977-1995

The illustration provided in Figure 1, below, of the UK business cycle and output gap over the 1977-2005 period offers a stark reminder of the economic trials and tribulations of the last three decades. Most notable are the recession of 1979-1980, over which the newly-formed Thatcher government presided; the continued economic malaise in the early 1980s during which unemployment soared and privatisation of state-controlled firms was the order of the day; the economic recovery of the late 1980s; the recession of the early 1990s which culminated in Britain's exit from the ERM in September 1992; and, in more recent times, the longest period of sustained economic growth in modern British history.[9] Needless to say, the business cycle movements experienced over the 1977-2005 period were formative events in the development of disparities in regional economic performance, particularly in perpetuating the north-south divide, which has featured prominently in British regional growth literature. It is the 1995-2004 period that will be scrutinized in detail in Sections 3-5. It can be seen from Figure 1 that the business cycle stance in 1995 differs from that in 2004, with 1995 showing no output gap and 2004 representing a business cycle peak.

Research into British regional growth patterns over the 1977-1995 period has identified a number of prominent features: Chatterji and Dewhurst (1996) and Bishop and Gripaios (2004) both conclude that Regional GDP per capita data yields no evidence of convergence exists over this time period, though the former point to some sub-periods that exhibit convergence (in periods where the economy as a

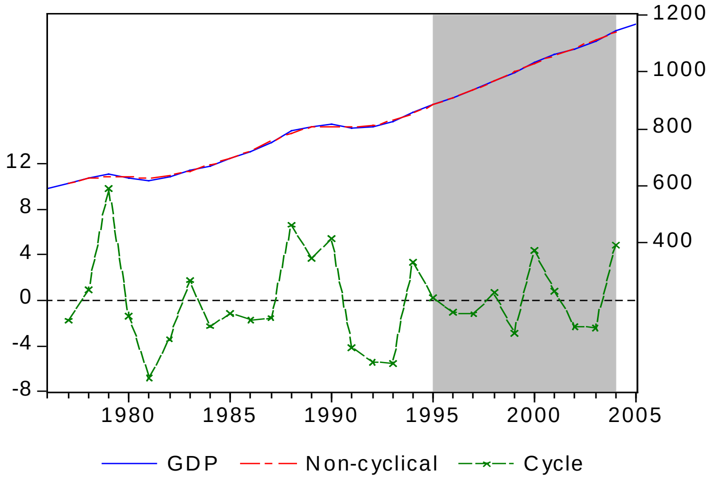
---

[8] Fure further details of the composition of the RPI series, see the ONS publication *Economic Trends 615*, February 2005.

[9] A detailed treatment of modern day British economic development can be found in Floud and Johnson (2004) "The Cambridge Economic History of Modern Britain".

whole was experiencing slow growth); Dewhurst (1998) finds evidence of the fore-mentioned "north-south divide"; Duranton and Monastiriotis (2002) and Monastiriotis (2006), using data from the New Earnings Survey, point to widening aggregate disparities throughout the 1980s and 1990s but a convergence in regional reward structures.[10] Henley (2006) has undertaken a spatial econometric analysis of NUT 3 level aggregate GVA data for the 1995-2001 period and concludes that British NUTS 3 sub-regions experienced divergence over this time period.

Figure 1: UK real GDP and Output Gap, 1977-2005



Note: The shaded region indicates the 1995-2004 period, the time period under consideration in Sections 3-5; GDP data (seasonally adjusted) is given in 2002 UK£ billion (right hand scale). The output gap is calculated using the Baxter and King (1999) time-varying asymmetric band-pass filter, with "Cycle" denoting the filtered series. The band pass filter isolates the cyclical component of GDP that lies between 2 and 8 years. The filtered series is measured in percent (left hand scale). A dashed line is included at zero on the left-hand scale for ease of interpretation.

The choice of dataset used in pre-1995 research has come under scrutiny recently in the aftermath of Cameron and Muellbauer (2000), who point out a discrepancy in the compilation of Regional GDP data over the 1977-1995 period when compared with New Earnings Survey data. In the UK National Accounts, estimates of income from employment and of wages and salaries are constructed from a one percent sample of tax and social security records (combined with estimated earnings of those below the relevant tax bands). Cameron and Muellbauer (2000) note, however, that between the 1982/83 tax year and 1989/90 tax years the British Inland Revenue were unsuccessful in allocating approximately 12% of UK tax records in their 1% sample to a particular region, with the under-allocation falling mostly on the South East region (information on the percentage unallocated, if any, for previous years is unavailable). By 1995 this source of discrepancy had been eliminated. With this caveat in mind, Table 1 provides estimation results of a Least Trimmed Squares (LTS) linear regression of average regional GDP per capita growth ($\Delta\%L_{1977\text{-}1995}$) of the 62 UK counties on initial, 1977, log GDP

[10] The New Earnings Survey (NES) is based on a 1 percent sample of employees in employment, information on earnings and hours is obtained in confidence from employers. has been conducted during April each year since 1971 by The Department of Enterprise, Trade and Investment.

($lnGDP_{1977}$) – the standard test for unconditional β-convergence, with a positive coefficient on initial GDP pointing to concentration (divergence) and a negative coefficient indicating de-concentration (convergence).[11] The Least Trimmed Squares (LTS) regression method is used, rather than Ordinary Least Squares (OLS), in order to control for the effects of possible outliers in the data. The LTS technique, introduced by Leroy and Rousseeuw (1987), uses a re-sampling algorithm to locate observations which are not outliers and then uses this "good data" to pinpoint the presence of outliers. Observations whose LTS standardized residuals are greater than 2.5 in magnitude are then dropped from the sample and an OLS regression is run using the remaining data.

Table 1: Total real GDP per capita growth on initial log GDP per capita, 1977-1995

Dependent variable: Average Growth of Total GDP per Capita 1977-1995		
	1977-1995	**1977-1995**
constant	0.118***	0.118***
	(0.00)	(0.00)
$lnGDP_{1977}$	-0.003	-0.004
	(0.49)	(0.44)
NS Dummy		0.03**
		(0.01)
R^2	0.01	0.12
Number of Obs	61	61

Note: P-values in brackets. GDP per capita deflated using price deflator with base year = 2002. Warwickshire is the only county whose LTS standardized residual exceeds 2.5 in magnitude.

It is apparent from Table 1 that there is no support for the hypotheses of absolute β convergence over the 1977-1995 period. This is consistent with the findings of Bishop and Gripaios (2004) who find no signs of convergence over the 1977-1995 period regardless of whether one uses National Accounts or New Earnings Survey data. Also, as per Bishop and Gripaios (2004), a similar regression for conditional β convergence, which includes a North-South *(NS)* dummy, does not point to convergence but yields a statistically significant dummy variable coefficient, indicative of the north-south divide.[12]

4. Trends in British regional GVA per capita over the 1995 – 2004 period

This section focuses on NUTS 3 level sub-regional GVA per capita data spanning from 1995-2004, not just for aggregate British GVA per capita, but also for secondary and services sectors, in order to gain an understanding of the driving force behind aggregate British convergence-divergence trends. Such disaggregation is intuitively appealing: if one were to find inconclusive evidence of aggregate

[11] OLS estimation yields results in keeping with the LTS results and are available from the author on request. See Barro and Sala-I-Martin (1995) for a full treatment of absolute and conditional convergence concepts.
[12] As per Monastiriotis (2006), "South" refers to South East, South West, Greater London, East of England, East Midlands. "North" consists of the remaining regions of England, as well as Wales and Scotland.

GDP per capita convergence-divergence over time it may be the case that this is concealing off-setting effects of, say, services sector divergence and manufacturing convergence or vice versa.[13]

In 2004 primary, secondary (including construction), and services as defined above accounted for approximately 1%, 22% and 75% of British GVA, while the equivalent shares in 1995 were 2%, 30% and 66%, respectively. Clearly, over the 1995-2004 the services sector has consolidated its dominance in the creation of British GVA, while the contribution of the secondary industry appears to be on the wane. This serves a further justification for adopting a disaggregated approach in analysing British real GVA per capita trends over time. In order to present as detailed a picture as possible of British sub-regional real GVA per capita trends over the 1995-2004, a number of complementary techniques are employed. An obvious starting point in shedding light on these regional trends at the sectoral level is the inclusion of colour-coded maps comparing the spatial dispersion of sub-regional real GVA per capita across the 128 British NUTS 3 regions at the beginning and end of the sample period. Kernel density estimation is then employed as a useful next step as it enables us to identify the stylised facts of the distribution of sub-regional real GVA per capita over time in a non-parametric manner that, as the old adage goes, "allows the data to speak for itself". This nonparametric analysis is complemented in Section 5 with techniques well-known to those familiar with the existing economic growth literature: cross-sectional "growth equations" for absolute convergence.

4.1 The spatial dispersion of real GVA per capita

In order to geographically illustrate the spatial disparity of regional real GVA per capita, a set of maps have been compiled, Figures 2- 4, of the real GVA per capita across British NUTS 3 sub-regions. Each map is colour coded, with the light shading denoting 0-100% of median real GVA per capita, medium shading denoting 100-125%, and dark shading denoting over 125% of median real GVA per capita. Each sub-region is shown relative to the median rather than the mean to mitigate the impact of outliers such as the services GVA of Inner London West. Figure 2 presents aggregate real GVA per capita for 1995 and 2004. Salient features include the high GVA per capita in greater London, Manchester-Liverpool, Edinburgh, Glasgow, and Aberdeen (near the North Sea oil fields); a clear expansion of the greater London high-GVA area over the period in question; the noticeable improvement of the Midlands; and the apparent falling back of Northern England and Scotland. One might wonder whether these impressions are reflected in the development of the secondary and services sectors over the 1995-2004 period. As illustrated in Figure 3, however, the secondary industry presents a more mixed picture: the North of England appears to fall back, relatively speaking; a belt of increased GVA per capita is apparent in the Midlands, while the South West and South East exhibit some shuffling of regions between the three categories, but no clear pattern. The services sector (Figure 4) highlights the

[13] This approach stems from the work of Desmet and Fafchamps (2004), who, using nonparametric methods, have recently examined the spatial distribution of employment, as opposed to GDP, across US counties between 1972 and 1992. Their results point to an increase in total employment concentration, with this aggregate dynamic driven by services sector divergence outweighing opposing primary and secondary influences

clear expansion of the high-GVA greater London area, increases in Liverpool-Manchester, but continued sluggishness in Northern England and Scotland. In all it would appear that it is the services industry which drives the expansion of the southern high GVA in the aggregate map. While the secondary sector does appear to be the more dispersed in terms of the highest GVA category; this trend seems to be eclipsed in the aggregate GVA map by the strong services performance.

Figure 2: Aggregate real GVA per capita 1995 (left) and 2004 (right)

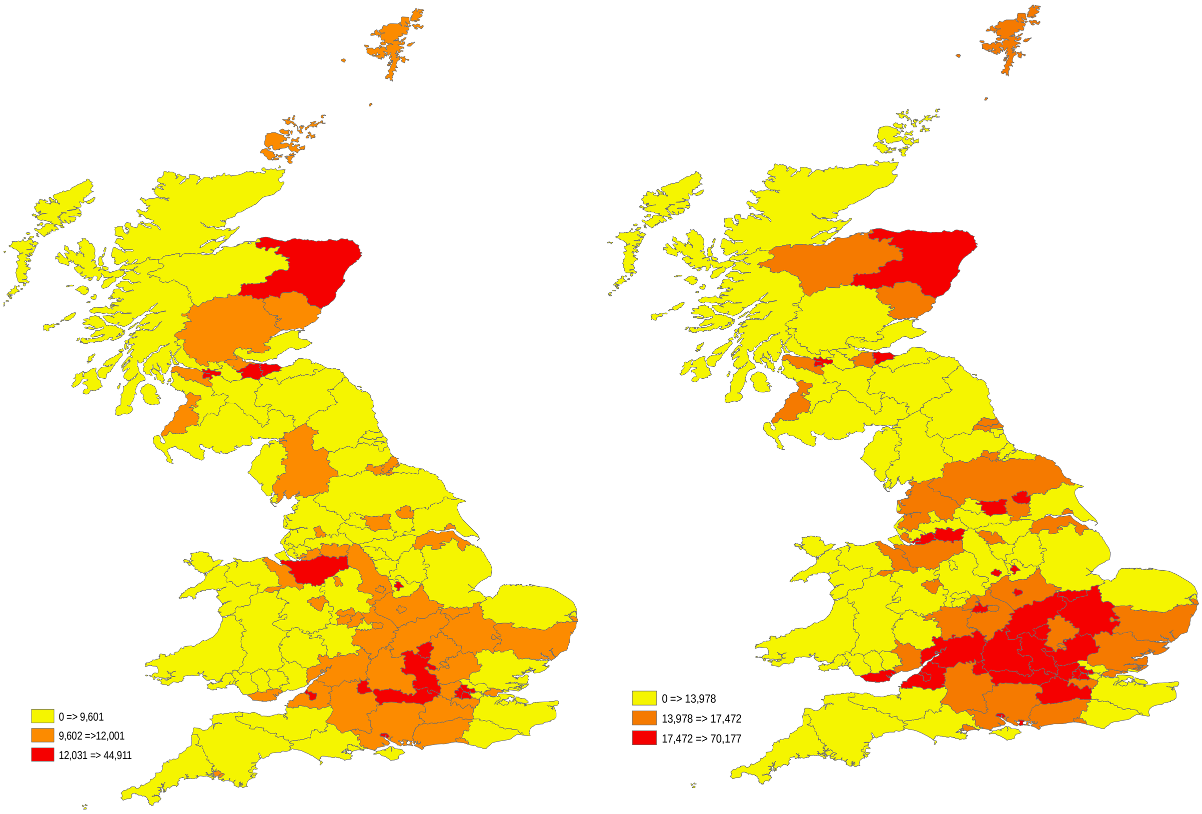


Figure 3: Secondary industry real GVA per capita 1995 (left) and 2004 (right)

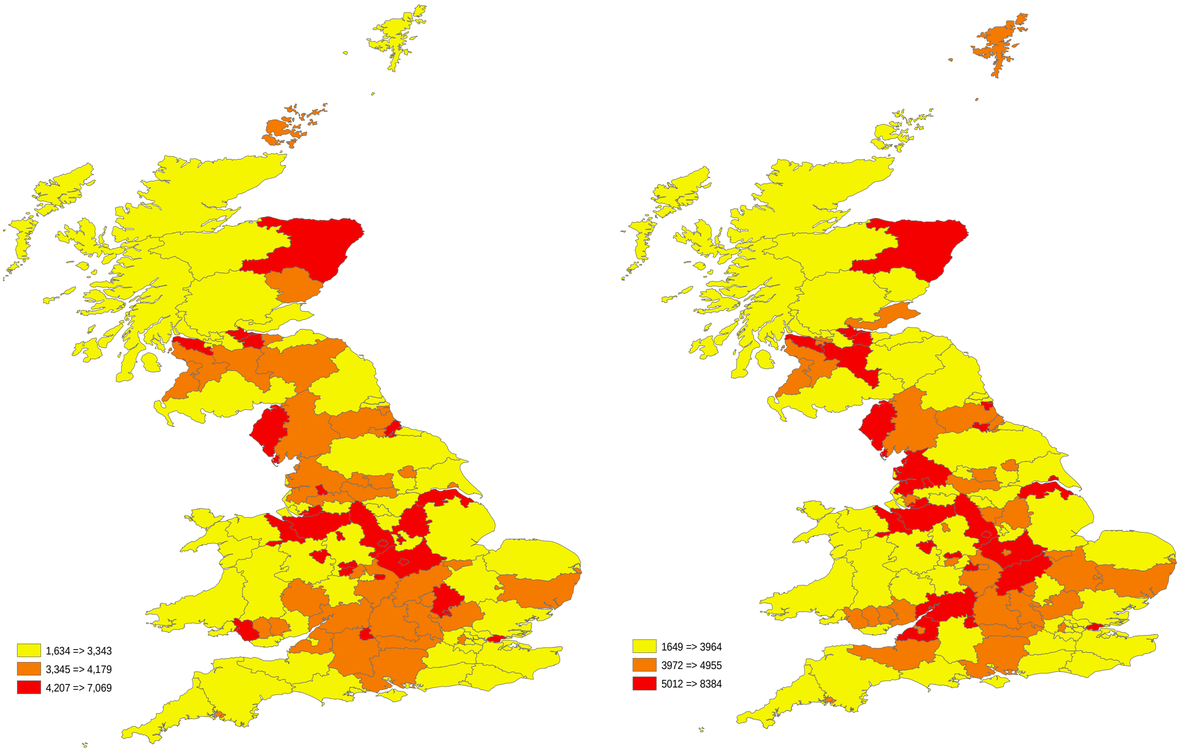
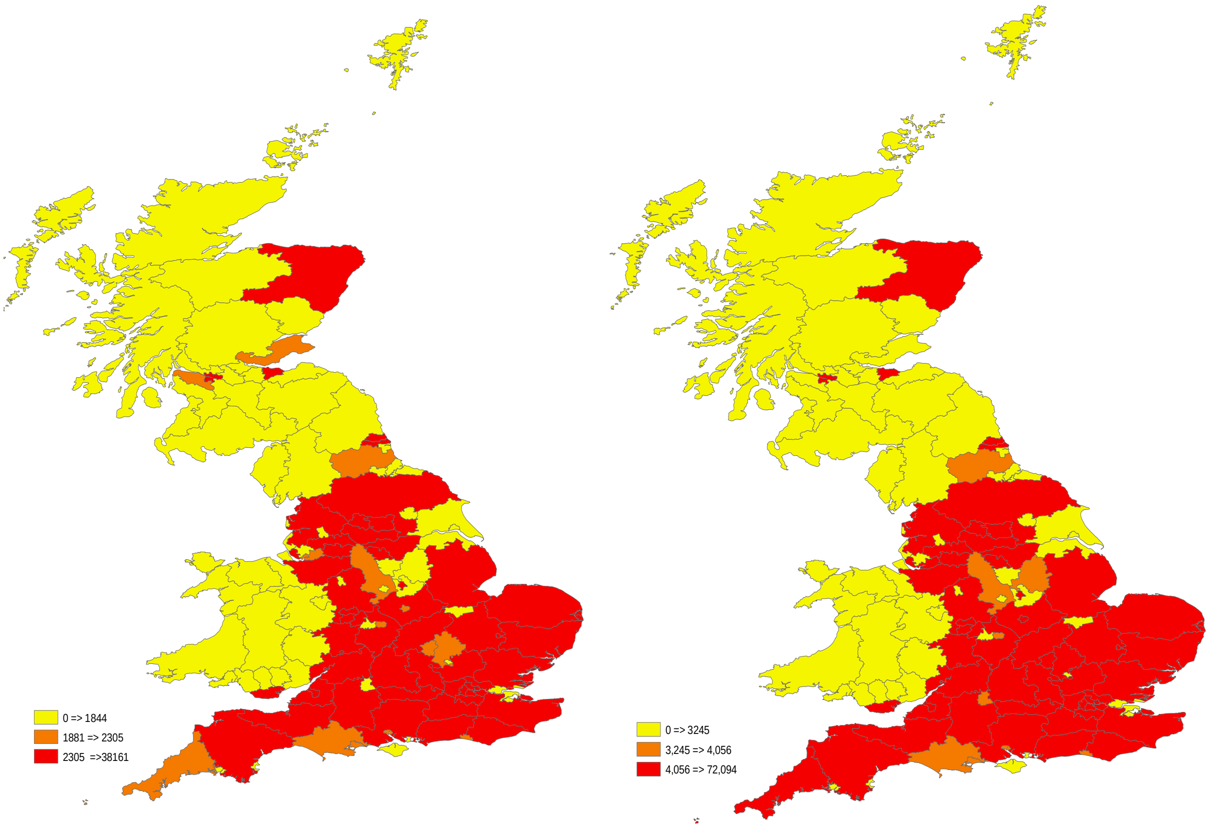


Figure 4: Services real GVA per capita 1995 (left) and 2004 (right)



4.2 Kernel density estimation and real GVA per capita

Nonparametric techniques, such as the Kernel density estimator, can also reveal interesting features of the data and therefore help to capture the stylised facts that need explanation. The kernel estimator for the density function $f(x)$ at point x is

$$(1) \qquad \hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right)$$

where $x = x_1, x_2, \ldots, x_n$, is an independent and identically distributed sample of random variables from a probability density $f(x)$ and $K(\cdot)$ is the standard normal kernel with window width h. The window width essentially controls the degree to which the data are smoothed to produce the kernel estimate. The larger the value of h, the smoother the kernel distribution. A crucial issue is the selection of this smoothing parameter. Here, the two-stage direct plug-in bandwidth selection method of Sheather and Jones (1991) is employed, which has been shown to perform quite well for many density types by Park and Turlach (1992) and Wand and Jones (1995).[14] The distributions have been fitted to the logarithm of real GVA per capita. In Figures 5, 6 and 7 are plotted the kernel density estimations for the (log) *GDP* from 1995 to 2004 obtained using the abovementioned bandwidth selection method and by transforming the income variable to the original scale.[15]

[14] Given the crucial role played by the bandwidth selection method, it is important to assess the performance of alternative bandwidth selectors. When the Silverman (1986) rule of thumb bandwidth selector has been used for the above kernel density estimation, similar trends are exhibited by the distributions. Detailed results are available from the author on request.

[15] In this paper, the visual impressions provided by the Kernel density estimators receive a more rigorous appraisal when compared to the findings of the absolute convergence regression analysis of Section 5. There are, on the other hand, a number of statistical techniques available for testing whether the estimated distributions are unimodal or multimodal. Among them are the Timm (2002) bimodality index, the Silverman (1981, 1986) multimodality test, and the nonparametric test of density time invariance using the test statistic of Li (1996). See Colavecchio, Curran, and Funke (2006) for a detailed discussion of these multimodality tests.

Figure 5: Aggregate GVA per capita Kernel Density Estimates 1995-2004

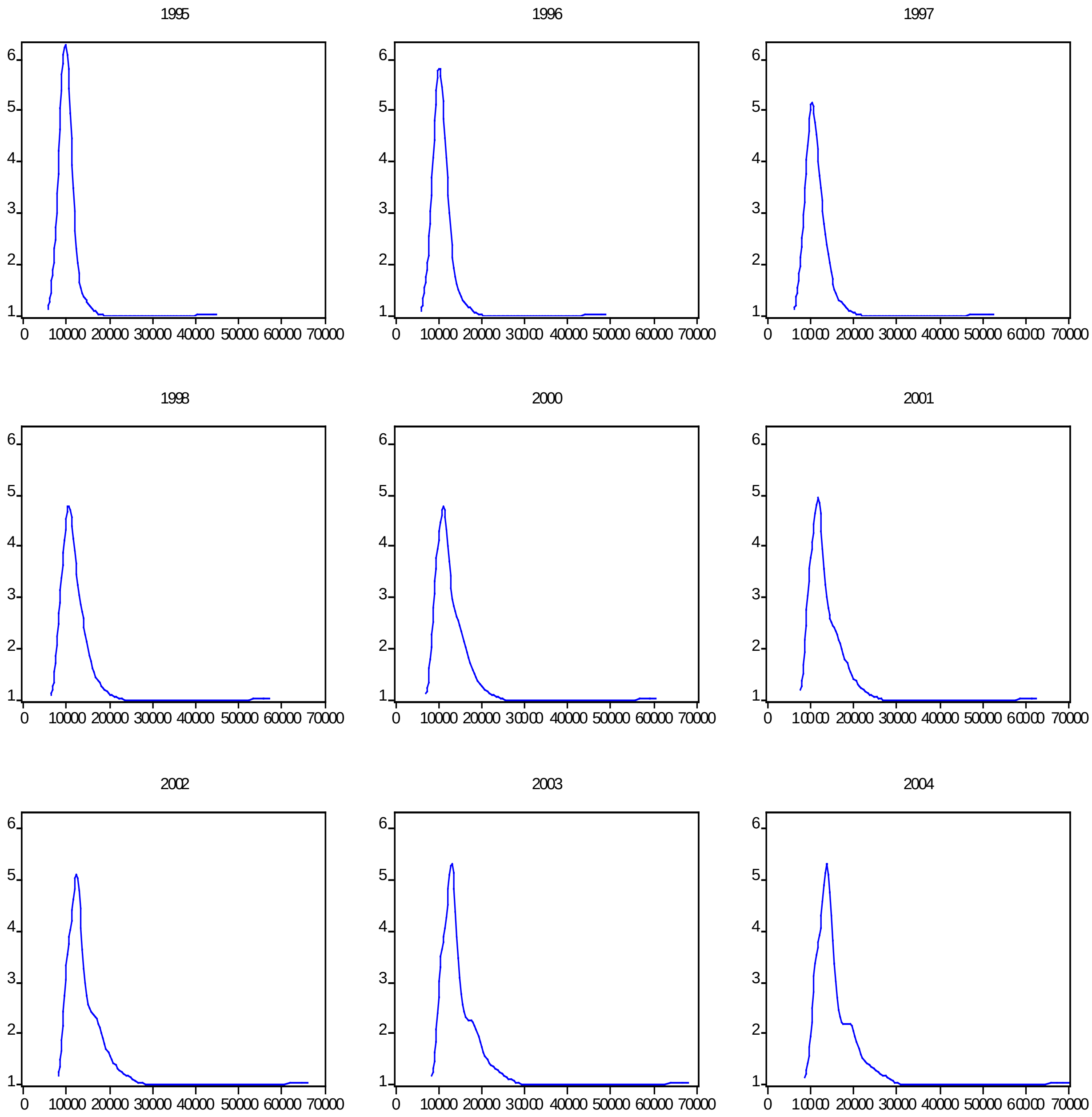


Figure 6: Secondary GVA per capita Kernel Density Estimates 1995-2004

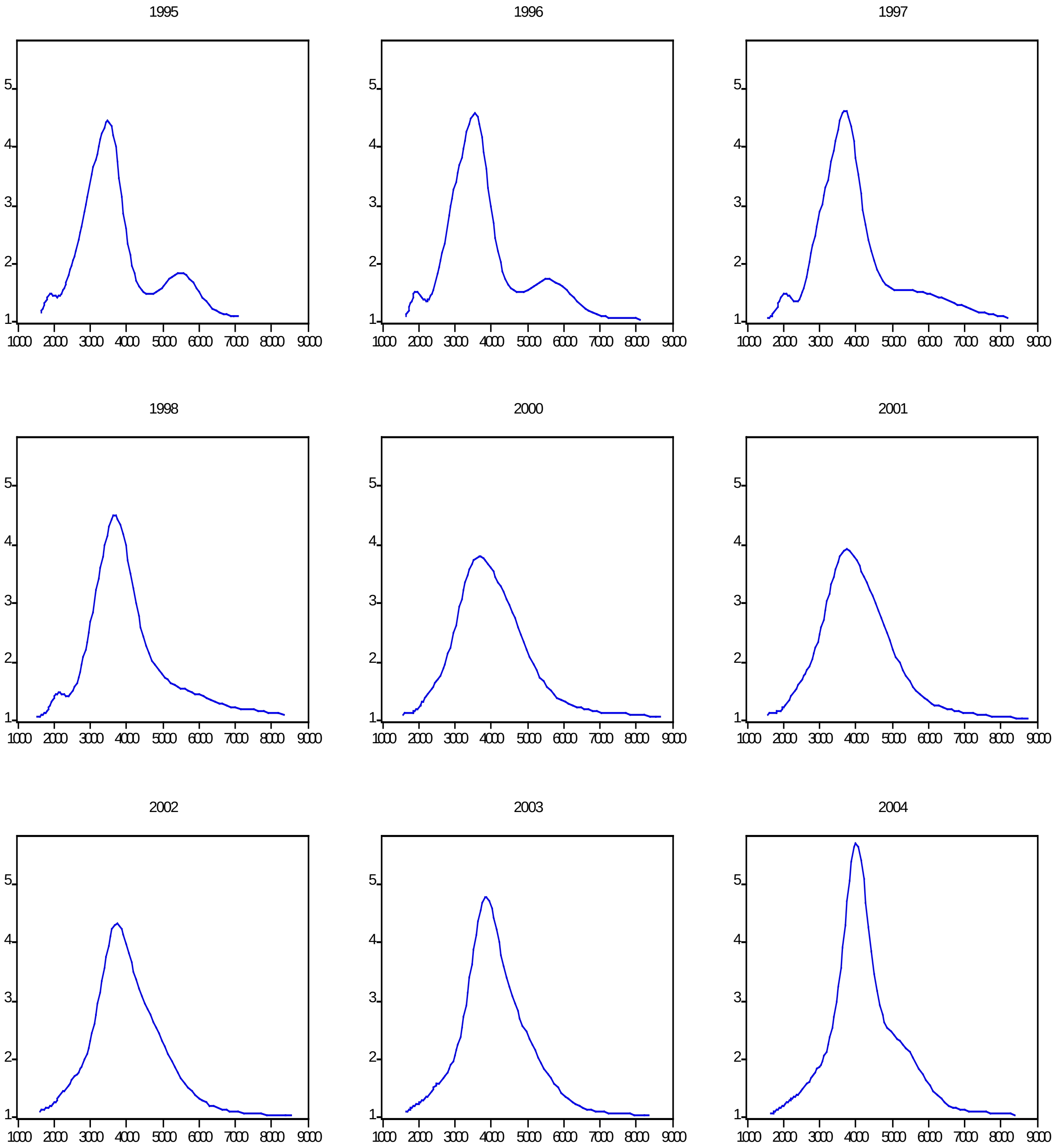
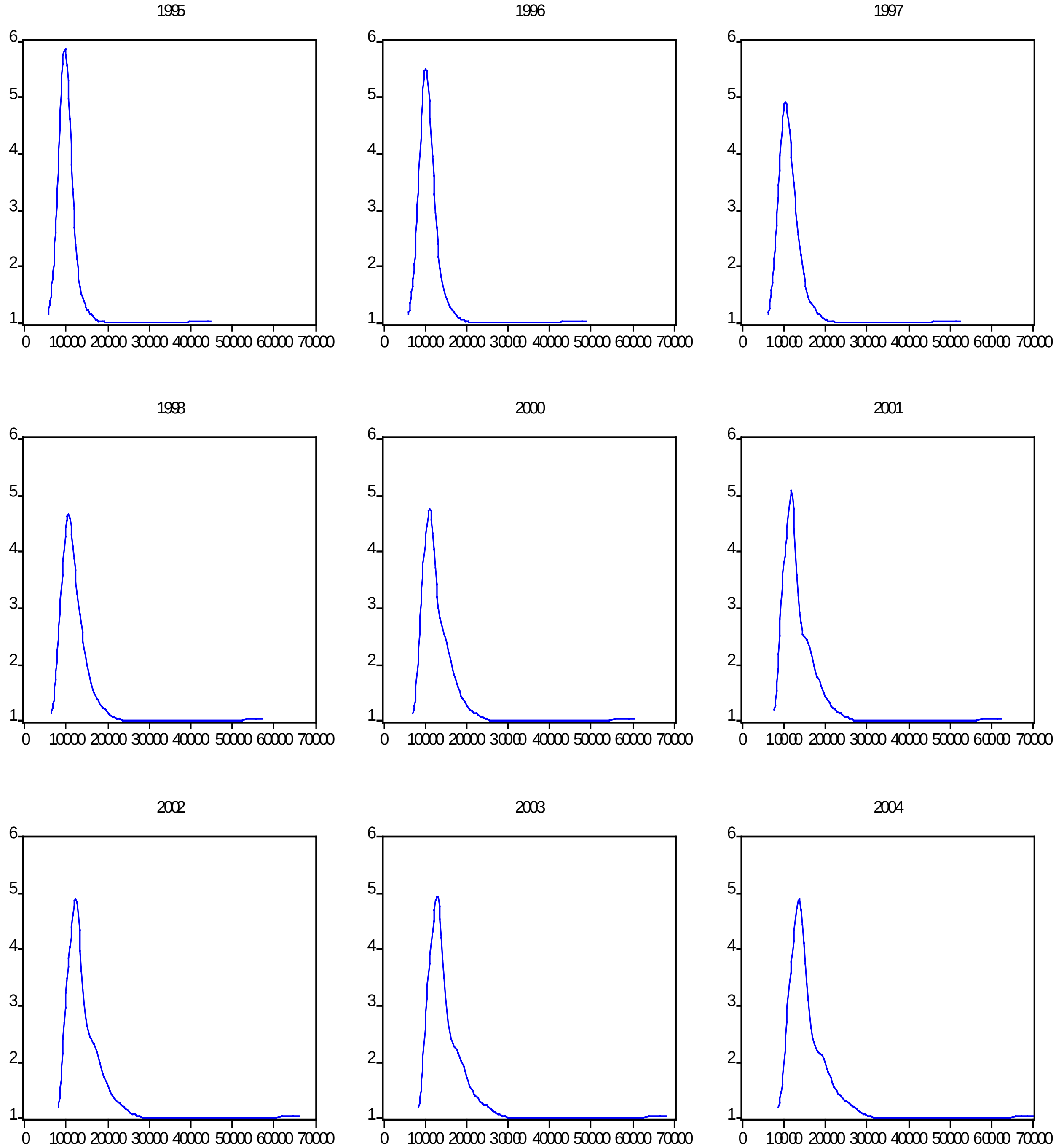


Figure 7: Services GVA per Capita Kernel Density Estimates 1995-2004



In the Kernel density estimation context, a convergence process occurs if, for instance, a bimodal density is detected at the beginning of the sample period and over time there is a tendency in the distribution to move towards unimodality. Alternatively, if there already is a unimodal distribution at the beginning of the time span in question, convergence occurs when the dispersion of this density and therefore per capita income declines over time. The unimodal distribution of Figure 5 reveals no discernable movement towards convergence in aggregate GVA per capita distribution over the 1995-2004 period. If anything, it hints at a greater dispersion of the unimodal density in 2003 and 2004.

This is in stark contrast with the distribution of secondary real GVA per capita (Figure 6) which appears to exhibit what may even be a bi-modal distribution in 1995, with a high GVA mode discernible, but converges quickly to a unimodal density as this high GVA mode recedes. This process culminates in a noticeably slender distribution in 2004, where sub-regions appear to be compressing towards the middle of the distribution. It is the kernel density distribution of the services sector that most resembles that of aggregate real GVA per capita. As with the aggregate distribution, the services sector (Figure 7) displays no tendency towards convergence but rather hints at a greater dispersion of the distribution in 2003 and 2004. The services sector also contains a noticeable outlier in the shape of the financial district of Inner London West, whose real GVA per capita in 2004 was almost seven times that of the median value. This services sector outlier is also visible in the aggregate GVA case (Figure 5).[16] In all, it would appear that despite the strong movement towards convergence in secondary sector real GVA per capita over the 1995-2004 period, it is the services sector that exerts the stronger influence of aggregate real GVA per capita during this period – consistent with the growth of the services sector's share of aggregate GVA over the 1995-2004 period.

Given the emergence of contrasting convergence and divergence trends for secondary and services real GVA per capita respectively, it is natural to wonder whether these sectors are converging and diverging to "good" or "bad" states. Descriptive statistics for these two sectors for 1995 and 2004 provide a clear indication of these states:

Table 2: Summary Statistics for Secondary and Services real GVA per capita, 1995 and 2004

Secondary Sector GVA per capita (2002 UK£)			**Services Sector GVA per capita (2002 UK£)**		
	1995	**2004**		**1995**	**2004**
Mean	3,517.29	4,031.723	Mean	6,422.84	11,261.36
Median	3,343.53	3,964.373	Median	5,828.70	9,708.08
Maximum	7,068.65	8,383.499	Maximum	41,398.86	64,654.04
Minimum	1,634.15	1,648.837	Minimum	3,050.08	5,766.21
Std. Dev.	1,162.03	1,168.167	Std. Dev.	3,574.93	6,023.20

The contrast between secondary and services sector GVA per capita developments over the 1995-2004 period is stark. The virtually unchanged mean, median, and standard deviation of secondary GVA per capita over the 10 year period, together with slight increases in the minimum and maximum GVA per capita figures suggest that any convergence experienced in the secondary sector has not been a buoyant one. Services GVA per capita, on the other hand, bears all the hallmarks of a sector on the move, with its mean and median showing marked increases over the 10 years and its widening standard deviation indicative of the absolute divergence hinted at by this section's kernel estimation analysis.

[16] In 2004 the highest NUTS 3 real GVA per capita was that of Inner London West with £64,654 (in 2002 UK£). Edinburgh resided in second place, with a real GVA per capita of £23,174. Median real GVA per capita was £9,708. In 1995 Inner London West enjoyed a real GVA per capita of £41,399 compared to Edinburgh's £12,870 and a median value of £5,828, indicating a slight narrowing of both the gap between Inner London West and Edinburgh and between Inner London West and the median over the 1995-2004 period.

Section 5. Distribution dynamics of British real GVA per capita

Having established the distribution of aggregate real GVA per capita and its constituent parts, the underlying process is further examined by considering the intra-distributional dynamics of secondary and services sector GVA per capita over the 1995-2004 period. This involves modelling directly the evolution of relative income distributions by constructing transition probability matrices that track changes over time in the relative position of districts within the distribution. This is an exercise that a number of authors have undertaken (see Quah, 1996a, 1996b, 1997, 2006). The modelling of distribution dynamics assumes that the density distribution ϕ_t has evolved in accordance with the following equation:

(2) $$\phi_{t+1} = M\,\phi_t,$$

where M is an operator that maps the transition between the income distributions for the periods t and t+1. Since the density distribution ϕ for the period t only depends on the density ϕ for the immediately previous period, this is a first-order Markov process.[17] In the estimates below it is assumed that the distribution ϕ has a finite number of states. For the Markov transition matrices it is assumed that the probability of variable s_t taking on a particular value j depends only on its past value s_{t-1} according to the first-order Markov chain

(3) $$P\{s_t = j \mid s_{t-1} = i\} = P_{ij},$$

where P_{ij} indicates the probability that state i will be followed by state j. As

(4) $$P_{i1} + P_{i2} + ... + P_{in} = 1$$

it is possible to construct what is known as the transition matrix

(5) $$P = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ P_{n1} & P_{n1} & \cdots & P_{nn} \end{bmatrix}$$

where line i and column j give the probability that state i will be followed by state j. In the following modelling approach, the probability P_{ij} measures the proportion of districts in regime i during the

[17] Equation (2) may be seen as analogous to a first-order autoregression in which point estimates are replaced by complete distributions.

previous period that migrate to regime j in the current period. According to Geweke et al. (1986), the maximum likelihood estimator for the transition probability $\hat{P}_{ij}$ is given by:

$$(6) \qquad \hat{P}_{ij} = \frac{\sum n_{ij}}{\sum n_i}$$

where $\sum n_{ij}$ is the number of districts that were in income category i in the previous period and have migrated to income category j in the current period, and $\sum n_i$ is the total of districts that were in income category i in the previous period. The main advantage of the transition matrix is that it allows one to summarise the random ups and down of regional fortunes in a handful of numbers.

The transition probability matrix in Tables 3 and 4 report transitions between the 1995 and 2004 distributions of the NUTS 3 level secondary and services sector real GVA per capita relative to the British median. The main diagonal of the matrix gives the proportion of districts that remained in the same range of the distribution throughout the period in question, while the off-diagonal probabilities are those associated with moving from one state to another. Tables 3 and 4 also provide information about n, the number of regions that begin their transitions in a given state. Furthermore, the classes that divide up the state space are provided.

Table 3: Transition probability matrix relative to the median secondary real GVA per capita

	GVA PER CAPITA 2004							
	n		*9*	*18*	*37*	*35*	*16*	*13*
GVA PER CAPITA 1995	*10*	[0-0.60]	0.60	0.40	0.00	0.00	0.00	0.00
	19	[0.60-0.80]	0.16	0.42	0.32	0.05	0.00	0.05
	35	[0.80-1.00]	0.00	0.11	0.63	0.26	0.00	0.00
	34	[1.00-1.20]	0.00	0.06	0.21	0.50	0.18	0.06
	8	[1.20- 1.40]	0.00	0.00	0.13	0.25	0.63	0.00
	22	[1.40- ∞]	0.00	0.00	0.05	0.27	0.23	0.45
			[0-0.60]	[0.60-0.80]	[0.80-1.00]	[1.00-1.20]	[1.20- 1.40]	[1.40- ∞]

The secondary sector transition probability matrix in Table 3 reveals a number of noteworthy behavioural patterns in the distribution of secondary sector real GVA per capita over time. It is clear from the probabilities that lie along the diagonal that some states are more susceptible to movement than others. In particular, regions in the second state (60-80% of the median) and those in the highest state (greater than 140% of the median) appear more likely to move than regions in the other states, as their probability of staying put is noticeably lower. The off-diagonal probabilities for the second state suggest that the movement is most likely a positive one towards the median, while the off-diagonal probability for the highest state points to reversal towards the median – similar to the convergence trends apparent in the previous section's kernel density estimation and absolute convergence regressions. This is further borne out by the number of sub-regions residing in each state in 1995 and 2004: there appears to be shuffling of sub-regions within the three lower states of the secondary sector but no great change in the number of sub-regions residing in each state, as anticipated by the maps presented in Section 2, while there is a clear levelling out of the number of sub-regions in the second highest and highest states between 1995 and 2004.

Table 4: Transition probability matrix relative to the median services real GVA per capita

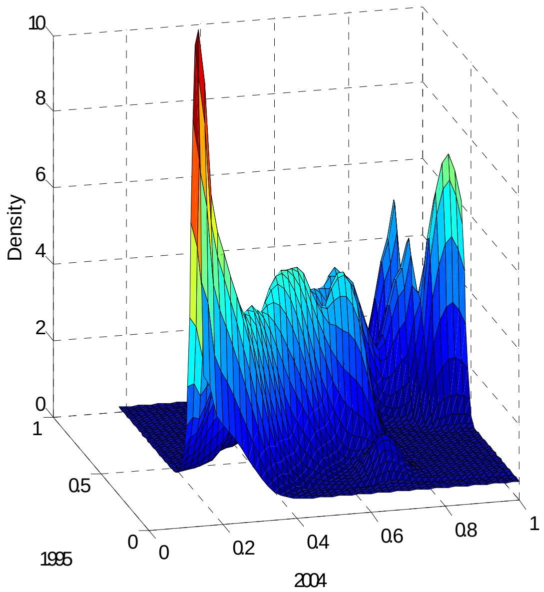
	GVA PER CAPITA 2004							
	n		*2*	*22*	*40*	*23*	*13*	*28*
GVA PER CAPITA 1995	*1*	[0-0.60]	1.00	0.00	0.00	0.00	0.00	0.00
	22	[0.60-0.80]	0.05	0.64	0.32	0.00	0.00	0.00
	41	[0.80-1.00]	0.00	0.20	0.66	0.12	0.02	0.00
	32	[1.00-1.20]	0.00	0.00	0.19	0.53	0.25	0.03
	13	[1.20- 1.40]	0.00	0.00	0.00	0.08	0.31	0.62
	19	[1.40- ∞]	0.00	0.00	0.00	0.00	0.00	1.00
			[0-0.60]	[0.60-0.80]	[0.80-1.00]	[1.00-1.20]	[1.20- 1.40]	[1.40- ∞]

In the services sector (Table 4), the most notable feature is the high probability of movement (62%) from the second highest state to the highest. Those sub-regions that enjoy relatively high services sector GVA per capita (say, over 120% of median service sector GVA per capita) at the beginning of the time period display a relatively high probability to either push forward to a higher state or, at least, maintain their status quo over the 1995-2004 period. The actual number of sub-regions enjoying services sector GVA per capita of over 140% of the median swells from 19 to 28 over the 1995 to 2004 period, as the sub-regions above the median level push forward. On the other hand both the probabilities and absolute number of sub-regions associated with the lower states indicate that sub-regions in these states have a high probability of retaining their relative positions but are less likely to move up to higher states. In all, the impression created by Tables 3 and 4 of a secondary sector compressing at the median level, coupled with the break-away group of high services GVA per capita sub-regions, is in keeping with that created by the colour-coded maps and kernel densities of the previous sections.

In Tables 3 and 4, the operator M has been constructed by assuming that the distribution ϕ_t has a finite number of states. This discrete modelling approach leads to the problem that the researcher has to determine the number of intervals and the limit values of each interval in an arbitrary and ad hoc way. Furthermore, the discretisation process may eliminate the property of Markovian dependence in the data, as Bulli (2001) has pointed out. The solution which addresses these shortcomings consists of carrying out a continuous analysis of transition, which avoids discretisation through the use of conditional densities that are estimated non-parametrically and referred to as stochastic kernels. A stochastic kernel amounts to a transition matrix with an infinite number of infinitely small ranges. The results from this tool are displayed as three-dimensional graphs in Figure 8 and a two-dimensional contour map in Figure 9.

Figure 8: Stochastic Kernel Estimates, 1995-2004

Secondary



Services

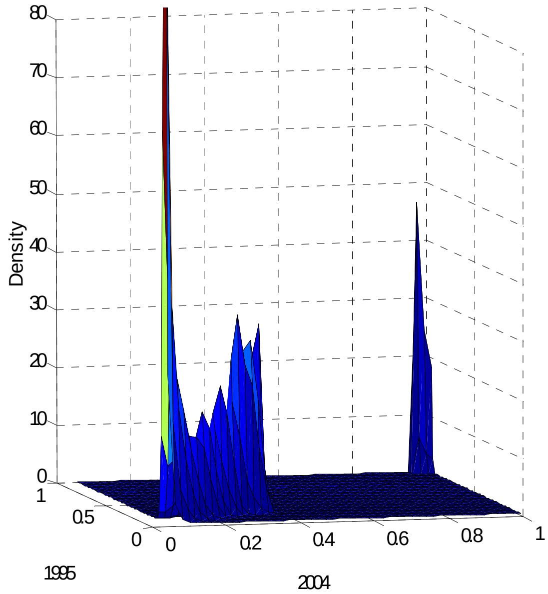
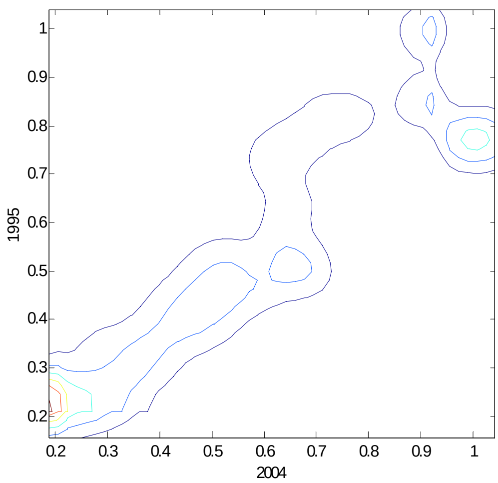
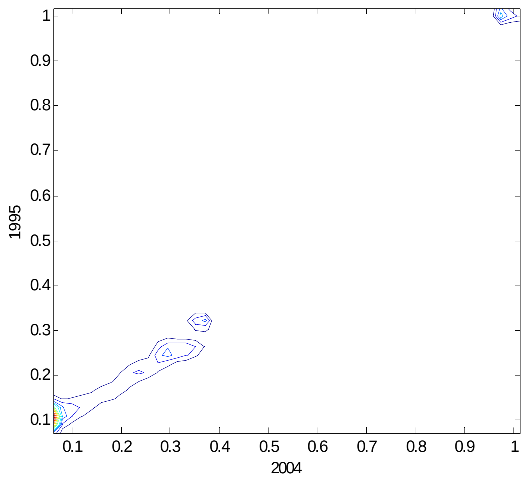


Figure 9: Stochastic Kernel Contours, 1995-2004

Secondary



Services



<u>Note:</u> In Figure 8 and 9 the NUTS 3 sub-region with the highest GVA per capita has been used as a numeraire. Scaling real GVA per capita relative to the median value has also been explored but yielded the same results.

The secondary and services sector three-dimensional stochastic kernel estimates of Figure 8, together with the associated stochastic kernel contours of Figure 9, tackle some of the shortcomings of the transition probability matrices as well as reiterating the main findings of the previous sections. Notwithstanding the pronounced peak at the beginning of the secondary sector stochastic kernel (capturing sub-regions that exhibit a very low a secondary sector real GVA per capita throughout the 1995-2004 period), the relatively low peaks in the middle of the distribution point to the probability of some shuffling of NUTS 3 sub-regions in the middle of the distribution. As for the peak at the upper end of the secondary sector real GVA per capita distribution, the stochastic contour plot of Figure 9 provides some useful insights. While the lower and middle sections of the of the contour plot follow a 45 degree line indicative of regions retaining their relative position over the 1995-2004 period, the upper section of the contour plot clearly appears to be off-diagonal (both above and below 45 degrees) – indicating mixed fortunes for what were high secondary sector GVA per capita sub-regions in 1995. The services sector real GVA per capita stochastic kernel estimate of Figure 8 is dominated by one characteristic – the outlier in the upper part of the distribution. As discussed in Section 4, this outlying observation is the real GVA per capita of Inner London West, the financial district of London and the NUTS 3 sub-region which has enjoyed the highest services sector real GVA per capita over the 1995-2004 period. The real GVA per capita of Inner London West has been almost three times that of the second placed NUTS 3 sub-region, Edinburgh, throughout the decade in question and almost seven times greater than the median value. As illustrated by the services contour graph, the lower part of the service sector distribution appears to veer below a 45 degree line trajectory, indicative of some relatively high services GVA sub-regions such as Berkshire, Milton Keynes, and Inner London East which, from the underlying real GVA per capita data, can be seen to have improved their relative positions over the 1995-2004 period. It is also splintered into a further cluster where Edinburgh would be expected to reside, as it retained its position as second highest services GVA per capita sub-region throughout the 1995-2004 period.

6. The growth equation approach

Having conveyed an impression of both the spatial disparity and the underlying distribution of British real GVA per capita, it is useful to verify that these impressions are consistent with the findings yielded by a test of absolute convergence over the 1995-2004 period. Table 5 provides a cross-sectional regression testing for absolute convergence over the period in question for aggregate real GVA per capita, as well as the secondary and services equivalents. As discussed in Section 3, there is at least one notable outlier in both the services sector and aggregate real GVA per capita NUTS 3 level data: the financial district of Inner London West. Furthermore, a number of remote NUTS 3 regions present in the data may also be outliers. The presence of such outliers may distort the fit of the OLS regression line. As in Section 3, the Least Trimmed Squares (LTS) regression method is utilized in order to address this problem.[18]

Table 5: LTS regression of sectoral GVA growth on initial sectoral GVA

Dependent variable: Average GVA Growth (1995-2004)			
	Secondary Sector	Services Sector	Aggregate
constant	0.277***	-0.083***	0.011
	(0.00)	(0.00)	(0.80)
$\ln GVA_{1995}$	-0.032***	0.016***	0.003
	(0.00)	(0.00)	(0.522)
R^2	0.40	0.18	0.00
Number of Obs	116	119	121

Notes: P-values in brackets. In the above regressions, outliers are those regions whose LTS standardized residuals 2.5 in magnitude.
Secondary sector outliers: Blackpool, Liverpool, Sefton, Derby, Solihull, Brighton and Hove, Isle of Wight, Torbay, Gwynedd, Central Valleys, Scottish Borders, Orkney Islands.
Service sector outliers: East Merseyside, East Derbyshire, South Nottinghamshire, Inner London West, Torbay, Isle of Anglesey, West Lothian, North Lanarkshire, Eileann Siar. Aggregate outliers: Solihull, Inner London West, Inner London East, Berkshire, Surrey, Edinburgh.

Looking at the cross-sectional "growth equation" for absolute convergence in Tables 5, aggregate GVA per capita trends across the 128 British NUTS 3 sub-regions over the 1995-2004 period appear to be inconclusive: aggregate GVA per capita does not exhibit a statistically significant movement towards divergence or convergence, and the explanatory power indicated by its R^2 is negligible. With regard to the individual sectors, much more pronounced trends emerge: services real GVA per capita displays a clear tendency towards divergence, while the secondary sector displays a strong convergence trend. The secondary sector's estimated annual speed of convergence is 3.15% per annum, while the service sector appears to be diverging at a rate of 1.61% per annum.[19] This suggests that the inconclusive aggregate GVA per capita trends observed over the 1995-2004 period conceal the contrasting divergence and convergence trends at play in the secondary and services sectors. When using the LTS estimator, it is, of course, important to examine the observations identified as outliers. The secondary sector outliers those regions which exhibit the lowest secondary sector real GVA per capita over the 1995-2004 time period and are for the most part either urban areas or extremities. The services sector outliers comprise of Inner London West, as expected, as well as a handful of regions scattered around the country which possess particularly low Services sector real GVA per capita. It is the outliers of the aggregate regression that appear to be a little less dispersed both in terms of wealth and location: all six are high GVA regions and four lie either in or adjacent to Greater London.

In all, it would appear that a coherent message emanates the maps, kernel density estimates, and absolute convergence test for the 1995-2004 period: clear convergence of the secondary sector as

[18] Unfortunately the time span of the available NUTS 3 GVA data (1995-2004) proved to be too short to undertake reliable panel data analysis.

[19] The speed of convergence, θ, is calculated as $\theta = \log(1-\beta)/k$, where k denotes the number of years in the time period.

secondary sector real GVA per capita compresses towards its average, and a trend towards divergence emerging from the services sector.

7. Conclusions and extensions

Having employed an array of diverse techniques to analyse the development of aggregate NUTS 3 real GVA per capita and its secondary and services components over the 1995-2004 period, it now remains to collect the various findings and identify any coherent trends which may emerge.

From the offset it is clear that breaking aggregate real GVA per capita in to its secondary and services components allows for a more accurate characterisation of British GVA growth as services share of aggregate GVA grew from 66% to 75% while the secondary (including construction) sector's share slumped from 30% to 22% over the 1996-2004 period. The merits of this approach can also been seen in the colour-coded maps of secondary, services, and aggregate real GVA per capita over the 1995-2004 period, where the dispersion of aggregate real GVA per capita across British sub-regions appears to be influenced much more strongly by services sector trends that by the secondary sector. The overview of the entire distribution of NUTS 3 level real GVA per capita provided by the kernel density estimates clearly indicate that aggregate GVA has not experienced a movement towards convergence over the 1995-2004 period – if anything, the upper tail of the distribution indicates a tendency towards divergence. Similar trends are visible for the services sector: no trace of a convergence process over the time period in question, and upper tail indicative of a divergence process. The secondary sector, on the other hand, exhibits clear convergence with what appears to be a multi-modal distribution becoming noticeably more unimodal over time.

Rather that checking for multi-modality using statistical tests, cross-sectional LTS growth equation regressions for absolute convergence are utilised to verify the non-parametric impressions. These absolute convergence regressions confirm aggregate real GVA's lack of convergence, the clear convergence of the secondary sector (with an estimated annual convergence speed of 3.5% per annum), and points to statistically significant divergence in the services sector (with an estimated annual divergence speed of 1.61% per annum). As for whether the secondary and services sectors are converging/diverging to "good" or "bad" states, the transition probability matrices along with the Stochastic Kernel estimates, as well as the summary statistics of 1995 and 2004 secondary and services GVA per capita, all point to an upwardly mobile services sector with high GVA NUTS 3 sub-regions inclined to pull away from chasing pack, while the secondary sector appears to be almost stagnant.

The techniques employed in this study are not without limitations. Three problems in particular that arise in this line of research are (i) the impact of commuter flows that inevitably permeate highly disaggregated regional data; (ii) the inconsistencies that arise between administrative districts, such as NUTS 3 sub-regions, and actual zones of economic activity. The boundaries of administrative districts are often influenced by tradition, local custom, or other arbitrary reason and may render these districts less than ideal for economic analysis; (iii) identifying and accurately controlling for spatial dependence between neighbouring geographic units.

It is hoped that a number of extensions to this work can help address some of these shortcomings: one useful approach, which has been advanced in regional growth studies by Casado-Diaz (2000) and Andersen (2003) is to bundle the NUTS 3 regions together into "Functional Economic Areas" using data on commuter flows to determine the size of each economic area. Such an approach would also tackle the difficulties posed by commuter flows, as each functional economic area would represent a district that is relatively self-contained in economic terms. In this way problems caused by the mismatch between the boundaries of administrative regions and economic regions can be addressed. With regard to spatial dependence, an array of diagnostic tests of spatial autocorrelation and spatial econometric techniques pioneered by Anselin (1988) have come to the fore in regional economic analysis in recent years. These developments, coupled with the compilation of a dataset of potential explanatory variables, offers a way forward which may yield a more detailed picture of British regional growth and the factors which influence the regional growth process.

While the most appropriate choice of technique for representing the empirics of regional British GVA per capita growth may yet be an open question, the newfound abundance of econometric approaches can only enhance the prospects of gaining a clearer insight into the British modern day regional development.

References

Andersen, A.K. (2002) Are Commuting Areas Relevant for the Delimitation of Administrative Regions in Denmark? *Regional Studies* 36(8), 833-844.

Anselin L. (1988) *Spatial Econometrics: Methods and Models*, Kluwert Academic.

Barro R. J. (1991) Economic Growth in a Cross-Section of Countries, *Quarterly Journal of Economics* 105, 407-443.

Baxter M, and R.G. King (1999) Measuring Business Cycles: Approximate Band-pass Filters for Economic Time Series, *The Review of Economics and Statistics* 81: 575–593.

Bishop P. and P. Gripaios (2004), Earnings Biases and Convergence in the UK: A County Level Analysis, *Applied Economics Letters* 11(1), pp.33-37.

Bishop P. and P. Gripaios (2006) Earnings Convergence in GB Counties: A Distribution Dynamics Approach, *Applied Economics Letters* 13, 29-33.

Boddy, M., Hudson, J., Plumridge A., Webber, J., (2005) Regional Productivity Differentials: Explaining the Gap, No 515, Discussion Papers from University of the West of England, School of Economics http://econpapers.repec.org/paper/uwewpaper/0515.htm

Bulli, S. (2001) "Distribution Dynamics and Cross-Country Convergence: A New Approach", *Scottish Journal of Political Economy* 100, 223-251.

Cameron G. and J. Muellbauer (2000) Earnings Biases in the UK Regional Accounts: Some Economic Policy and Research Implications, *Economic Journal* 110, pp.412-419.

Casado-Diaz J. M. (2000) Local Labour Market Areas in Spain: a Case Study, Regional Studies 34(9), 843–56.

Chatterji, M. and J. H. L Dewhurst (1996) Convergence Clubs and Relative Economics Performance in Great Britain: 1977–1991, *Regional Studies*, 30, pp. 31–40.

Christiano L. J. and T.J. Fitzgerald (1999) The Band Pass Filter. NBER Working Paper No. 7257, Cambridge, MA.

Department of Trade and Industry (2003) UK Productivity and Competitiveness Indicators Department of Trade and Industry, London.

Department of Trade and Industry (2004) Regional Competitiveness and the State of the Regions Department of Trade and Industry, London.

Dewhurst J. (1998) Convergence and Divergence in Regional Household Incomes per Head in the United Kingdom, 1984-93, *Applied Economics* 30, pp.31-35.

Duranton, G. and V. Monastiriotis (2002) Mind the Gap: the Evolution of Regional Earnings Inequalities in the UK, 1982-1997, *Journal of Regional Science* 42 (2): 219-256.

DTI (2003) A Modern Regional Policy for the UK, TSO, London.

Floud, R. and P. Johnson (Eds.) (2004) The Cambridge Economic History of Modern Britain (Volume 3) Cambridge University Press.

Henley, A. (2005) On Regional Growth Convergence in Great Britain, *Regional Studies* 39, 1245-1260

Li, Q. (1996) Nonparametric Testing of Closeness Between Two Unknown Distribution Functions, *Econometric Reviews* 15, 261-274.

Monastiriotis, V. (2006) Sub-regional Disparities in Britain: Convergence, Asymmetries and Spatial Dependence", Research Papers in Environmental and Spatial Analysis No. 12, European Institute, London School of Economics.

ODPM (2003) Reducing Regional Disparities in Prosperity, ODPM Housing, Planning, Local Government and the Regions Committee, TSO, London.

Office of National Statistics (2002) Annual Respondents Database, Office of National Statistics, London.

Park, B.U. and B.A. Turlach (1992) Practical Performance of Several Data Driven Bandwidth Selectors (with Discussion), *Computational Statistics* 7, 251-285.

Quah, D. T. (1996a) Twin Peaks: Growth and Convergence in Models of Distribution Dynamics, *The Economic Journal* 106, 1045-1055.

Quah, D.T. (1996b) Convergence Empirics Across Economies with (some) Capital Mobility, *Journal of Economic Growth* 1, 95-124.

Quah, D.T. (1997) Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs, *Journal of Economic Growth* 2, 27-59.

Quah, D. T. (2006) Growth and Distribution, unpublished manuscript, London School of Economics (see http://econ.lse.ac.uk/staff/dquah/p/gnd.pdf).

Rice, P., and A. J. Venables (2004) Spatial Determinants of Productivity: Analysis for the Regions of Great Britain, *CEPR Discussion Papers* 4527, C.E.P.R. Discussion Papers.

Leroy, A.M. and P.J. Rousseeuw (1987) *Robust Regression and Outlier Detection*, Wiley: New York.

Silverman, B.W. (1981) Using Kernel Density Estimates to Investigate Multimodality, *Journal of the Royal Statistical Society* 43, 97-99.

Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*, London (Chapman and Hall).

Timm, N.H. (2002) *Applied Multivariate Analysis*, Berlin (Springer Verlag).

Wand, M.P. and M.C. Jones (1995) *Kernel Smoothing*, London (Chapman and Hall).

Sectoral Trends and British Regional Economic Growth – A Spatial Econometric Perspective

Declan Curran

1. Introduction

The emergence of an impressive array of spatial econometric techniques in recent years has helped give geographic factors a more realistic characterisation in regional economic analysis. The importance of this spatial dimension has never been in doubt, but now the tools are available to give a more vivid depiction of the impact on an area from its core or peripheral location, proximity to natural resources, and spillover effects from neighbouring regions. This paper builds upon the work of Henley (2005) and Monastiriotis (2006) and employs these spatial techniques to shed light on the regional growth process occurring in Britain over the 1995-2004 period. Regional disparities have been synonymous with modern day British economic development and their influence can still be seen in current regional growth trends. In 2005 the gross value added (GVA) per head of population for the UK was £17,700, with London had the highest regional GVA per head of population (£24,100), and South East following with £20,400.[20] The East of England (£18,900) was the only other region to have a GVA per head of population higher than the national average.[21] Wales had the lowest GVA per head of population at £13,800.[22] That said, there have been signs recently that these disparities may be lessening: in 2005 the North East enjoyed, along with the East Midlands and London, the strongest GVA per head growth (4.4 per cent), while the lowest growth rate (3.5 per cent) was experienced in the South East. [23]

This analysis of British regional economic development focuses on NUTS 3 real GVA per capita data spanning from 1995-2004, not just for aggregate British GVA per capita but also for secondary and services sectors. In this way, it is possible to look beneath the surface of the British sub-regional aggregate GVA growth process experienced over the period 1995-2004, by examining to what extent this process may have been driven by the differing growth dynamics of the secondary and services

[20] Throughout this paper, the term "regions" denotes British NUTS 1 level disaggregation, "counties" denote British NUTS 2 disaggregation, and "sub-regions" denote British NUTS 3 level disaggregation. The term "regional economic growth" is used in a general sense to refer to the field of literature to which this paper belongs.

[21] GVA is defined as follows: Under European System of Accounts 95 (ESA95), the term GVA is used to denote estimates that were previously known as Gross Domestic Product (GDP) at basic prices. Under ESA95 the term GDP denotes GVA plus taxes (less subsidies) on products, i.e. at market prices.

[22] Data available from the Office of National Statistics (ONS) at http://www.statistics.gov.uk/CCI/nugget.asp?ID=420&Pos=&ColRank=1&Rank=374.

[23] The quantity of real GDP generated each geographic unit, scaled by that unit's population, is a standard proxy for the productivity in the face of data constraints at high levels of disaggregation. It is not intended to represent income per capita. For a treatment of regional productivity differentials based on individual business units see Boddy et al (2005).

sectors. This approach also finds support in the work of Boddy et. al. (2005) who, in their study of productivity differentials based on individual business units, find that "the scale of difference in productivity between particular sectors is very considerable". Two problems often emerge in studies utilising highly disaggregated regional data: (i) neglect of the impact of commuter flows and (ii) the administrative delineation of regions may not reflect self-contained economic areas. This paper attempts to address these two issue by constructing a set of functional economic regions for Britain, where the 128 NUTS 3 regions are aggregated together using a method based on commuter flow data. These functional economic areas provide a means for checking the robustness of results emanating from the econometric analysis carried out on the NUTS 3 level data. While the time-span (1995-2004) considered in this paper is dictated by data availability, this decade is nonetheless an importance one. It captures a period of time where regional growth in many developed countries has been impacted by the move towards outsourcing of manufacturing and the absorption of phenomenal technological advances. Britain is no exception to this trend: in 2004 primary, secondary, and services as defined above accounted for approximately 1%, 22% and 75% of British GVA, while the equivalent shares in 1995 were 2%, 30% and 66%, respectively.[24] This surge in services sector output, accompanied by a falling off of secondary output, justifies a more disaggregated approach to the convergence/divergence debate.

This paper is organised as follows: Section 2 provides a description of the data used in this paper, as well as a brief review of the literature on British regional growth in the years prior to 1995. The spatial dispersion of British real GVA per capita is also discussed, with a set of colour-coded maps provided. A description of how β-convergence analysis has been augmented to include a number of spatial econometric methods is provided in Section 3. The section concludes with an outline of the approach adopted in this paper for allocating British NUTS 3 regions to functional economic regions. The results of the spatial econometric analysis testing for absolute and conditional convergence are reported in Section 4. Conclusions are then presented in Section 5.

2. Data Issues and Background

This paper is primarily focused on NUTS 3 level gross value added (GVA) data, which is available for the period 1995-2004. Turning to the 1995-2004 period, unadjusted (constrained to headline NUTS2) aggregate GVA by NUTS3 area at current basic prices for the years 1995 to 2004 is available from the Office of National Statistics (www.statistics.gov.uk), as well as being disaggregated for 1) agriculture, hunting and forestry 2) Industry, including energy and construction and 3) service activities, including Financial Intermediation Services Indirectly Measured (FISIM). These three categories are henceforth referred to as "primary", "secondary", and "services", respectively. Estimates of workplace based GVA allocate income to the region in which commuters work. Per capita estimates can then be

[24] Calculations based on National Accounts GVA data available from Office of National Statistics, as discussed in Section 2.

constructed using NUTS 3 level population data available from Nomis Labour Market Statistics (www.nomisweb.co.uk). Unfortunately, regional deflators such as the Retail Price Index (RPI) are only available for UK for the years 2000, 2003, and 2004, and the methodology for this index is still in a formative stage. One could just use the yearly national deflator for each NUTS 3 region. However, this is unsatisfactory as it makes no allowance whatsoever for regional price differences – particularly British secondary, services, and aggregate GVA per capita exhibit clear regional trends, as illustrated in Figures 1-3. In this study, regional deflators for each year have been constructed by weighting the 1995-99 national RPI figure by the 2000 regional RPI weights. Similarly for 2001-2002 regional RPI the 2003 regional RPI figures are used as weights. The basket used to calculate the RPI figures include both consumer goods and services such as household services, personal services, and leisure services.[25]

By way of background, it should be noted that studies of British regional growth patterns over the 1977-1995 period, based on National Accounts GDP per capita data for the 62 British counties and New Earnings Survey data, have identified a number of prominent features.[26] Chatterji and Dewhurst (1996) conclude that Regional GDP per capita data yields no evidence of convergence exists over this time period, though they do identify some sub-periods that exhibit convergence (in periods where the economy as a whole was experiencing slow growth). Bishop and Gripaios (2004) find no signs of convergence over the 1977-1995 period, regardless of whether one uses National Accounts or New Earnings Survey data. A further insight to emerge from this line of research has been the influence of geographic location and spatial factors on British regional growth. Dewhurst (1998) finds evidence of the influence of the fore-mentioned "north-south divide" on British regional growth patterns and Bishop and Gripaios (2004) also find a significant "north-south divide" effect, which acts to the detriment of the northern areas. More recently a whole range of spatial economic techniques have become available, allowing for a more refined characterisation of the spatial dimension in the regional growth process. When this spatial component is controlled for in convergence analysis, there are signs that not only did Britain not experience regional convergence in recent decades, but there may even have been a process of divergence in action. Monastiriotis (2006), using wage data from the New Earnings Survey, points to widening aggregate wage disparities throughout the 1980s and 1990s when the issue of spatial dependence is taken into account. Henley (2006) has undertaken a spatial econometric analysis of NUT 3 level aggregate GVA data for the 1995-2001 period and concludes that British NUTS 3 sub-regions experienced divergence over this time period.

In order to provide a visual impression of the spatial dispersion of real GVA per capita across British NUTS 3 sub-regions a set of maps are presented (Figures 1-3). Each map is colour coded, with the light shading denoting 0-100% of median real GVA per capita, medium shading denoting 100-125%, and dark shading denoting over 125% of median real GVA per capita. Each sub-region is shown

[25] Fure further details of the composition of the RPI series, see the ONS publication *Economic Trends 615*, February 2005.

[26] For the purposes of this study, only Great Britain is considered, i.e. Northern Ireland is not included.

relative to the median rather than the mean to mitigate the impact of outliers such as the services GVA of Inner London West. Figure 1 presents aggregate real GVA per capita for 1995 and 2004. Salient features include the apparent spatial clustering of high GVA per capita in greater London, Manchester-Liverpool, Edinburgh, Glasgow, and Aberdeen (near the North Sea oil fields); a clear expansion of the greater London high-GVA area over the period in question; the noticeable improvement of the Midlands; and the apparent falling back of Northern England and Scotland. One might wonder whether these impressions are reflected in the development of the secondary and services sectors over the 1995-2004 period. As illustrated in Figure 2, however, the secondary industry presents a more mixed picture: the North of England appears to fall back, relatively speaking; a belt of increased GVA per capita is apparent in the Midlands, while the South West and South East exhibit some shuffling of regions between the three categories, but no clear pattern. The services sector (Figure 3) highlights the strength of the high-GVA greater London area, increases in Liverpool-Manchester, but continued sluggishness in Northern England and Scotland. In all it would appear that it is the services industry which drives the expansion of the southern high GVA in the aggregate map. While the secondary sector does appear to be the more dispersed in terms of the highest GVA category; this trend seems to be eclipsed in the aggregate GVA map by the strong services performance.

Further descriptive evidence of sub-regional GVA per capita trends can be gleaned from the summary statistics presented in Table 1.

Table 1: Summary Statistics for Secondary and Services real GVA per capita, 1995 and 2004

Secondary Sector GVA per capita (2002 UK£)			**Services Sector GVA per capita (2002 UK£)**		
	1995	**2004**		**1995**	**2004**
Mean	3,517.29	4,031.723	Mean	6,422.84	11,261.36
Median	3,343.53	3,964.373	Median	5,828.70	9,708.08
Maximum	7,068.65	8,383.499	Maximum	41,398.86	64,654.04
Minimum	1,634.15	1,648.837	Minimum	3,050.08	5,766.21
Std. Dev.	1,162.03	1,168.167	Std. Dev.	3,574.93	6,023.20

The contrast between secondary and services sector GVA per capita developments over the 1995-2004 period is stark. The virtually unchanged mean, median, and standard deviation of secondary GVA per capita over the 10 year period, together with slight increases in the minimum and maximum GVA per capita figures suggest that any convergence experienced in the secondary sector has not been a buoyant one. Services GVA per capita, on the other hand, bears all the hallmarks of a sector on the move, with its mean and median showing marked increases over the 10 years and its widening standard deviation indicative of the absolute divergence hinted at by this section's kernel estimation analysis.

Figure 1: Aggregate Real GVA Per Capita, 1995 (left) and 2004 (right)

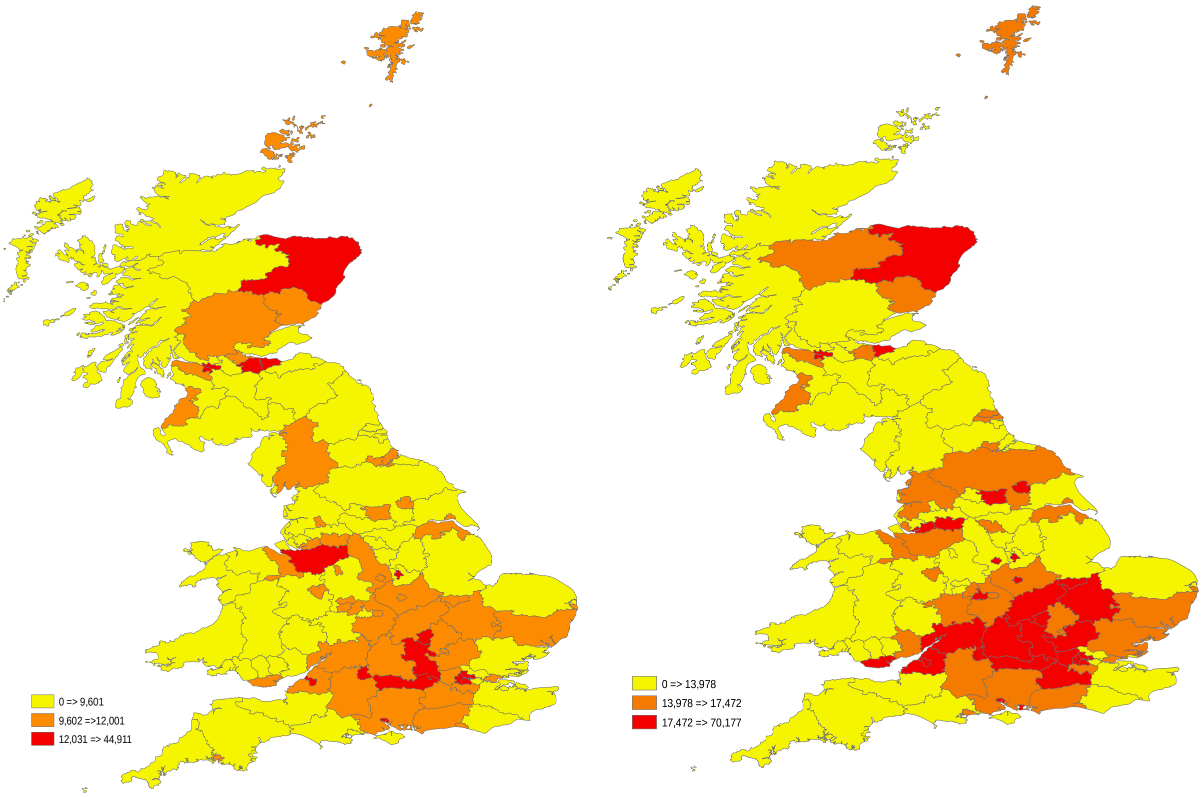


Figure 2: Secondary Sector Real GVA Per Capita, 1995 (left) and 2004 (right)

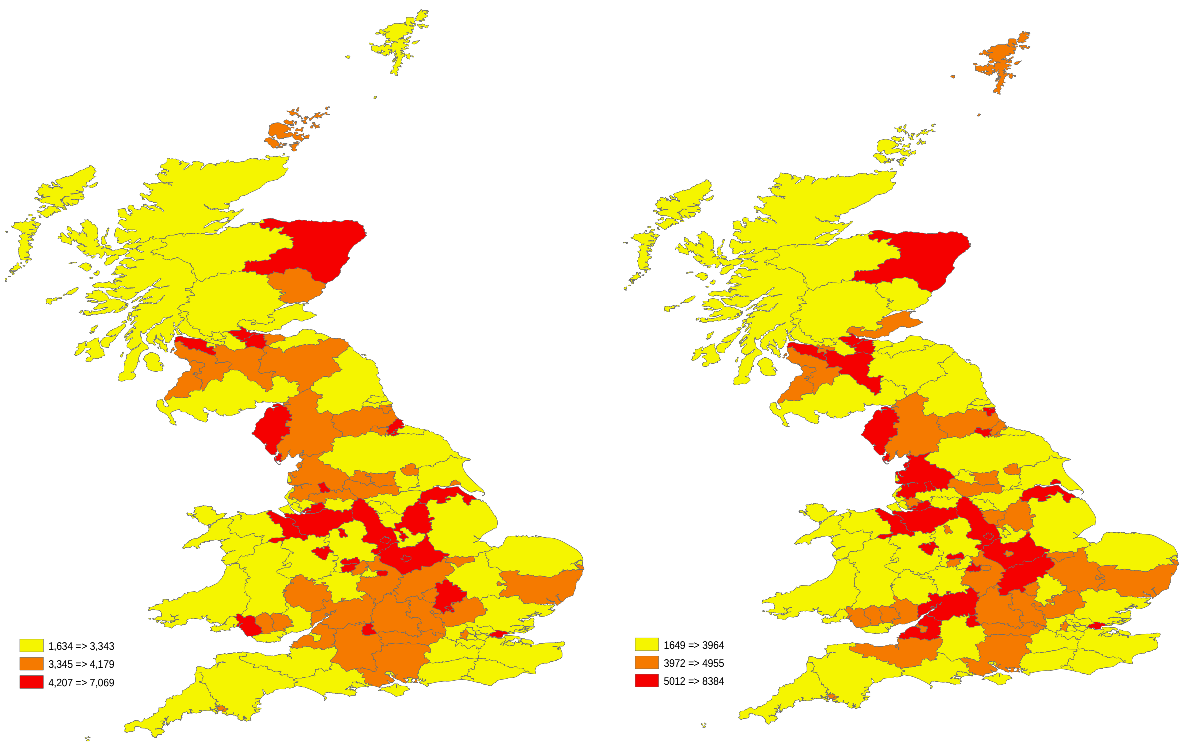
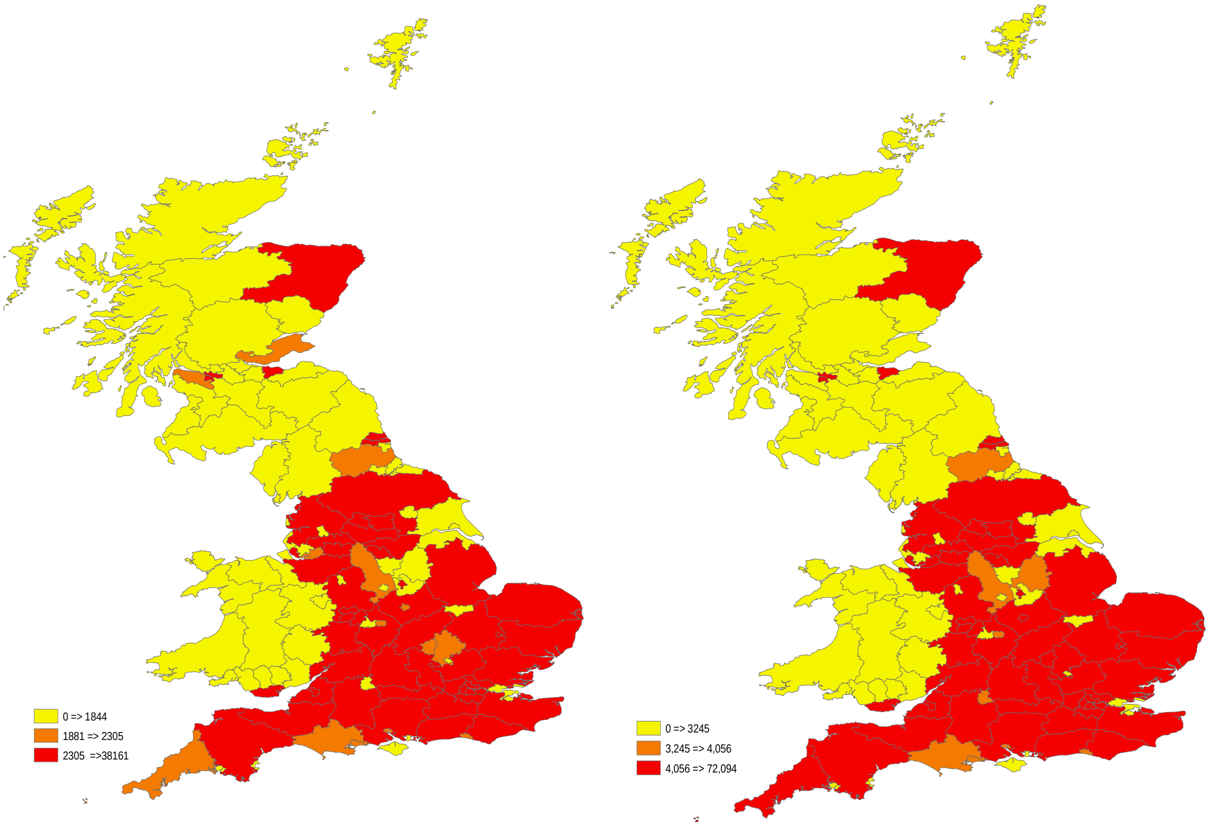


Figure 3: Services Sector Real GVA Per Capita, 1995 (left) and 2004 (right)



In Sections 3 and 4 a number of additional data sources are drawn upon. NUTS 3 level commuter flow data used in the construction of British functional economic areas is available from the Labour Force Survey Data Service (lfs.dataservice@ons.gov.uk). The explanatory variables introduced in the conditional convergence analysis of Section 4 utilises average primary school pupil-teacher ratio per county and the average A-level pass rate achieved by pupils in each county, both of which are available from the ONS publication *Regional Trends*. The number of businesses registered for Value Added Tax and female employment in expressed as a proportion of people aged 16+ are both available from Nomis Labour Market Statistics (www.nomisweb.co.uk). Capital expenditure for British sub-regions is available from the ONS series *Regions in Figures*.[27]

3. Regional Convergence and the Spatial Dimension

This section begins with a brief description of how β-convergence analysis, as developed by Baumol (1986), Barro and Sala-I-Martin (1992), and Mankiw et al. (1992), has been augmented to include a number of spatial econometric methods. When considering regional convergence, various empirical

[27] *Region in Figures* has now been discontinued. The final edition was Winter 2004/05 (volume 9). It has now been replaced by a new publication, *Regional Snapshot*.

approaches have been implemented in the literature: from simple plots of measures of dispersion over time to intra-distributional dynamics using Markov chains applied to GDP per capita. It is β-convergence analysis, however, that has lent itself most easily to spatial econometric analysis. This section then discusses methods for constructing functional economic areas from administrative regions. The section concludes with an outline of the approach adopted in this paper for allocating British NUTS 3 regions to functional economic regions.

3.1 Spatial Convergence and the Modelling of Regional Growth

While a variety of distinct convergence concepts have emanated from the economic growth literature, one form of convergence which has received particular attention over the last two decades has been that of β-convergence. This form of convergence occurs when poor regions grow faster than richer regions, resulting in a catching-up process where the poor regions close the economic gap that exists between their richer counterparts. The now-standard specification of β-convergence can be expressed in vector form as follows:

$$(1) \qquad \mathbf{ln}\left(\frac{y_{t+k}}{y_t}\right) = \alpha + \mathbf{(}1 - e^{-\lambda k}\mathbf{)}\,\mathbf{ln(}y_t\mathbf{)} + \varepsilon_t$$

where $y_{i,t}$ is per capita income of state i in year t; α represents the intercept term, and $(1\text{-}e^{-\lambda k})$ is the convergence coefficient, which is usually reparametrized as $\beta = (1\text{-}e^{-\lambda k})$. The β coefficient is then estimated using Ordinary Least Squares (OLS), and the speed of convergence, λ, can then be calculated. A negative estimate for β indicates that growth rates of per capita income over the k years is negatively correlated with initial incomes – a finding which is interpreted as a support for the hypothesis of convergence. It is assumed that the error terms from different regions are independent:

$$(2) \qquad E[\varepsilon_t \varepsilon_t'] = \sigma_t^2 I\ .$$

This unconditional β-convergence specification can then be augmented, as per Barro and Sala-I-Martin (1992), to include a range of control variables (such as differences in human capital accumulation, infrastructure disparities, industrial structure, as well as dummy variables reflecting different regional characteristics) which may capture differences in the paths of steady-state GVA per capita.

Equations (1) and (2) can be augmented to capture interactions across space, a refinement which reflects more accurately the realities of the growth process across regions. As Henley (2006) notes, this spatial dimension can exert its influence on regional growth through numerous channels: adjustment costs and barriers to labour and capital mobility, spatial patterns in technological diffusion, the ability of regions to pursue independent regional growth policies, and the extent to which neighbouring regions interact and benefit from spillover effects. Any analysis which ignores the influence of spatial location on the growth process runs the risk of producing biased results. Following from Anselin (1988), spatial dependence has been incorporated into the β-convergence specification in two ways: it can be included as an explanatory variable in the specification or it can be modelled as operating through the error process.[28] The former, known as a Spatial Autoregressive Model (SAR), depicts a region's growth as being directly affected by growth in neighbouring regions. This direct spatial effect is independent of the exogenous variables and is captured by including a spatial autoregressive parameter, ρ, and a spatial weight matrix, W, in the specification:

(3) $$\ln\left(\frac{y_{t+k}}{y_t}\right) = \alpha + (1 - e^{-\lambda k})\ln(y_t) + \rho W \ln\left(\frac{y_{t+k}}{y_t}\right) + \varepsilon_t$$

In equation (3), the growth of a given region is influenced by the growth rate of adjacent regions. This "spatial lag" approach can also be utilised where a region's growth rate is though to be influenced by the initial income level of adjacent regions, a specification which Rey and Montouri (1999) refer to as a spatial cross-regressive model:

(4) $$\ln\left(\frac{y_{t+k}}{y_t}\right) = \alpha + (1 - e^{-\lambda k})\ln(y_t) + \tau W \ln(y_t) + \varepsilon_t \cdot$$

It may be the case that, rather being directly affected by the growth rate of its neighbours, a region's growth rate may be influenced by a complex set of random, unexpected shocks transmitted across space. Such unexpected shocks take the form of spillovers associated with technology or consumer tastes. In this SEM case, the spatial influence does not enter the systematic component of the specification. Instead, it is captured in an error term which contains a spatial error coefficient, ζ, and an idiosyncratic component, u, where $u \sim N(0, \sigma^2 I)$.

(5) $$\ln\left(\frac{y_{t+k}}{y_t}\right) = \alpha + (1 - e^{-\lambda k})\ln(y_t) + \varepsilon_t \qquad \text{where} \quad \varepsilon_t = \zeta W \varepsilon_t + u_t$$

Section 4 reports results for cross-sectional growth equation regressions which test for absolute and conditional convergence using the SAR and SEM specifications.

3.2 Functional Economic Areas

It is entirely possible that the administrative areas into which a country is divided may not coincide with patterns of economic activity on the ground. Administrative areas may differ from areas of economic activity due to factors such as local democracy or local customs, and these differences may

[28] For more detailed treatment of spatial autoregressive and spatial error models, see Bernat (1996), Rey and Montouri (1999), and Fingleton and Lopez-Baso (2006).

be perpetuated over time. These areas of economic activity have been termed *functional economic areas* (or *local labour market areas*, or *commuting areas*, or *travel to-work-areas*) and have been focus of much research, as illustrated by Coombes et al. (1986), Casado-Diaz (2000), and Andersen (2002), to name but a few. A functional economic area can be characterised by a high frequency of intra-regional interaction, for example, intra-regional trade in goods and services or labour commuting As Andersen(2002) notes, the divergence between administrative and functional economic areas may lead to tensions between administrative authorities, inefficient planning of infrastructure, or sub-optimal labour market policies. This mismatch between administrative and functional economic areas may also have repercussions for those interested in studying the regional economic growth process: findings based on data disaggregated along administrative lines may not fully reflect the economic realities at regional level. It is understandable then that some effort should be invested in checking whether findings based on administrative-area data are consistent with those that would emerge if a functional economic delineation were used.[29]

The problem, of course, is how to identify functional economic areas and delineate them in a meaningful, consistent way. Karlsson and Olsson (2006) outline three theoretical approaches to delineating functional economic areas: (i) the local labour market approach, where one-way commuting data can be used to indicate the existence of wage differentials between areas. Focal regions are identified which are self-contained in terms of commuter flow and then the remaining areas are assigned to these cores, based on commuter flows. The borders are found when areas have equal attraction to both of the closest foci. (ii) the commuting zone approach, which is similar to (i), but hinges less on urban foci and instead considers the existing mutual dependencies of regions. The interaction between regions is calculated using commuter flows in both directions (iii) the accessibility approach, which uses commuting time to proxy the potential interaction between areas. The approach used to delineate functional economic areas in this paper is in keeping with (ii) above, as it uses commuter flows in both directions as a means to approximating the interaction between administrative areas. The methodology adopted in this paper is now described in more detail.

The methodology used here for delineating functional economic regions owes its origins to Coombes et al. (1986), who use micro-level data to divide Britain up into Travel-to-Work Areas (TTWAs). These TTWAs incorporate commuting data in their definition and utilise census data in the delineation process. The algorithm discussed presently, originally constructed in Coombes et al. (1986), consists of three phases: i) possible foci are identified; ii) unallocated units are assigned to these foci; iii) the process is iterated until all regions are deemed self-contained or “closed” in an economic sense, as

[29] The broader issue is that of the Modifiable Areal Unit Problem (MAUP), which occurs whenever arbitrarily defined boundaries are used for measurement and reporting of spatial phenomena. This problem may be alleviated by analysing data at various levels of disaggregation or by taking highly disaggregated spatial units and aggregating them in a context driven by an economic or demographic factor that is not arbitrary. In recent times a number of GIS computational methods have been developed in order to provide a consistent, uniform method for addressing Modifiable Areal Unit Problem (MAUP). For a more detailed treatment of MAUP, see Openshaw (1984). Efforts to address MAUP using GIS technology are discussed in Openshaw and Alvanides (1999).

defined in the methodology. As outlined presently, this methodology has been further refined by Eurostat (1992), Casado- Diaz (2000), and Andersen (2002).

The three phases of the methodology are undertaken in the following manner:

(i) Identification possible foci: at least one area or "couple" constitutes a focus to which all other areas will be assigned. A "couple" occurs where two areas have the highest total of in-commuting and out-commuting with each other. As per Coombes et al. (1986) and Casado-Diaz (2000), foci are identified using a supply-side and a demand-side self-containment condition. The supply-side self-containment condition captures the extent to which the resident working population work in their area of residence, while the demand-side self-containment condition captures the extent to which jobs in a given area were filled by residents of that area.

The supply-side self-containment condition expresses the number of residents who live and work in area i as a proportion of the total number of workers in area i (residents who live and work in area i plus inward commuters). Let C_{ji} denote the number of commuters travelling from area j to area i, and C_{ij} denoting those commuting in the opposite direction. The total number of inward commuters to area i can then be represented as $\Sigma_{i=1} C_{ji}$ and outward commuters from area i can be represented as $\Sigma_{i=1} C_{ij}$. The total number of residents who live and work in area i is denoted as TR. The supply-side containment condition can be stated as follows:

(6) $$\frac{TR}{TR + \sum_{i=1} C_{ij}},$$

and the demand-side self-containment condition, which expresses the number of residents who live and work in area i as a proportion of the total number of jobs in area i, can be stated as:

(7) $$\frac{TR}{TR + \sum_{i=1} C_{ji}}.$$

In keeping with Eurostat (1992), a self-containment level of 70% or over for both conditions is required for an area to qualify as a focus and areas must have a population of over 50,000.

(ii) Assignment of the remaining areas to the focus with which they exhibit the highest interaction. The interaction of area i with area j (or any potential focus) is approximated by the sum of commuter flows in both direction between areas i and j expressed as a proportion of total commuter flows to and from area i:

(8) $$\frac{C_{ij} + C_{ji}}{\sum_{i=1} C_{ij} + \sum_{i=1} C_{ji}}$$

(iii) Having assigned all areas to potential foci, it now remains to be seen if each newly constructed functional economic area is now sufficiently "closed". Andersen (2002) has developed a measure of how closed a functional economic area is. The number of residents who live and work in this newly constructed functional economic area (TR_{FEA}) is expressed as a proportion of that functional economic area's total commuter inflow and outflows. This ratio, denoted in equation (9) as κ must then exceed a certain threshold value for the functional economic area to be deemed "closed".

(9) $$\kappa = \frac{TR_{FEA}}{\sum_{i=1} C_{ij} + \sum_{i=1} C_{ji}}$$

The choice of the threshold value, κ, is exogenous and is ultimately data-driven – clearly an unsatisfactory situation. However, as Andersen (2002) notes, it seems reasonable to argue that this threshold should not be less than 1, because a value less than 1 would suggest that commuter flows into and out of the functional economic area were greater than the number of residents living and working in the functional economic area. In light of this, the rather lax but intuitively understandable lower bound threshold value, κ=1, is used in this paper to determine whether the functional economic areas are "closed". Examination of the data used in this paper, Great Britain's 128 NUTS 3 sub-regions, also supports using this threshold value as it ensures that the number of functional economic areas obtained is not at the extremely low or high end of the 128 NUTS 3 sub-region total. Where functional economic areas are not deemed to be sufficiently "closed", the procedure is iterated until all these "not closed" areas are assigned to another focus and the threshold is attained. In the case of remote areas which may not have any interaction with a foci, these remote areas are assigned to the area with which they share the highest interaction. The final delineation of the 64 British functional economic areas, constructed using NUTS 3 level commuter flow data for the year 2000, is displayed in Figure 4, and a complete list of the functional economic areas is provided in Appendix 1.

Figure 4: Functional Economic Areas for Britain Based on Commuter Flow Data

Having illustrated the 64 functional economic areas in Figure 4, it is natural to wonder how satisfactory these constructed areas are in capturing the reality of British regional patterns "on the ground". An intuitive indication might be gained from comparing the functional economic areas (and their underlying NUTS 3 sub-region components) with existing urban conurbations. Here, the six English metropolitan counties and Greater London are used to give an impression of the performance of the functional economic areas.[30] As outlined in Table 2, these metropolitan counties (and Greater London) envelope a number of NUTS 3 sub-regions. How do the functional economic areas handle these NUTS 3 regions? One would expect that sensible functional economic areas should not separate NUTS 3 regions that the metropolitan counties suggest should be grouped together. As illustrated in Table 2, the functional economic areas are consistent with the amalgamation of metropolitan counties

[30] The English metropolitan county sub-division was created by the Local Government Act, 1972. The administrative area of Greater London is not a Metropolitian county, as it was created earlier (Local Government Act, 1963). For completeness, Edinburgh (ED) and Glasgow (GL) are indicated in Figure 4.

(and Greater London) NUTS 3 regions, with a number of neighbouring sub-regions added to these metropolitan counties where the commuter flow data deemed appropriate.

Table 2: Comparison of Metropolitan Counties and Functional Economic Areas

Metropolitan county	**Constituent NUTS 3 Sub-regions**	**Comparable Functional Economic Area(s)**
Greater Manchester (MN)	1.Greater Manchester North 2. Greater Manchester South	• North Manchester, South Manchester, and Cheshire
Merseyside (LV)	1.Liverpool, 2. Sefton, 3. East Merseyside, 4. Wirral	• Liverpool, Sefton, East Merseyside, Wirral, Halton and Warrington
South Yorkshire (SW)	1. Sheffield, 2. Barnsley, Doncaster, Rotherham	• Barnsley Doncaster and Rotherhan, Sheffiled, East Derbyshire
West Yorkshire (WY)	1. Leeds, 2. Bradford, 3. Wakefield, Calderdale, and Kirklees	• Leeds and Calderdale, Kirklees, and Wakefield • Bradford
Tyne and Wear (TN)	1. Sunderland, 2.Tyneside	• Northumberland and Tyneside • Sunderland
West Midlands (W)	1. Birmingham, 2. Coventry, 3. Wolverhampton, Walsall 4. Dudley, Sandwell, 5. Solihull,	• Birmingham, Solihull, Dudley, Sandiwell, Wolverhampton and Walsall • Warwickshire, Coventry
Greater London (LN)	1. Inner London West 2. Inner London East 3. Outer London East and North East 4. Outer London South 5. Outer London West and North West	• Inner London West ,Inner London East Outer London East and North East, Outer London South, Outer London West and North West, Hertfordshire, Buckinghamshire, Surrey

Note: Abbreviations in brackets identify the metropolitan counties in Figure 4.

One curious functional economic area in Table 2 is the city of Bradford. This raises the issue of the plausibility of the smaller functional economic areas visible in Figure 4. An inspection of the supply-side and demand-side self containment criteria derived from commuter flows (equations (6) and (7) above) and the measure of self containment (equation (9)) for the smallest, in geographical terms, of these areas suggests that as Bradford, Swindon, Plymouth, and Swansea do appear to be quite self-contained, though Sunderland and York are more debatable.[31]

Table 3: Self-Containment of Smallest Functional Economic Areas

Functional Economic Area	Supply–Side Condition (70% Threshold)	Demand–Side Condition (70% Threshold)	Measure of Self-containment (κ=1)
Bradford	0.77	0.79	1.79
Swindon	0.79	0.77	1.74
Plymouth	0.93	0.79	2.96
Swansea	0.84	0.80	2.28
Sunderland	0.73	0.71	1.29
York	0.71	0.72	1.26

[31] The full version of the data given in Table 3 for all 64 functional economic areas constructed in this paper is available from the author on request.

In all, the functional economic areas constructed in this paper appear to serve as a useful basis for checking the robustness of regional econometric analysis undertaken in Section 4, based on administrative NUTS 3 level units.

4. Spatial Analysis of β-convergence

The focus now turns to establishing the empirics of regional growth and β-convergence across British sub-regions, in the presence of possible spatial dependence. The first step is to statistically test for the presence of spatial autocorrelation in sub-regional secondary, services and aggregate real GVA data. From Figures 1-3 it appears that clear spatial patterns exist in the geographic dispersion of secondary, services and aggregate real GVA across British sub-regions. In order to confirm this, the well-known diagnostic for global spatial autocorrelation, Moran's I statistic, is utilised. Once the presence of spatial autocorrelation has been established, the issue of convergence across sub-regions is then considered. As outlined in Section 3, the cross-sectional growth equations which test the hypotheses of absolute conditional convergence are easily augmented to incorporate spatial autoregressive (SAR) components and spatial error (SEM) components. What is more, the inclusion of a set of explanatory variables in the conditional convergence growth equation allows one to identify those factors which may explain the trends observed in British sub-regional growth over the 1995-2004 period.

4.1. Diagnostic Test for Spatial Autocorrelation

The Moran's I statistic for spatial autocorrelation yields a test statistic which can be defined as follows:

(10) $$I_t = \left(\frac{n}{s}\right)\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} y_{it} y_{jt}}{\sum_{i=1}^{n}\sum_{j=1}^{n} y_{ij} y_{jt}}$$

where w_{ij} represents the elements of the spatial weighting matrix W, and n denotes the total number of sub-regions. The results of this diagnostic test for spatial autocorrelation on secondary, services and aggregate log real GVA per capita for 1995 and 2004, as well as for real GVA per capita growth over the 1995-2004 period, are reported in Table 4. The test has been carried out using two different types of spatial weighting matrix: i) a binary contiguity matrix, where $w_{ij} = 1$ if sub-regions are geographically adjacent, and $w_{ij} = 0$; ii) a distance-based spatial weighting matrix, where w_{ij} denotes the distance between sub-regions i and j.

Table 4: Moran's *I* Global Spatial Autocorrelation Statistic

	Secondary		Services		Aggregate	
	Binary *W*	Distance *W*	Binary *W*	Distance *W*	Binary *W*	Binary *W*
Log real GVA per capita 1995	0.115**	0.115**	0.200**	0.111***	0.114**	0.114**
Log real GVA per capita 2004	0.156**	0.156**	0.238***	0.141***	0.197**	0.197**
GVA Growth 1995-2004	0.017	0.017	0.198***	0.043**	0.123**	0.123**

Note: Significance at ***1%, **5%, and *10% level.

It is clear from Table 4 that secondary, services, and aggregate real GVA per capita do indeed exhibit strong spatial autocorrelation across sub-regions in both 1995 and 2004, the start- and end-point of the dataset used in this paper. However, when one considers growth rates over the 1995-2004 period, it is just services and aggregate GVA per capita that exhibit spatial autocorrelation, which suggests that it aggregate GVA growth spatial autocorrelation over the 1995-2004 period has been driven by that of the services sector. These findings appear to be robust to the type of spatial weighting matrix used in the Moran's *I* statistic.

4.2. Absolute β-convergence

Tables 3 and 4 below present spatial autoregressive (SAR) and spatial error and (SEM) cross-sectional regressions of secondary, services, and aggregate GVA per capita growth on initial, 1995, log GVA per capita $(lnGVA_{1995})$ – as outlined in Section 3. This is the standard test for absolute β-convergence (augmented to capture two distinct types of spatial influence), where a negative significant coefficient on initial log GVA indicates convergence and a positive significant coefficient indicates divergence. GVA per capita data for the full set of 128 NUT 3 sub-regions are used in the specifications in Table 5. The results reported in Table 6 relate to the 64 functional economic areas constructed in Section 3.3. For this purpose NUTS 3 level real GVA data have been allocated into functional economic areas and then divided by the relevant population figure. In keeping with the notation of Section 3, ρ and τ represent the spatial autocorrelation coefficient and spatial error coefficient, respectively. The spatial weighting matrix used in throughout this section is the binary contiguity matrix. Appendix 2 replicates this regression analysis using the distance spatial weighting matrix.[32]

[32] Higher R^2 values and lower log-likelihood values suggest that the specifications using the binary contiguity spatial weighting matrix are superior to those using the distance spatial weighting matrix.

Table 5: Absolute Convergence Regressions for British NUTS 3 Sub-regions, 1995-2004

Dependent variable: Average GVA Growth (1995-2004)						
	Spatial Autoregressive Model (SAR)			**Spatial Error Model (SEM)**		
	Secondary	Services	Aggregate	Secondary	Services	Aggregate
constant	0.199	0.016	-0.004	0.215	0.042	0.019
	(0.036)***	(0.024)	(0.037)	(0.037)***	(0.025)*	(0.038)
$\ln GVA_{1995}$	-0.023	0.004	0.004	-0.025	0.002	0.003
	(0.004)***	(0.003)	(0.004)	(0.005)***	(0.003)	(0.004)
ρ (SAR)	0.001	0.001	0.203			
	(0.001)	(0.001)*	(0.083)**			
τ (SEM)				0.003	0.007	0.408
				(0.001)**	(0.001)***	(0.099)***
R^2	0.17	0.04	0.04	0.18	0.12	0.05
Log Likelihood	-243.06	-181.80	394.55	-242.67	-176.23	394.79
Number of Obs	128	128	128	128	128	128

Note: Standard errors are given in parenthesis. Significance at ***1%, **5%, and *10% level

Table 6: Absolute Convergence Regressions for Functional Economic Areas, 1995-2004

Dependent variable: Average GVA Growth (1995-2004)						
	Spatial Autoregressive Model (SAR)			**Spatial Error Model (SEM)**		
	Secondary	Services	Aggregate	Secondary	Services	Aggregate
constant	0.238	-0.079	-0.177	0.259	-0.040	-0.115
	(0.048)***	(0.034)**	(0.078)**	(0.046)***	(0.031)	(0.066)*
$\ln GVA_{1995}$	-0.028	0.015	0.024	-0.030	0.011	0.017
	(0.006)***	(0.004)***	(0.008)***	(0.006)***	(0.004)***	(0.007)**
ρ (SAR)	0.169	0.035	-0.010			
	(0.12)	(0.047)	(0.079)			
τ (SEM)				0.398	0.660	0.460
				(0.120)***	(0.082)***	(0.112)***
R^2	0.23	0.22	0.14	0.27	0.33	0.18
Log Likelihood	194.35	223.05	203.50	195.55	227.34	204.86
Number of Obs	64	64	64	64	64	64

Note: Standard errors are given in parenthesis. Significance at ***1%, **5%, and *10% level

With regard to regional growth convergence, a number of finding emerge from Tables 5 and 6. First, it is clear that there is no absolute convergence in aggregate real GVA per capita growth over the 1995-2004 period. In fact, the functional economic area SEM and SAR specifications indicate divergence in aggregate real GVA per capita growth – a finding supported by Henley (2005) and Monastiriotis (2006). Second, services sector GVA growth does not show signs of convergence. As with the aggregate data, it appears be experiencing a process of divergence when analysed in the functional area context. Finally, secondary sector GVA exhibits strong convergence across all specifications, with an estimated annual speed of convergence of ranging from 2.3-3.1%. This, as suggested in Section 2, may reflect a process of sub-regional secondary GVA per capita being sucked towards the average, due to the sector's near stagnant growth performance over the 1995-2004 period. As for the competing spatial specifications, both yield similar findings but it is the SEM specification which results in higher R^2 values and lower log-likelihood values.

4.3. Conditional β-convergence

The cross-sectional specifications used to test for absolute convergence are now augmented with a set of explanatory variables, which may capture differences in the paths of steady-state GVA per capita. The explanatory variables introduced to the analysis address a number of key features which have emerged from the literature as being influential in the economic growth process. Foremost amongst these are initial education levels and human capital formation, which are necessary to raise productivity.[33] Regarding human capital, this paper follows the approach of Henley (2005) which includes two variables, each capturing distinct aspects of human capital accumulation process: (i) the county average primary school pupil-teacher ratio (*Pupil_Teacher*) and (ii) the average A-level pass rate (*grades*) achieved by pupils in each county. It is this exam which enables pupils to enter university. As 1995 data is unavailable for both of these variables, data dating from 1993 is used instead. As these variables are unavailable at sub-regional level, the data for each county is applied to the sub-region residing in that county. As discussed in Section 2, location and geographic proximity have been identified as key drivers of the British regional growth process – a feature which has been typified by the oft-cited "north-south divide". In order to capture this, a set of dummy variables for the eleven NUTS 1 regions has been constructed. Furthermore, the rural/urban orientation of each sub-region is captured through the inclusion of a variable representing each sub-region's 1995 agricultural real GVA as a proportion of aggregate real GVA *(Agri)*. Unfortunately, data on the capital stock residing in each sub-region at the start of the 1995-2004 period is unavailable. That said, data on the number of businesses registered for Value Added Tax (VAT) is available and is disaggregated for secondary and services sectors. A similar approach is taken by Hart and McGuinness (2003), where the stock of enterprises is used as a proxy for capital utilization. These variables weighted by the population of their relevant sub-region and included in the unconditional convergence specifications *(No. of Businesses)*. A further control variables, females in employment in 1995 expressed as a proportion of people aged 16+ (*Fem Emp'ment*) is included in order capture differences in local labour market conditions (such as the tightness of the labour market) at the beginning of the 1995-2004 period. Capital expenditure for each sub-region *(Capital Expenditure)* in 1995, deflated as described in Section 2 and weighted by population, is also included in the specifications.[34] As in Sub-section 4.2, ρ and τ represent the spatial autocorrelation coefficient and spatial error coefficient, respectively, and the spatial weighting matrix used is the binary contiguity matrix. Appendix 2 reports results for Tables 5-8 when the distance-based spatial weighting matrix is used. Table 7 reports results for the 128 NUTS 3 level sub-regions and Table 8 reports results for the 64 functional economic regions constructed in this paper. The set of NUTS 1 regional dummies are omitted from the specifications in Table 8 as the sub-

[33] See Mankiw et al. (1992) and Barro and Sala-I-Martin (1995, pp. 420-445) for a detailed discussion regarding the inclusion of control and environmental variables in conditional convergence regressions.

[34] Capital expenditure data for the 11 NUTS 3 regions of Wales was unavailable for 1995. As a proxy, the capital expenditure per worker figure for the NUTS 1 region, Wales, is weighted by each NUTS 3 region.

regions which form functional economic regions do not necessarily all belong to the same NUTS 1 region.

The econometric evidence provided in Tables 7 and 8 offers a number of insights into both the tendency (or lack of) towards convergence of sub-regional GVA and the influential factors in sub-regional GVA growth process over the 1995-2004 period. Similar to the absolute convergence case, the results reported in Tables 7 and 8 clearly show that there is no evidence of convergence of aggregate real GVA growth per capita over the 1995-2004 period. The functional economic area regressions of Table 8 even point to divergence in aggregate data over this time period – just as they did in the absolute convergence case. In the case of the services sector, across the specifications there appears to be strong support for the hypothesis that the services sector has also experienced divergence over the 1995-period. A further feature that the unconditional convergence results have in common with their absolute counterparts is the clear secondary sector convergence, but this time with the estimated annual speed of convergence residing within a 2-3.6% range. In all, these finding along with those of the absolute convergence specifications point to a situation where aggregate real GVA per capita has been strong influenced by the tendency towards divergence emanating from the services sector.

The conditional convergence regressions also provide some insights into the factors which have driven these growth trends over the 1995-2004 period. Reflecting its lack of convergence in Tables 5-8, aggregate real GVA growth per capita appears to have been negatively associated with sub-regions whose GVA contains a relatively large agricultural content. What is more, the relatively peripheral NUTS 1 regions of Wales and the North East. The functional economic area regressions of Table 8 also suggests that sub-regions with a higher proportion of female employment enjoyed aggregate GVA growth – indicative of a divergence process where GVA growth becomes increasingly concentrated in those functional economic areas with tighter labour markets. The negative significant coefficient on the number of VAT-registered businesses in Table 8 may reflect the substantial contribution of a relatively small number of large firms to functional economic area GVA per capita growth. Suprisingly, the inclusion of 1995 capital employed per capita in the specifications below does not appear to exhibit a significant impact on GVA growth per capita between sub-regions over the 1995-2004 period. The spatial autocorrelation coefficient does not appear to be significant for aggregate GVA growth within the functional economic zone – an indication, perhaps, that these areas are indeed relatively self-contained.

Table 7: Conditional Convergence Regressions for British NUTS 3 Sub-regions, 1995-2004

Dependent variable: Average GVA Growth (1995-2004)						
	Spatial Autoregressive Model (SAR)			**Spatial Error Model (SEM)**		
	Secondary	Services	Aggregate	Secondary	Services	Aggregate
constant	0.014	0.002	0.003	0.014	0.001	0.003
	(0.327)	(0.009)	(0.010)	(0.015)	(0.009)	(0.010)
$\ln GVA_{1995}$	-0.021	0.006	0.002	-0.020	0.006	0.002
	(0.004)***	(0.002)**	(0.002)	(0.004)***	(0.002)**	(0.003)
Grades	0.001	0.0001	0.0002	0.001	-0.0001	0.0002
	(0.001)	(0.0003)	(0.0004)	(0.001)	(0.0004)	(0.608)
Pupil_Teacher	0.004	0.0001	0.001	0.004	0.0002	0.001
	(0.001)***	(0.001)	(0.001)	(0.001)***	(0.001)	(0.001)
Agri	-0.004	-0.091	-0.118	-0.009	-0.065	-0.111
	(0.050)	(0.026)***	(0.036)***	(0.048)	(0.026)**	(0.036)***
No. of Businesses	1.021	-0.118	0.160	0.369	-0.130	0.161
	(1.051)	(0.174)	(0.112)	(0.974)	(0.168)	(0.110)
Capital Expenditure	0.070	0.039	-0.091	0.054	0.034	-0.079
	(0.161)	(0.096)	(0.105)	(0.158)	(0.714)	(0.449)
Female Emp'ment	0.002	0.0003	0.0001	0.002	0.0002	0.0001
	(0.001)***	(0.0004)	(0.0004)	(0.001)***	(0.0004)	(0.0004)
NE	0.006	-0.007	-0.010	0.004	-0.008	-0.009
	(0.009)	(0.005)	(0.005)*	(0.011)	(0.006)	(0.005)*
NW	-0.001	-0.001	-0.004	-0.005	-0.001	-0.004
	(0.007)	(0.004)	(0.004)	(0.009)	(0.005)	(0.004)
YH	0.010	0.0002	-0.003	0.007	0.003	-0.002
	(0.007)	(0.004)	(0.005)	(0.009)	(0.005)	(0.005)
EM	0.012	0.007	0.0001	0.010	0.009	0.001
	(0.008)	(0.004)*	(0.004)	(0.008)	(0.005)*	(0.005)
WM	0.005	0.003	-0.004	0.002	0.003	-0.004
	(0.007)	(0.004)	(0.004)	(0.008)	(0.004)	(0.004)
EE	0.002	0.002	-0.003	0.002	0.001	-0.003
	(0.006)	(0.004)	(0.439)	(0.007)	(0.004)	(0.004)
L	0.001	-0.006	-0.003	0.002	-0.006	-0.004
	(0.009)	(0.006)	(0.006)	(0.009)	(0.006)	(0.006)
SW	0.008	0.001	-0.001	0.007	0.001	-0.001
	(0.006)	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)
W	-0.0004	0.000	-0.011	-0.002	0.003	-0.010
	(0.007)	(0.004)	(0.004)**	(0.008)	(0.005)	(0.004)**
S	0.016	-0.003	-0.006	0.019	0.001	-0.005
	(0.007)**	(0.004)	(0.004)	(0.010)*	(0.005)	(0.004)
ρ (SAR)	0.363	-0.091	-0.076			
	(0.093)***	(0.027)***	(0.344)			
τ (SEM)				0.520	0.410	0.030
				(0.087)***	(0.097)***	(0.117)
R^2	0.26	0.38	0.31	0.27	0.38	0.30
Log Likelihood	344.96	413.85	401.18	345.41	413.84	400.90
Number of Obs	125	125	125	125	125	125

Note: Standard errors are given in parenthesis. Significance at ***1%, **5%, and *10% level. The NUTS 1 level regional dummy variables included are North East (NE), North West (NW), Yorkshire and the Humber (YH), East Midlands (EM), West Midlands (WM), East England (EE), London (L), South West (SW), Wales (W), and Scotland (S). South East is the base region.

Table 8: Conditional Convergence Regressions for Functional Economic Areas, 1995-2004

Dependent variable: Average GVA Growth (1995-2004)						
	Spatial Autoregressive Model (SAR)			**Spatial Error Model (SEM)**		
	Secondary	Services	Aggregate	Secondary	Services	Aggregate
constant	0.252	-0.040	-0.191	0.256	-0.040	-0.182
	(0.062)***	(0.041)	(0.073)**	(0.059)***	(0.038)	(0.064)***
$\ln GVA_{1995}$	-0.036	0.006	0.024	-0.036	0.008	0.025
	(0.007)***	(0.005)	(0.008)***	(0.007)***	(0.004)*	(0.007)***
Grades	-0.001	0.0001	-0.0003	-0.001	0.0001	0.0001
	(0.0004)*	(0.0002)	(0.0003)	(0.0003)**	(0.0003)	(0.0003)
Pupil_Teacher	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.0005)*	(0.001)	(0.001)	(0.001)	(0.001)
Agri	-0.012	-0.105	-0.113	0.003	-0.085	-0.110
	(0.046)	(0.032)***	(0.033)***	(0.042)	(0.030)**	(0.033)***
No. of businesses	-0.020	0.265	-5.539	-0.475	0.301	-7.373
	(1.213)	(0.361)	(1.524)***	(1.061)	(0.365)	(1.335)***
Capital Expenditure	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female Emp'ment	0.002	0.001	0.001	0.002	0.0002	0.0002
	(0.001)***	(0.001)	(0.0006)*	(0.001)***	(0.001)	(0.001)
ρ (SAR)	-0.097	0.047	0.005			
	(0.120)	(0.043)	(0.066)			
τ (SEM)				-0.417	0.485	0.602
				(0.136)***	(0.109)***	(0.092)***
R^2	0.39	0.41	0.46	0.41	0.44	0.48
Log Likelihood	201.61	232.08	218.34	202.62	233.43	219.43
Number of Obs	64	64	64	64	64	64

Note: Standard errors are given in parenthesis. Significance at ***1%, **5%, and *10% level

The explanatory variables in the services sector regressions also reflect the divergence trends evident in Tables 7 and 8. Again, agricultural output is negatively associated with GVA per capita growth. The East Midlands is the only NUTS 1 region that turns out to be significant, displaying a positive relationship with services GVA growth. The spatial autocorrelation coefficient is negatively significant, while the spatial error term is positively significant – suggesting that bordering a region which enjoys strong services GVA growth does not enhance one's own prospects of services sector growth. The secondary sector explanatory variables cannot be so easily interpreted in a manner consistent with the convergence process identified in Tables 5-8, particularly the positive significant pupil-teacher ratio and female employment coefficients. One hypothesis would be that, while traditional large scale manufacturing has suffered over recent years due to outsourcing and loss of competitiveness in global markets, small-scale manufacturing located relatively close to urban services hubs may have benefited. However, McGuinness and Hart (2003) find that urban locations have a negative impact on the growth of small manufacturing firms due to higher operating and labour costs. That said, McGuinness and Hart (2003) find that government regional assistance positively impacts small manufacturing firm growth through the enhancement of infrastructure networks. The positive pupil-teacher ratio coefficient may actually reflect the positive impact of public sector infrastructure investment from previous years on 1995-2004 secondary sector GVA growth. Government regional assistance may also explain the positive significant coefficient of the Scotland regional dummy. As Barry and Curran (2004) note that, up until the late 1990s, Scotland enjoyed great success in attracting significant FDI inflows in the computer assembly sub-sector, with infrastructure spending and regional preferential assistance being important determinants of new firm location.

5. Conclusions

The objective of this paper is to look beneath the surface of the British sub-regional aggregate GVA growth process experienced over the period 1995-2004, by examining to what extent this process may have been driven by the differing growth dynamics of the secondary and services sectors. As the importance of geographic and locational factors in shaping British sub-regional growth has been widely accepted, spatial econometric methods would appear to be well suited for this task.

From the colour-coded maps of secondary, services, and aggregate real GVA per capita across Britain over the 1995-2004 period, a clear pattern in the spatial dispersion of these sectors is apparent: the secondary sector resides predominantly in the north, while the services sector is very much concentrated in the south –the “north-south divide”. A statistical test for spatial autocorrelation (Moran’s *I*) across the British NUTS 3 sub-regions confirms this spatial dependence. What is more, aggregate real GVA per capita appears to be influenced to a far greater extent by the services sector than by the secondary sector. This is not suprising given the growing proportion of aggregate GVA attributable to the services sector over the 1995-2004, not to mention the sluggish performance of the secondary sector in general over this period. The spatial econometric analysis undertaken in this paper serves a number of purposes: it allows one to (i) test for aggregate real GVA per capita convergence, as well as services and secondary convergence; (ii) characterise spatial influence as a “spatial lag” directly effecting neighbouring regions (SAR) or as an indirect, random spillover effect between regions (SEM); (iii) check the robustness of finding emanating from administrative NUTS 3 level data and with those arising from the use of functional economic areas; (iv) control for the impact of commuter flows, as the functional economic areas are constructed using commuter flow data; (v) test for the robustness of results to different types of spatial weighting matrices; and (vi) incorporate a set of explanatory variables into the analysis which shed light on influential factors in the sub-regional growth process.

The key findings of this paper are robust to the specification of spatial component, the choice of weight matrix, and the delineation of British sub-regions. Aggregate real GVA per capita over the 1995-2004 period exhibits no signs of convergence, either absolute or conditional. The services sector also exhibits no signs of either absolute or conditional convergence. Secondary sector real GVA per capita shows clear signs of both absolute and conditional convergence, with an estimated annual convergence rate of approximately 2-3% depending on the choice of specification. Regarding the aggregate and services sectors, there is strong evidence across the various specifications that these sectors have actually experienced a process of divergence over the 1995 and 2004 period, both in absolute and conditional terms. It is also clear across the specifications that the inclusion of a spatial term is justified and adds to the explanatory powers of those specifications. Furthermore, the insignificance of the spatial autocorrelation coefficient in the functional area specifications suggests that these constructed areas do serve their purpose of approximating self-contained economic areas.

The explanatory variables included in the tests for conditional convergence illustrate the differing forces at play in the various sectoral growth processes. While a more rigorous micro-level analysis is required to fully uncover the drivers of growth in various sectors, a number of insights can be gleaned from this exercise. The total aggregate GVA per capita of sub-regions appears to be influenced negatively by high agricultural GVA content, peripheral location (such as the North East and Wales) and the presence of a large number of VAT-registered businesses (perhaps capturing high levels of capital utilization). The functional economic area approach adds to this in indicating the positive influence of a high proportion of female employment (assumed to capture tight local labour markets), suggesting perhaps a demand-driven move by firms towards the market place.

The secondary sector developments over the 1995-2004 period appear to be driven by other factors: the positive influences of the both the parent-teacher ratio and the Scotland dummy variable suggest that public investment of previous years and regional financial assistance may have influenced secondary sector growth trends. The services sector developments over the 1995-2004 period reflects factors that are in keeping with its lack of convergence: the negative influence of high agricultural GVA content and the positive coefficient of the East Midlands dummy variable could both be seen as signs of a services industry which is slow to move beyond its urban, predominantly southern, stronghold. That said, any such inferences should come with a caveat attached, as the services industry is known to be more heterogeneous in its composition than the secondary industry.

In all, it would appear that analysing aggregate GVA data alone is insufficient to identify the underlying trends in British sub-regional growth. Incorporating a sectoral breakdown of British GVA growth and characterising accurately its spatial dimension offer the potential of richer insights into this growth process.

Appendix 1 British Functional Economic Areas (64)

Hartlepool, Stockton-on-Tees, and South Teeside
Darlington and Durham
Northumberland and Tyneside
Sunderland
West Cumbria
East Cumbria
North and South Manchester, and Cheshire
Lancashire, Blackburn, and Blackpool
Greater Liverpool and Halton and Warrington
Kingston upon Hull and East Riding of Yorkshire
North and North East Lincolnshire
York
North Yorkshire CC
Barnsley Doncaster and Rotherhan, Sheffiled, East Derbyshire
Bradford
Leeds and Calderdale, Kirklees, and Wakefield
Derby, South West Derbyshire
Nottingham, NS Nottinghamshire
Leicester and Leicestershire
Northamptonshire
Lincolnshire
Herefordshire
Worcestershire
Shropshire CC, Telford and Wrekin
Staffordshire and Stoke
Warwickshire, Coventry
Birmingham, Solihull, Dudley, Sandiwell, Wolverhampton and Walsall
Peterborough and Cambridgeshire
Norfolk and Suffolk
Bedfordshire and Luton
Essex, Thurrock, Southend
Greater London , Hertfordshire, Buckinghamshire, Surrey
Berkshire
Milton Keynes
Oxfordshire
East Sussex CC
West Sussex, Brighton and Hove
Hampshire, Portsmouth, Southampton, Isle of Wight
Kent and Medway
North East Somerset, S Gloucestershire, Bristol
Gloucestershire
Swindon
Wiltshire CC
Dorset, Bournemouth and Poole
Somerset
Cornwall and Isles of Scilly
Plymouth
Devon and Torbay
Isle of Angelsey, Gwynedd
South West Wales
Central Valleys
Gwent, Monmouthshire, Newport
Bridgend and Neath Port Talbot
Swansea
Cardiff and Vale of Glamorgan
Conwy, Denbighshire, Flintshire, Wrexham
Powys
Abberdeen, Aberdeenshire, and Angus
Clackmannanshire and Fife
Edinburgh, West Lothian, and Scottish Borders
Falkirk, Perth, Kinross, and Sterling
Dumfries and Galloway
Glasgow, East and West Dumbartonshire, Inverclyde, East Renfrewshire and Renfrew, North and South Lanakshire
Highlands and Islands

Appendix 2: Spatial Regression Analysis Using the Distance-Based Spatial Weighting Matrix

Table A2.1.: Absolute Convergence Regressions for British NUTS 3 Sub-regions, 1995-2004

Dependent variable: Average GVA Growth (1995-2004)						
	Spatial Autoregressive Model (SAR)			**Spatial Error Model (SEM)**		
	Secondary	Services	Aggregate	Secondary	Services	Aggregate
constant	0.186	-0.004	-0.062	0.199	0.011	-0.022
	(0.036)***	(0.025)*	(0.038)*	(0.057)***	(0.037)	(0.049)
$\ln GVA_{1995}$	-0.023	0.005	0.007	-0.023	0.005	0.001
	(0.004)***	(0.003)*	(0.004)*	(0.004)***	(0.003)*	(0.004)*
ρ (SAR)	0.010	0.001	0.969			
	(0.0002)***	(0.0002)***	(0.022)			
τ (SEM)				0.010	0.010	1.032
				(0.0002)***	(0.0002)***	(0.023)***
R^2	0.16	0.01	0.01	0.16	0.01	0.01
Log Likelihood	-242.66	-182.14	393.79	-242.66	-182.14	393.79
Number of Obs	128	128	128	128	128	128

Note: Standard errors are given in parenthesis. Significance at ***1%, **5%, and *10% level

Table A2.1.: Absolute Convergence Regressions for Functional Economic Areas, 1995-2004

Dependent variable: Average GVA Growth (1995-2004)						
	Spatial Autoregressive Model (SAR)			**Spatial Error Model (SEM)**		
	Secondary	Services	Aggregate	Secondary	Services	Aggregate
constant	0.210	-0.137	-0.177	0.222	-0.083	-0.174
	(0.049)***	(0.034)***	(0.079)**	(0.053)***	(0.036)**	(0.070)**
$\ln GVA_{1995}$	-0.026	0.016	0.024	-0.025	0.016	0.024
	(0.006)***	(0.004)***	(0.008)***	(0.006)***	(0.004)***	(0.007)**
ρ (SAR)	0.938	0.938	-0.010			
	(0.045)***	(0.045)***	(0.079)			
τ (SEM)						
				1.067	1.067	1.067
				(0.048)***	(0.048)***	(0.048)***
R^2	0.20	0.19	0.14	0.19	0.18	0.11
Log Likelihood	194.56	223.52	203.50	194.56	223.52	204.20
Number of Obs	64	64	64	64	64	64

Note: Standard errors are given in parenthesis. Significance at ***1%, **5%, and *10% level

Table A2.3.: Conditional Convergence Regressions for British NUTS 3 Sub-regions, 1995-2004

Dependent variable: Average GVA Growth (1995-2004)						
	Spatial Autoregressive Model (SAR)			**Spatial Error Model (SEM)**		
	Secondary	Services	Aggregate	Secondary	Services	Aggregate
constant	0.002	-0.052	-0.037	0.016	0.001	0.003
	(0.015)	(0.009)***	(0.010)***	(0.045)	(0.026)	(0.028)
$\ln GVA_{1995}$	-0.020	0.005	0.002	-0.020	0.005	0.002
	(0.004)***	(0.002)**	(0.400)	(0.004)***	(0.002)**	(0.003)
Grades	0.001	0.0002	0.0002	0.001	0.0002	0.0002
	(0.001)	(0.0004)	(0.0004)	(0.370)	(0.0004)	(0.0004)
Pupil_Teacher	0.004	-0.00003	0.001	0.004	0.00004	0.0007
	(0.001)***	(0.001)	(0.001)	(0.001)***	(0.001)	(0.001)
Agri	-0.022	-0.079	-0.113	-0.022	-0.078	-0.112
	(0.052)	(0.026)***	(0.002)***	(0.051)	(0.026)***	(0.002)***
No. of businesses	1.399	-0.127	0.166	1.388	-0.126	0.165
	(1.07)	(0.176)	(0.112)	(1.067)	(0.175)	(0.111)
Capital Expenditure	-0.002	0.054	-0.079	-0.002	0.053	-0.079
	(0.164)	(0.097)	(0.105)	(0.163)	(0.096)	(0.452)
Female Emp'ment	0.002	0.0003	0.0001	0.002	0.0003	0.0001
	(0.001)**	(0.0004)	(0.811)	(0.001)**	(0.0004)	(0.001)
NE	0.005	-0.006	-0.009	0.005	-0.006	-0.009
	(0.009)	(0.005)	(0.005)*	(0.009)	(0.005)	(0.005)*
NW	-0.002	-0.0002	-0.004	-0.002	-0.0003	-0.004
	(0.007)	(0.004)	(0.004)	(0.007)	(0.004)	(0.004)
YH	0.009	0.0003	-0.003	0.009	0.0003	-0.003
	(0.008)	(0.004)	(0.005)	(0.007)	(0.004)	(0.005)
EM	0.009	0.007	0.001	0.009	0.007	0.001
	(0.008)	(0.004)*	(0.005)	(0.244)	(0.004)*	(0.005)
WM	0.004	0.003	-0.004	0.004	0.003	-0.004
	(0.007)	(0.004)	(0.004)	(0.007)	(0.004)	(0.004)
EE	-0.001	0.002	-0.003	-0.001	0.002	-0.003
	(0.007)	(0.004)	(0.464)	(0.007)	(0.004)	(0.004)
L	0.002	-0.006	-0.004	0.002	-0.006	-0.004
	(0.009)	(0.278)	(0.006)	(0.009)	(0.006)	(0.006)
SW	0.007	0.001	-0.001	0.007	0.001	-0.001
	(0.006)	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)
W	-0.004	0.0003	-0.010	-0.004	0.0003	-0.010
	(0.007)	(0.004)	(0.004)**	(0.007)	(0.004)	(0.004)**
S	0.015	-0.003	-0.005	0.015	-0.003	-0.005
	(0.008)**	(0.004)	(0.004)	(0.008)**	(0.478)	(0.004)
ρ (SAR)	0.968	0.968	0.968			
	(0.023)***	(0.023)***	(0.023)***			
τ (SEM)				1.033	1.033	1.033
				(0.023)***	(0.023)***	(0.023)***
R^2	0.23	0.37	0.29	0.23	0.37	0.29
Log Likelihood	343.83	413.84	401.60	343.83	413.84	401.60
Number of Obs	125	125	125	125	125	125

Note: Standard errors are given in parenthesis. Significance at ***1%, **5%, and *10% level. The NUTS 1 level regional dummy variables included are North East (NE), North West (NW), Yorkshire and the Humber (YH), East Midlands (EM), West Midlands (WM), East England (EE), London (L), South West (SW), Wales (W), and Scotland (S). South East is the base region.

Table A2.4.: Conditional Convergence Regressions for Functional Economic Areas, 1995-2004

Dependent variable: Average GVA Growth (1995-2004)						
	Spatial Autoregressive Model (SAR)			**Spatial Error Model (SEM)**		
	Secondary	Services	Aggregate	Secondary	Services	Aggregate
constant	0.250	-0.096	-0.235	0.261	-0.042	-0.192
	(0.063)***	(0.042)**	(0.072)***	(0.065)***	(0.043)	(0.073)**
$\ln GVA_{1995}$	-0.038	0.007	0.025	-0.037	0.007	0.024
	(0.007)***	(0.005)	(0.008)***	(0.007)***	(0.005)	(0.008)***
Grades	-0.001	0.000	-0.0003	-0.001	0.000	-0.0003
	(0.0004)	(0.000)	(0.0003)	(0.0003)	(0.000)	(0.0003)
Pupil_Teacher	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Agri	-0.017	-0.107	-0.114	-0.016	-0.105	-0.113
	(0.046)	(0.033)***	(0.033)***	(0.046)	(0.032)***	(0.032)***
No. of businesses	-0.030	0.307	-5.596	-0.030	0.303	-5.514
	(1.229)	(0.368)	(1.534)***	(1.211)	(0.363)	(1.512)***
Capital Expenditure	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female Emp'ment	0.002	0.001	0.001	0.002	0.001	0.001
	(0.001)**	(0.001)	(0.0005)*	(0.001)**	(0.001)	(0.001)*
ρ (SAR)	0.938	0.938	0.938			
	(0.044)***	(0.045)***	(0.045)***			
τ (SEM)				1.067	1.067	1.067
				(0.065)***	(0.048)***	(0.048)***
R^2	0.37	0.38	0.44	0.37	0.38	0.44
Log Likelihood	202.49	232.27	219.05	202.48	232.27	219.05
Number of Obs	64	64	64	64	64	64

Note: Standard errors are given in parenthesis. Significance at ***1%, **5%, and *10% level.

References

Andersen, A.K. (2002) Are Commuting Areas Relevant for the Delimitation of Administrative Regions in Denmark? *Regional Studies* 36(8), 833-844.

Anselin L. (1988) *Spatial Econometrics: Methods and Models*, Kluwert Academic.

Barro R. J. (1991) Economic Growth in a Cross-Section of Countries, *Quarterly Journal of Economics* 105, 407-443.

Barry F. and D. Curran (2004) Enlargement and the European Geography of the Information Technology Sector, *The World Econom*y 27(6) 901:922

Baumol WJ (1986) Productivity Growth, Convergence, and Welfare: What the Long-run Data Show, *American Economic Review* 76: 1072-1085.

Bernat A. (1996) Does Manufacturing Matter? A Spatial Econometric View of Kaldor's Laws, *Journal of Regional Science* 36: 463-477.

Bishop P. and P. Gripaios (2004), Earnings Biases and Convergence in the UK: A County Level Analysis, *Applied Economics Letters* 11(1), 33-37.

Bishop P. and P. Gripaios (2006) Earnings Convergence in GB Counties: A Distribution Dynamics Approach, *Applied Economics Letters* 13, 29-33.

Boddy M., Hudson, J., Plumridge A., and D.Webber (2005) Regional Productivity Differentials: Explaining the Gap, No 515, Discussion Papers from University of the West of England, School of Economics http://econpapers.repec.org/paper/uwewpaper/0515.htm

Chatterji, M. and J. H. L Dewhurst (1996) Convergence Clubs and Relative Economics Performance in Great Britain: 1977–1991, *Regional Studies*, 30, 1996, pp. 31–40.

Casado-Diaz J. M. (2000) Local Labour Market Areas in Spain: a Case Study, Regional Studies 34(9), 843–56.

Coombes M. G., Green A. E. and S. Openshaw (1986) An Efficient Algorithm to Generate Official Statistical Reporting Areas: the Case of the 1984 Travel-to-Work Areas Revision in Britain, *Journal of Operational Research Society*; 37, 943–53.

Dewhurst J. (1998) Convergence and Divergence in Regional Household Incomes per Head in the United Kingdom, 1984-93, *Applied Economics* 30, 31-35.

Eurostat (1992) Etude sur les Zones d'Emploi, Document E/LOC/20. Office for Official Publications of the European Communities, Luxembourg.

Fingleton B. and E. Lopez-Baso (2006) Empirical Growth Models with Spatial Effects, *Papers in Regional Economic Science* 85(2)

Henley, A. (2005) On Regional Growth Convergence in Great Britain, *Regional Studies* 39, 1245-1260

Karlsson C. and Olsson M. (2006) The Identification of Functional Regions: Theory, Methods, and Applications, *Annual of Regional Science* 40: 1-18

Mankiw G, Romer D. and D. Weil (1992) A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics* 107: 407-437.

McGuinness S. and M. Hart (2003) Small Firm Growth in the UK Regions 1994-1997: Towards an Explanatory Framework, *Regional Studies* 37(2): 109-122

Monastiriotis, V. (2006) Sub-regional Disparities in Britain: Convergence, Asymmetries and Spatial Dependence", Research Papers in Environmental and Spatial Analysis No. 12, European Institute, London School of Economics.

Openshaw, S. (1984). *The Modifiable Areal Unit Problem*. Norwich

Openshaw, S., and Alvanides, S. (1999) Applying Geocomputation to the Analysis of Spatial Distributions, in P.A Longley, M.F. Goodchild, D.J. Maguire and D.W. Rhind, (eds) (1999) *Geographical Information Systems, Proinciples, Techniques, Management and Applications*, New York: John Wiley.

Rey S., and B. Montouri (1999) U.S. Regional Income Convergence: A Spatial Econometric Perspective, *Regional Studies* 33: 143-156.

Drifting Together Or Falling Apart? The empirics of Regional Economic Growth in Post-Unification Germany

Roberta Colavecchio, Declan Curran, and Michael Funke

1. Introduction

When, on October 3rd 1990, the 60 million Germans in the West were formally re-united with the 16 million Germans in the East, the two parts could hardly have been more different. Despite a common culture and language, after forty years of development with radically different economic institutions and incentives, the Federal Republic of Germany and the German Democratic Republic (GDR) were characterized by substantial disparities in physical and human capital, labour productivity, incomes and wealth. According to Sinn and Sinn (1992), GDP per person in East Germany in 1989 was only 60 percent of the West German level. The West was one of the technologically most advanced and richest countries in the world; the East was economically shattered after four decades of communism and nearly bankrupt. In the years leading up to unification, real GDP growth was steady in the former West Germany and the unemployment rate was stable. After unification, the western states experienced sharper business cycle fluctuations: a modest upturn in 1990-91 was followed by a sharp recession in 1992-93, both of which were mainly due to the unification process. The initial economic boom was led by "exports" to the eastern states, where consumers were switching to cheaper and better quality goods produced in the West. Moreover, the German government financed its initial transfers to the GDR by borrowing, a choice which stimulated an economy already near its output potential and triggered a widening fiscal deficit. The subsequent recession was also closely related to unification. Restrictive measures were implemented to reduce the fiscal deficit and the Bundesbank tightened monetary policy to cap the rising inflation. These policy responses, coupled with a contraction in foreign demand, had a dampening effect on the economy and the post-unification boom gradually turned into a deep recession, with *GDP* growth rates well below the historical average for western Germany. Indeed, the major cost of unification for western regions in the years immediately following unification was a lower income growth rate and thus the issue of convergence of *GDP* growth rates among German regions in the last decade is still an open question. In particular, the problem of uneven regional developments has been closely monitored in economic policy debates and in recent years there has been a surge in empirical work on growth and convergence.

When considering regional convergence, various empirical approaches have been implemented in the literature: from simple plots of measures of dispersion over time to intra-distributional dynamics using Markov chains applied to *GDP* per capita. Numerous studies have revealed persistent differences in per capita income among regions. Evidence shows that some regions managed to sustain high per capita income over a long time span while other regions seemed to be trapped in a low income growth path. These persistent differences are strikingly at odds with the standard neoclassical growth model, which predicts that poorer countries usually develop faster than richer ones and that there is a tendency toward convergence in levels of *GDP* per capita. A key feature of the neoclassical growth model has been the assumption of identical production functions for all regions. As a consequence, a single dynamic model is adequate to characterise all cross-region growth behaviour. On the other hand there exists an opposing growth paradigm [see, for example, Azariadis and Drazen (1990)] explaining multiple steady states in the growth rate of per capita income. According to Azariadis and Drazen (1990) and Aghion and Howitt (1998, chapter 10), multiple locally stable equilibria can be attributed to differences in initial conditions. Faini (1984) has initially considered multiple steady states in the context of regional development issues.[35] In all these models, different initial conditions may cause regions to get stuck at different self-perpetuating levels of economic activity. As suggested by Quah (1996, 1997) and Paap and van Dijk (1998), this may lead to a polarisation into clubs of rich and poor countries or regions.[36]

Research on convergence has accommodated cross-regional heterogeneity in a sequence of stages. At first, conventional cross-section analysis [see, for example, Barro (1991) and Mankiw et al. (1992)] assumed complete homogeneity in steady state growth rates. Recently, Lee et al. (1997, 1998) allowed complete heterogeneity in steady state growth rates. However, as pointed out by Islam (1998), extensions that allow varying growth rates run the risk of robbing the concept of convergence of any economic meaning. Instead of assuming complete heterogeneity, we set a structure of an intermediate form: we advocate techniques which focus on the evolution of the entire cross-sectional distribution in addressing the question of convergence across German districts in the first decade after German unification.[37] In this context, a convergence process occurs if, for instance, a bimodal density is detected at the beginning of the sample period and over time there is a tendency in the distribution to move towards unimodality. Alternatively, if there already is a unimodal distribution after German unification, convergence occurs when the dispersion of this density and therefore per capita income declines over time.[38] To the best of our knowledge, no papers have attempted to formally test the

[35] Nelson (1956) is the grandfather of low-level equilibrium trap models.

[36] The obvious difficulty here is to figure out in the data which countries are in the bad and which ones are in the good equilibrium. Barrier to getting out of such a trap can be the lack of a "big push" [see Murphy et al. (1989)]. Rodrik (1996) has argued that the East Asian miracle may have depended on a state-assisted process of overcoming coordination failure, and a consequent shift between two different equilibrium output levels (or a virtuous circle). It is also worth noting that the possibility of non-uniqueness is discussed informally even in Solow´s (1956) original exposition of the neoclassical growth model.

[37] In this paper we add to the contributions of Bianchi (1997), Corrado et al. (2005), Jüßen (2005), Lopez-Bazo et al. (1999) and Pittau (2005) testing for "two-club" or "twin-peak" convergence of GDP per capita across countries and EU regions by analysing data which do not overlap with the data of existing papers. Magrini (2004), p. 2744) maintains that "[…] the distributional approach to convergence – particularly when based on nonparametric kernel estimations – appears to be generally more informative than convergence empirics within the regression approach, and therefore represents a more promising way forward".

[38] Economic and social cohesion is embedded in the German constitution („Verfassung"). Regional income inequalities are therefore a major concern for policy-makers and substantial fiscal transfers are offered to less developed regions which aim at reducing undesired income disparities across regions.

convergence club hypothesis across East and West German regions after unification.[39] It is our purpose to detect whether clubs exist and which regions are associated with which clubs. A natural approach to assess the evolution over time of the dispersion of the regional per capita income is to estimate the cross-section distributions by using kernel density estimation.

The remainder of the paper is structured as follows. Section 2 describes the data set used for this study together with the non-parametric estimates of the per capita regional *GDP* over time. To support the visual impression given by kernel density estimates, and to provide further insight on the features of the underlying density, we have performed several statistical tests, whose results are presented in section 3 and 4. In particular, section 3 reports the outcome of a non-parametric test for multimodality, along with a set of maps which provide an illustration of the spatial structure of the real *GDP* per capita across German districts. Section 4 presents the results of a parametric test of density time invariance. Suggestions for future research appear in Section 5 and Section 6 concludes.

2. Data issues and empirical evidence from non-parametric density estimation

The opportunity to assess spatial disparity trends in per capita income indicators is limited by the availability of consistent and comparable data. Long and dense time series for small geographic units are difficult to obtain, and in many cases not existent. In this section, we briefly present the spatial distribution of our data which are at the heart of our analysis. There are three levels of administration in Germany: (1) the Federal Republic at the national level; (2) 16 federal states (*Bundesländer*) on the regional level and (3) 439 districts (*Kreise*) or towns with autonomous administration (*kreisfreie Städte*), both on the local level. Smaller municipalities belong to the districts. In our empirical work below we focus on these 439 districts covering the entire economy.[40] Our data run from 1992 to 2001; 2001 is the latest year available to us and data prior to German unification are not available. Data for 1993 are missing. The source of our data is the "Arbeitskreis Volkswirtschaftliche Gesamtrechnung der Länder".[41] The *GDP* per capita data are at constant 1995 prices and are obtained dividing the *GDP* of the German districts by their population. Ideally, we should deflate district-level per capita incomes using district-level deflators but, since district-level price indices are not available, we follow the usual practice and simply use the 16 state-level GDP deflators.

[39] Funke and Niebuhr (2005) have demonstrated the existence of two clubs across West German regions prior to unification using threshold estimation techniques.

[40] We focus on district-level data because state-level data tend to „aggregate away" important differences between smaller geographic entities within the 16 states. For example, in the dataset that we analyse below, the ratio of GDP per capita between the richest (*Hamburg*) and the poorest state (*Sachsen-Anhalt*) was 2.63 in 2001, while the corresponding ratio for the richest (*Landkreis München*) and the poorest district (*Mittlerer Erzgebirgskreis*) was 7.30. On the other hand, one has to be aware that district-level GDP per capita figures may be affected by a commuting bias. Especially, commuters could overstate GDP per capita in agglomerations and city regions. Hamburg and Berlin are classified as a single region. This was forced on us because of lack of district-level data for both states. We also run the Kernel estimates excluding Berlin and Hamburg. Qualitatively, results are unchanged and the pattern is not much affected. Brakman et al. (2004) have analysed the spatial distribution of wages across Germany using district-level data.

[41] See http://www.statistik-bw.de/Arbeitskreis_VGR/publikationen.asp.

Nonparametric density estimations can reveal several features of the data and therefore help to capture the stylised facts that need explanation, exploring which specifications match with the data. The kernel estimator for the density function $f(x)$ at point x is

(1) $$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right)$$

where $x = x_1, x_2, \ldots, x_n$, is an independent and identically distributed sample of random variables from a probability density $f(x)$ and $K(\cdot)$ is the standard normal kernel with window width h. The window width essentially controls the degree to which the data are smoothed to produce the kernel estimate. The larger the value of h, the smoother the kernel distribution. A crucial issue is the selection of this smoothing parameter. In order to solve the trade-off between oversmoothing and undersmoothing, i.e. the trade-off between bias and variance, we have first used Silverman´s (1986) "first generation" rule-of-thumb for a Gaussian kernel.[42] Additionally, we consider the two-stage direct plug-in bandwidth selection method of Sheather and Jones (1991), which has been shown to perform quite well for many density types by Park and Turlach (1992) and Wand and Jones (1995).[43]

The distributions have been fitted to the logarithm of real per capita income. In Figures 1 and 2 are plotted the kernel density estimations for the (log) *GDP* from 1992 to 2001 obtained using the two abovementioned bandwidth selection methods and by transforming the income variable to the original scale.[44] The figures show similar patterns, validating the fact that the estimates are robust with respect to the bandwidth specification. Nevertheless, as expected, the Silverman (1986) rule of thumb returns a slightly larger optimal smoothing parameter and therefore the relative density estimate (Figure 1) appears oversmoothed compared to the one obtained used the Sheather and Jones (1991) plug-in method (Figure 2).

[42] The properties of this rule may be seen in Silverman (1986), pp. 45-48. In the estimates below we have used the modified "Silverman´s rule of thumb", as in (3.31), p. 48 of Silverman (1986).

[43] Our primary objective was to choose a bandwidth selection procedure that performs well for heavy tailed densities. Another concern is that some selectors have excellent asymptotic properties but very poor performance with small samples. This is why it is important to look more deeply in comparing alternative bandwidth selectors.

[44] Before plunging into the calculations, it is worthwhile to stress the limits of the purely statistical devices. They are useful in identifying interesting patterns and regularities but, by their very nature, do not uncover the ultimate reasons why some districts are much richer than others, even at a modest growth accounting level. Instead, it should be understood as a diagnostic tool - just as medical tests can tell one whether or not he is suffering from a certain ailment but cannot reveal the causes of it. This, of course, does not make the test any less useful.

Figure 1: Non-parametric densities of per capita GDP (constant 1995 prices) across German districts using Silverman´s (1986) "first generation" rule-of-thumb

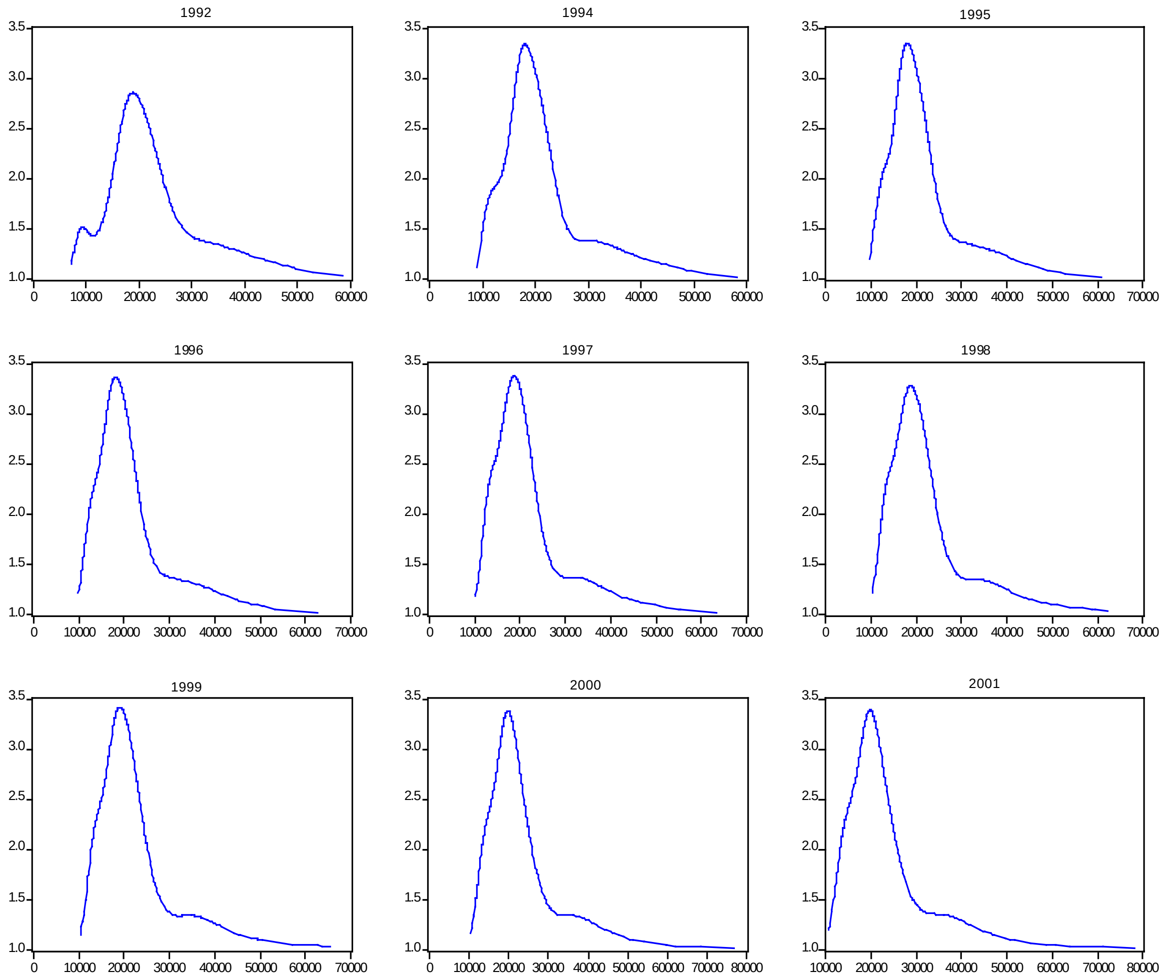
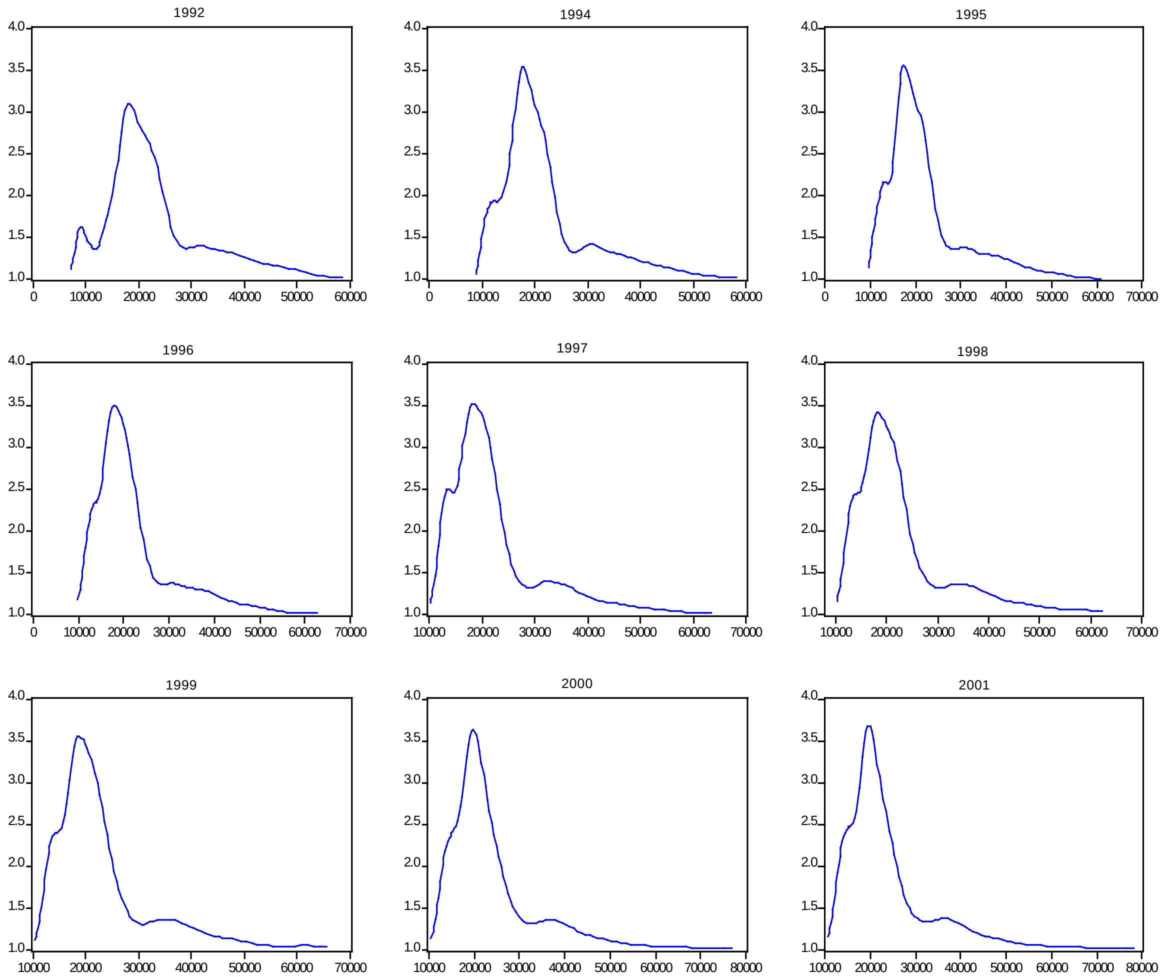


Figure 2: Non-parametric densities of per capita GDP (constant 1995 prices) across German districts using the plug-in bandwidth selection method of Sheather and Jones (1991)



A preliminary inspection of the estimated densities reveals several noteworthy aspects. First, the snapshots show pronounced triple peakedness at the beginning of the considered time span. This evidence indicates that the German districts in 1992 can be separated into three groups, poor, rich and middle income. Second, as time passes this triple peakedness becomes less visible as the mode corresponding to low-income level recedes somewhat, without disappearing entirely as Figure 1 would have us believe.[45] As we will see, this bimodal/trimodal ambiguity recurs later when we utilize statistical tests for multimodality. Either way, this smoothing of the third mode is indicative of an improvement in economic conditions of the German poorest districts. In particular, this smoothing of the initial trimodality supports the notion of a catching-up process of eastern Germany at the beginning of the 1990s, i.e. the poorest districts did not stay as poor as they were immediately after unification. That said, despite the tendency of initially poor units to increase relative incomes, on

[45] A "mode" is meant here to be a point on the empirical density estimate around which the tangent to the curve changes its slope from positive to negative.

average, over the considered decade, several districts experienced negative growth rates.[46] Third, there is a visible tendency for the remaining two peaks to move apart, with the third mode moving to the right towards a higher income level. Moreover, the variability of the "low-mean distribution" has been declining over the decade from 1992 to 2001 and in 2001 appears to be considerably smaller than the spread of the "high-mean distribution". This evidence reveals that cross regional income disparity has become larger rather than smaller as predicted by absolute convergence.

Furthermore, we use the methodology of distributional dynamics to model the evolution of the relative distribution of per capita incomes for Germany districts. This approach models directly the evolution of relative income distributions by constructing transition probability matrices that track changes over time in the relative position of districts within the distribution. This is an exercise that a number of authors have undertaken (see Quah, 1996a, 1996b, 1997, 2006). The modelling of distribution dynamics assumes that the density distribution ϕ_t has evolved in accordance with the following equation:

(2) $$\phi_{t+1} = M\,\phi_t,$$

where M is an operator that maps the transition between the income distributions for the periods t and t+1. Since the density distribution ϕ for the period t only depends on the density ϕ for the immediately previous period, this is a first-order Markov process.[47] In our estimates below we have assumed that the distribution ϕ has a finite number of states. For the Markov transition matrices we assume that the probability of variable s_t taking on a particular value j depends only on its past value s_{t-1} according to the first-order Markov chain

(3) $$P\{s_t = j \mid s_{t-1} = i\} = P_{ij},$$

where P_{ij} indicates the probability that state i will be followed by state j. As

(4) $$P_{i1} + P_{i2} + \ldots + P_{in} = 1$$

we may construct the so-called transition matrix

[46] In particular, the growth rates of the real GDP per capita over the decade from 1992 to 2001 were negative in 66 districts. Out of these 66, seven districts (Delmenhorst, Landkreis Holzminden, Landkreis Sigmaringen, Landkreis Soltau-Fallingborstel, Landkreis Unterallgäu, Neustadt an der Weinstrasse, Wilhelmshaven) have even experienced two-digit negative growth rates. Following Jones (1998, p. 4) these districts might be labeled “growth disasters”.

[47] Equation (2) may be seen as analogous to a first-order autoregression in which we replace point estimates by complete distributions.

(5) $$P = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ P_{n1} & P_{n1} & \cdots & P_{nn} \end{bmatrix}$$

where line i and column j give the probability that state i will be followed by state j. In our modelling approach, the probability P_{ij} measures the proportion of districts in regime i during the previous period that migrate to regime j in the current period. According to Geweke et al. (1986), the maximum likelihood estimator for the transition probability $\hat{p}_{ij}$ is given by:

(6) $$\hat{P}_{ij} = \frac{\sum n_{ij}}{\sum n_i}$$

where $\sum n_{ij}$ is the number of districts that were in income category i in the previous period and have migrated to income category j in the current period, and $\sum n_i$ is the total of districts that were is income category i in the previous period. The main advantage of the transition matrix is that it allows to summarise the random ups and down of regional fortunes in a handful of numbers.

The transition probability matrix in Table 1 reports transitions between the 1992 and 2001 distributions of *GDP* per capita relative to the German average.[48] The main diagonal of the matrix gives the proportion of districts that were in the same range of the distribution immediately after German unification as a decade later. Table 1 also provides information about n, the number of districts that begin their transitions in a given state. Furthermore, we provide the classes that divide up the state space.

The salient characteristics of the transition probability matrix in Table 1 reveal a number of noteworthy behavioural patterns in the distribution of real GDP over time. First, as indicated by the first element of the main diagonal (0.03), districts which originally reside in the lowest range of the distribution (i.e. with a GDP per capital of 50% or less of the German average) appear to be very unlikely to remain in this category at the end of the period in question. Such districts display a strong tendency to either move forward to the second category (0.68) or jump to the third category (0.27). Second, the third and fourth elements of the main diagonal (a real GDP of 65%-80% and 80%-100% of the German average, respectively) indicate a relatively high probability for the regions within this range to maintain their status quo over the period. That said, regions in the third category appear to be relatively open to backward or forward movements (with probabilities of 0.13 and 0.22 respectively) while those in the fourth seem decidedly more backward looking, as illustrated by the 0.26 probability of moving a step back but only a mere 0.04 probability of moving forward one step. Finally, the

[48] Only districts that were part of the dataset at the beginning of the sample period are included in the calculation.

districts residing in the fifth category (with a real GDP of 100-125% of the average) appear to be more likely to either retain this position or fall back by one category. These districts marked inability to move forward (a probability of 0.02) suggests there comes a point where incremental increases in real *GDP* become harder and harder to sustain. Furthermore, those districts that reside in the highest income category at the beginning of the time period display a very high probability (0.83) of consolidating their position of affluence.

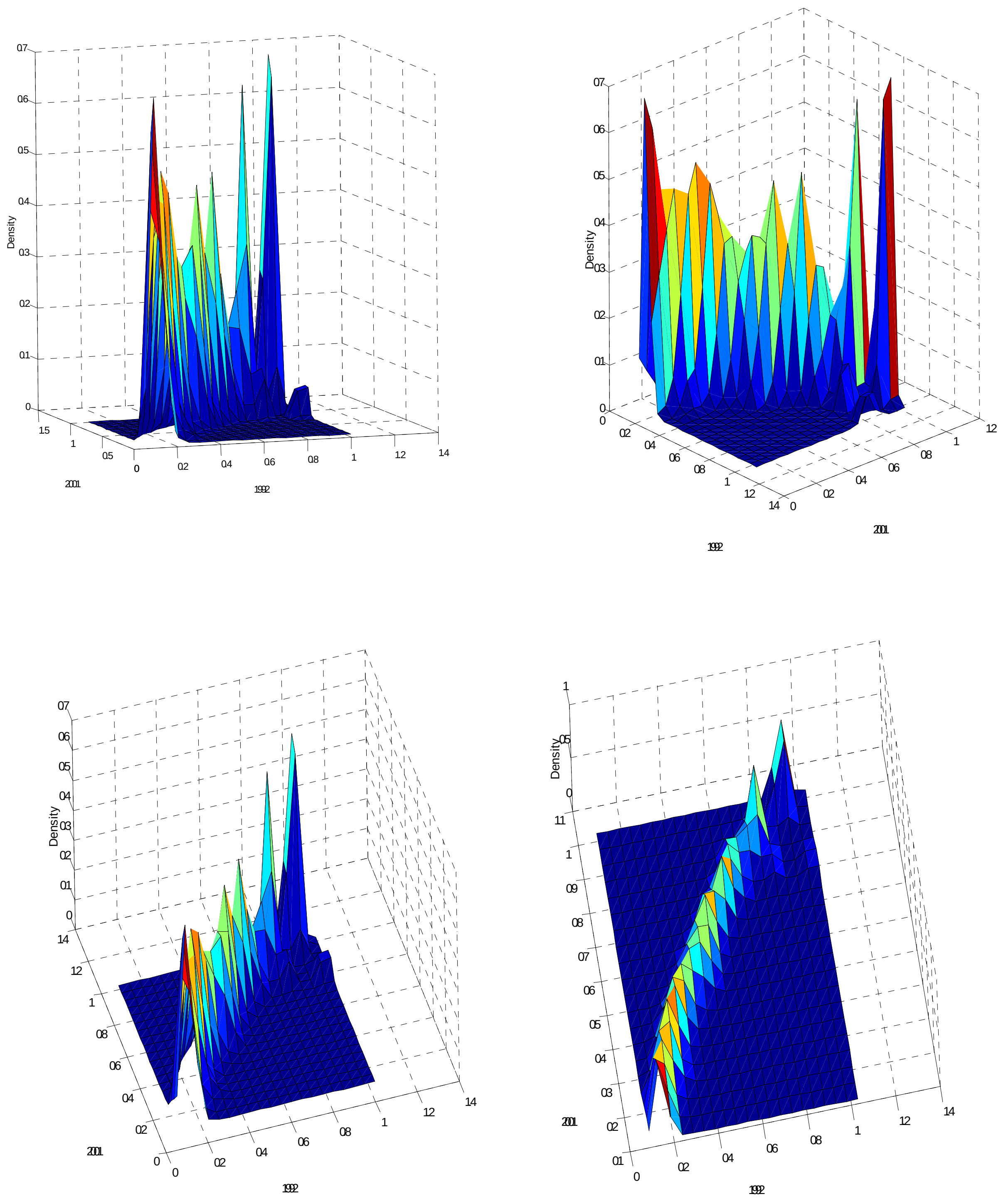
Table 1: Transition Probability Matrix Relative to the German Average

	GDP PER CAPITA 2001							
	n		*4*	*61*	*79*	*128*	*74*	*73*
GDP PER CAPITA 1992	*63*	[0-0.5]	0.03	0.68	0.27	0.02	0.00	0.00
	30	[0.5-0.65]	0.07	0.40	0.33	0.17	0.03	0.00
	45	[0.65-0.8]	0.00	0.13	0.53	0.22	0.11	0.00
	106	[0.8-1.00]	0.00	0.00	0.26	0.70	0.04	0.00
	90	[1.00-1.25]	0.00	0.00	0.00	0.41	0.57	0.02
	85	[1.25-∞]	0.00	0.00	0.00	0.01	0.15	0.83
			[0-0.5]	[0.5-0.65]	[0.65-0.8]	[0.8-1.0]	[1.0-1.25]	[1.25-∞]

The forementioned characteristics support the findings of kernel density estimation, namely: the tendency of the poorest districts to catch-up; the middle income districts retaining their status quo (despite a small number of their ranks back-peddling); and the consolidation of the richest districts of their position.

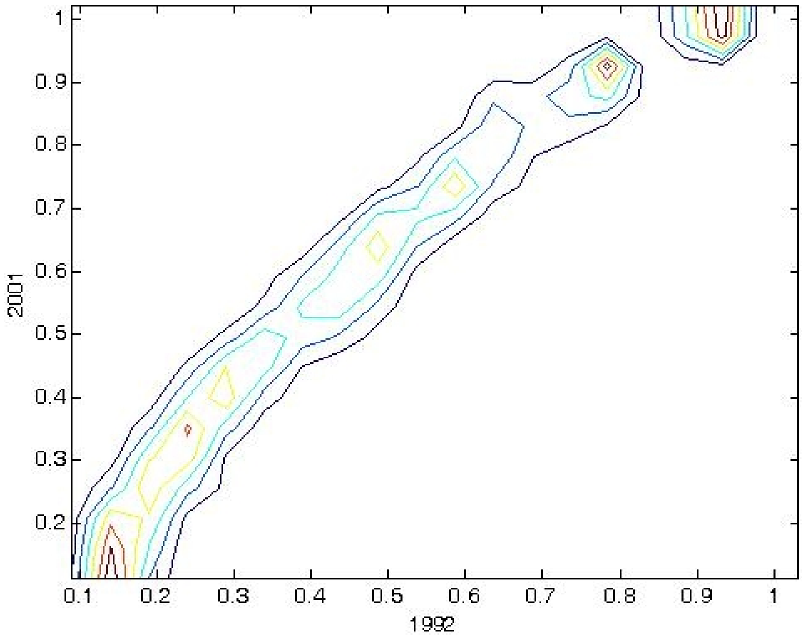
In Table 1, the operator *M* has been constructed by assuming that the distribution ϕ_t has a finite number of states. This discrete modelling approach leads to the problem that the researcher has to determine the number of intervals and the limit values of each interval in an arbitrary and ad hoc way. Furthermore, the discretisation process may eliminate the property of Markovian dependence in the data, as Bulli (2001) has pointed out. The solution addressing these shortcomings consists of carrying out a continuous analysis of transition, which avoids discretisation through the use of conditional densities that are estimated non-parametrically and referred to as stochastic kernels. A stochastic kernel amounts to a transition matrix with an infinite number of infinitely small ranges. The results from this tool are displayed as three-dimensional graphs in Figure 3 and a two-dimensional contour map in Figure 4.

Figure 3: Stochastic Kernel Estimates, 1992 - 2001



Note: In Figure 3 and 4 we have used the region with the highest per capita income as a numeraire. The choice is arbitrary but has no impact upon the qualitative results.

Figure 4: Stochastic Kernel Contours



The three dimensional stochastic kernel graph yields a number of valuable insights, which are both additional and complementary to those of the static kernel illustrations of Figures 1 and 2. In order to fully exploit the information content of this construct we firstly adjust the viewer's perspective by rotating the illustration (Figure 3, top left and top right). We then provide further insights by tilting the graph downwards, as if looking down on the three dimensional distribution from above. This "aerial view" is further enhanced by means of contour images of the distribution (Figure 4). Rotating the graphs in Figure 3 (top right and top left) highlights two features: the pronounced peaks at the beginning and end of the distribution; and middle section of the distribution which, while relatively lower, still suggests the possibility of either slippage or enhancement of one's relative position.

This aerial view of the income distribution, highlighting as it does the diagonal pattern of the distribution over time, illustrates the tendency of regions residing in low income categories in 1992 to remain there in 2001, while high income regions retain their affluent status throughout the period in question. That said, a further more subtle nuance can be gleaned from Figures 3 and 4. The hint of concavity visible in both the three dimensional graph (bottom right) and contour representation are indicative of upward movement in the status of the lower income regions; a finding which is explored further below.

However, the discussion above has relied heavily on the visual impression and shape of the non-parametric income densities. In practical terms, looking at Figure 1 - 4, the question to ask is: are those districts randomly drawn from an unimodal distribution, a bimodal distribution or is there any kind of trimodality? In order to shed further light on the structure of the underlying density function and in particular about the number of unobserved modes, in the following two sections we perform several statistical tests connected to kernel density estimates.

3. Tests for multimodality

The aim of this section is to give a short overview of the statistical method applied in this paper to test for convergence clubs.

In order to assess the issue of multimodality, we first calculate the Timm (2002) bimodality index (*BM*) for the real GDP per capita for each year of the considered time sample. The BM index is defined by:

$$(7) \qquad BM = \frac{m_3^2 + 1}{m_4 + \frac{3(n-1)^2}{(n-2)(n-3)}},$$

where m_3 is the skewness coefficient, m_4 is the kurtosis coefficient, and n is the number of observations. Values bigger than 0.55 indicate the existence of bimodal or multimodal distributions. The results are reported in in Table 2. The values of the index are all above the critical threshold of 0.55, which indicates the presence of multimodality in the distribution of the considered variable.

Having confirmed the existence of multimodality, we would now like to ascertain the actual number of modes present in our estimated density functions. Silverman (1981, 1986) has emphasised the proper modelling of the number of modes and has presented a test for multimodality and peakedness.[49] The test may help interpreting the evolution of regional inequalities across German districts.

Table 2: Timm´s (2002) Bimodality Index

YEAR	TEST STATISTIC
1992	1.23
1994	1.11
1995	1.23
1996	1.29
1997	1.38
1998	1.42
1999	1.56
2000	1.68
2001	1.76

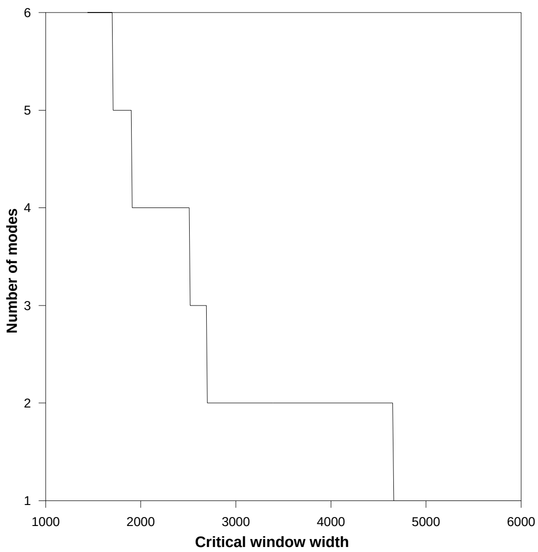
The non-parametric procedure tests the null hypothesis that a density *f* has *k* modes (or peaks, bumps) against the alternative that *f* has more than *k* modes, where *k* is a non-negative integer. The test statistic in this case is the critical window width, defined by

[49] This test of multimodality has been used by Bianchi (1997) to test the hypothesis of income convergence for a group of 119 countries between the years of 1970 and 1989. Bianchi (1997) rejects the hypothesis of convergence in favour of the formation of convergence clubs.

(8) $$h_{crit}(k) = \inf\{h \mid \hat{f} \text{ has at most } k \text{ modes}\}.$$

For $h < h_{crit}(k)$, the estimated density has at least k+1 modes. The basic idea of the test is intuitive and simple. Specifically, if the series has k modes, then $h_{crit}(k\text{-}1)$ should be 'large' because substantial smoothing is required to generate a (k-1)-mode density. For example, if the data possess two strong modes, a large value of h will be needed to obtain an unimodal estimate. An illustrative calculation of the critical window widths h and the corresponding number of modes (peaks) in the kernel density estimates for the year 1999 is plotted in Figure 5.

Figure 5: Number of Modes in the Kernel Density Estimate as a Function of the Window Width Size *h*, 1999



Thus, the technique forms a natural hypothesis-testing framework since large numbers of $h_{crit}(k)$ indicate more than k modes. The crucial question, then, becomes how large is "large" when the chosen bandwidth is concerned. The value of $h_{crit}(k)$ is computed through a binary search algorithm, and its significance level can be assessed by the bootstrap procedure attributable to Efron (1979). In particular, the bootstrap test requires a statistic test $t(x)$ and an estimated null distribution for the data under H_0. Given these, the p-value of the test is

(9) $$p_{\text{bootstrap}} = \text{prob}_{\hat{F}_0}\{t(x^*) > t(x)\}$$

where

(10) $$x^* = \left(x_1^*, x_2^*, \ldots, x_n^*\right)'$$

is the bootstrap drawn from the null distribution $\hat{F}_0$. To approximate $p_{\text{bootstrap}}$, bootstrap samples have to be drawn from a rescaled density estimate obtained by setting

$$x_i^* = \bar{y}^* + \sqrt{1 + \frac{h^2}{\hat{\sigma}^2}}\left(y_i^* - \bar{y}^* + h\varepsilon\right), \tag{11}$$

where $\sqrt{1 + h^2/\hat{\sigma}^2}$ is the rescaling factor, y_i^* are sampled with replacement from the original sample, $\bar{y}^*$ its mean, $\hat{\sigma}^2$ its variance and ε is assumed to be distributed as a standard normal since the kernel is Gaussian.[50] In the spirit of the analysis of Hall (1992), the bootstrap method treats the available sample as the population, and through repeated re-sampling of this sample, obtains the distribution of statistics of the test.[51] A sample is taken of the original series (with replacement) and transformed to have the same first and second moments. Critical values are then obtained by generating a large number of samples.[52] This is not a nested test and its results should therefore be interpreted as a hierarchical set of significance tests.[53]

We execute the Silverman (1981, 1986) test for each year, with null hypotheses of one, two and three modes (hence alternative hypotheses of more than one, more than two and more than three modes).

[50] Rescaling is necessary since the kernel estimation artificially increases the variance of the estimate [see Efron and Tibshirani (1993)]. Since the procedure samples from a smooth estimate of the population, it is called smooth bootstrap.

[51] Strictly speaking, the data can only be resampled with replacement if they are i.i.d. If there is (spatial) dependence, the bootstrap procedure needs to be modified to accommodate dependence. However, note that, while resampling blocks generally increase the efficiency of the bootstrap in this case, the available evidence on this issue indicates that the efficiency gains are very small [see, e.g., Hall et al. (1995)].

[52] In our simulations we set the number of bootstrap replications to 3000. It is well-known [see Izenman and Sommer (1988) and Hall and York (2001) for a detailed account] that the Silverman test tends to suffer from low power and accordingly probability values higher than conventional ones are typically used. One therefore has to be aware that inference on the state number retains some judgemental element.

[53] Although the Silverman (1981, 1986) test is flexible in its hypothesis, it does have the disadvantage of not being a nested test. For example, it could fail to reject the null hypothesis of having k modes, but reject the null of having k-p modes, where k-$p \geq 0$.

Table 3: Silverman´s Multimodality Test

YEAR	CRITICAL BANDWIDTHS AND *P*-VALUES			$k*$
	$h_{crit(1)}$	$h_{crit(2)}$	$h_{crit(3)}$	
1992	2490 [0.00]	2240 [0.00]	1780 [0.16]	3
1994	2530 [0.00]	2170 [0.12]	1600 [0.21]	2
1995	3120 [0.00]	2960 [0.08]	1590 [0.19]	3
1996	3760 [0.00]	2640 [0.13]	1810 [0.26]	2
1997	3060 [0.00]	3200 [0.06]	1930 [0.10]	3
1998	3910 [0.07]	2570 [0.19]	1860 [0.35]	2
1999	4660 [0.00]	2700 [0.10]	1860 [0.08]	2
2000	3875 [0.00]	2530 [0.10]	1710 [0.14]	2
2001	3620 [0.00]	3130 [0.05]	2710 [0.09]	3

Notes: Bootstrapped *p*-values in [·]s.

The results are listed in Table 3, with the first row for any given year indicating the values of $h_{crit}(k)$; the *p*-values associated with the corresponding critical value widths are given in parentheses; and $k*$ representing the number of modes detected.

Taken together, the BM index establishes the presence of multimodality while the Silverman test suggests a persistent ambiguity between trimodality and bimodality over the time period, consistent with the "eye-ball evidence" drawn from Figures 1 - 4.[54]

In order to geographically illustrate the clusters detected in the Kernel density estimations, we have produced a set of maps, Figures 6 and 7, which create a visual impression of the spatial structure of the real GDP per capita across German districts. The categories are defined such that in each income range there resides an equal number of districts. To be consistent with the results of our empirical analysis we have chosen to identify three and six categories of the real GDP per capita in 1992 and in 2001, the first and the last year of the considered time span. The presented maps provide evidence that spatial clusters do exist for the variable under consideration.[55] In particular, as one would expect, the poorest district are concentrated in East Germany. In 1992, all districts, except Berlin, Kreisfreie Stadt

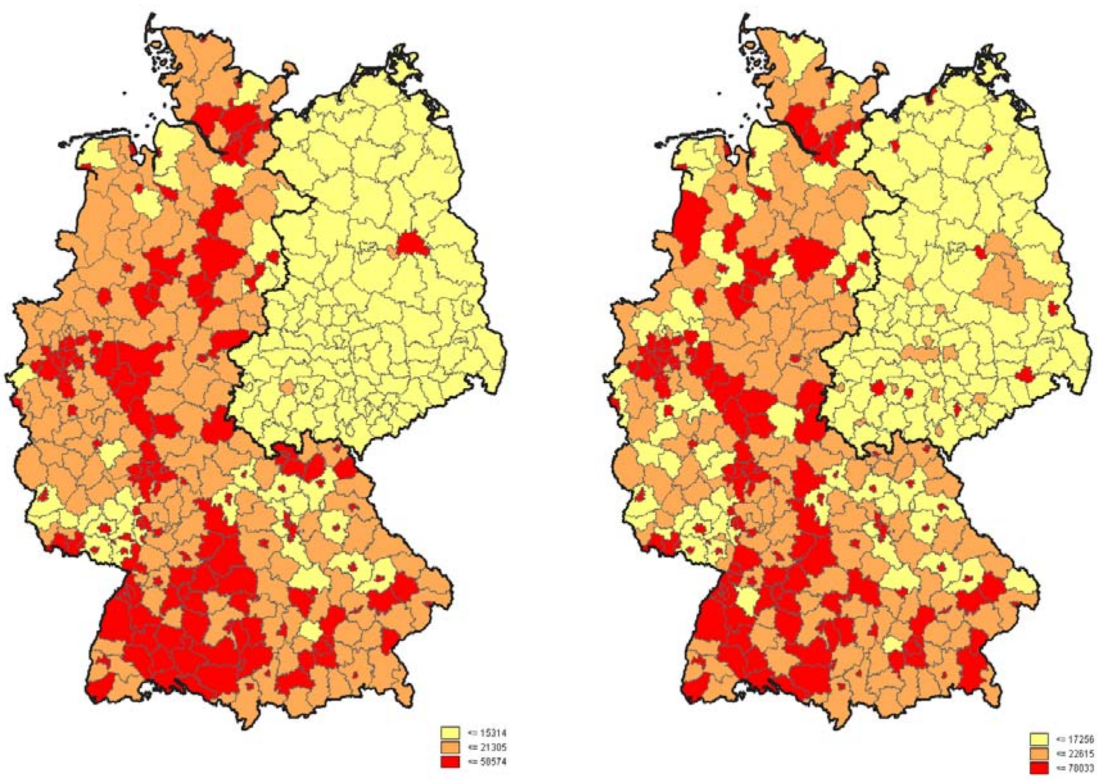
[54] We have also examined the distribution of each district´s per capita income relative to Hamburg´s income. This does not change the shape of the distribution, and the results are virtually identical.

[55] For the correct interpretation of the maps it is important to bear in mind that they are not suitable to assess the absolute growth performance of the 439 German districts: in particular, it is not possible to say whether over the last decade the poorest areas caught up with the richest ones or whether some areas got richer or poorer as they switched from a cluster to another. (The reason for that is that the thresholds defining the identified categories have changed over time). Looking at those thresholds – which all rose considerably – it is indeed possible to state that the average German GDP has risen between 1992 and 2001.

Potsdam and Kreisfreie Stadt Erfurt, belong to the "poorest" group,[56] whereas only 10 percent of the West districts are included in this low-income cluster. By 2001, the proportion of eastern districts that still reside in this same cluster has shrunk to 80 percent of the total eastern districts. However, a number of districts have switched to the richer groups, showing an improvement in their relative income level. In particular, the districts Landkreis Teltow-Fläming and Landkreis Dahme-Spreewald in the greater Berlin area have moved from the low- to the middle-income group. Furthermore, 8 percent of the eastern German districts (Kreisfreie Stadt Dresden, Kreisfreie Stadt Rostock, Kreisfreie Stadt Cottbus, Kreisfreie Stadt Neubrandenburg, Kreisfreie Stadt Jena, Kreisfreie Stadt Erfurt, Kreisfreie Stadt Schwerin, Kreisfreie Stadt Zwickau, Kreisfreie Stadt Potsdam) have gained a foothold amongst the richest elite by 2001.

Figure 6: Real GDP per capita 1992 and 2001

1992 **2001**



Notes: The thin lines indicate the regional boundaries of the 439 districts, the thick lines indicate the East German states and districts.

[56] Data for Mecklenburg-Vorpommern (18 Kreise) is not available for 1992. However, their per capita income level is found to reside within the lowest income category as soon as these figures become available in 1996.

A further piece of information that can be gleaned from the visual inspection of the maps is that approximately 40 percent of West Germany belongs to the high-income cluster, both in 1992 and in 2001, with a high concentration of rich districts localized in the Hamburg area to the North, as well as in the western and southern parts of West Germany. One surprising feature that emerges is the marked downturn in the fortunes of 24 western districts, who experienced an erosion of their per capita GDP from above 21,300 euros in 1992 to 17,200 euros in 2001.[57]

All in all, the comparison between 1992 and 2001 shows that the spatial structure of the real GDP per capita of German districts over the last decade has indeed changed. Figure 6 shows that the relative income position of the East German districts has remained at the bottom of the ranking whereas districts located in the South-West and in the Hamburg area were still included in the richest group. In other words, the relatively "poorer" districts have remained clustered in the eastern part of Germany and the "wealthier" areas have remained localized in the South-West. That said, the emergence of a number of wealthier eastern districts concurrently with the fall-back experienced by a pocket of western regions suggests that the overall picture may be more complex than first thought.

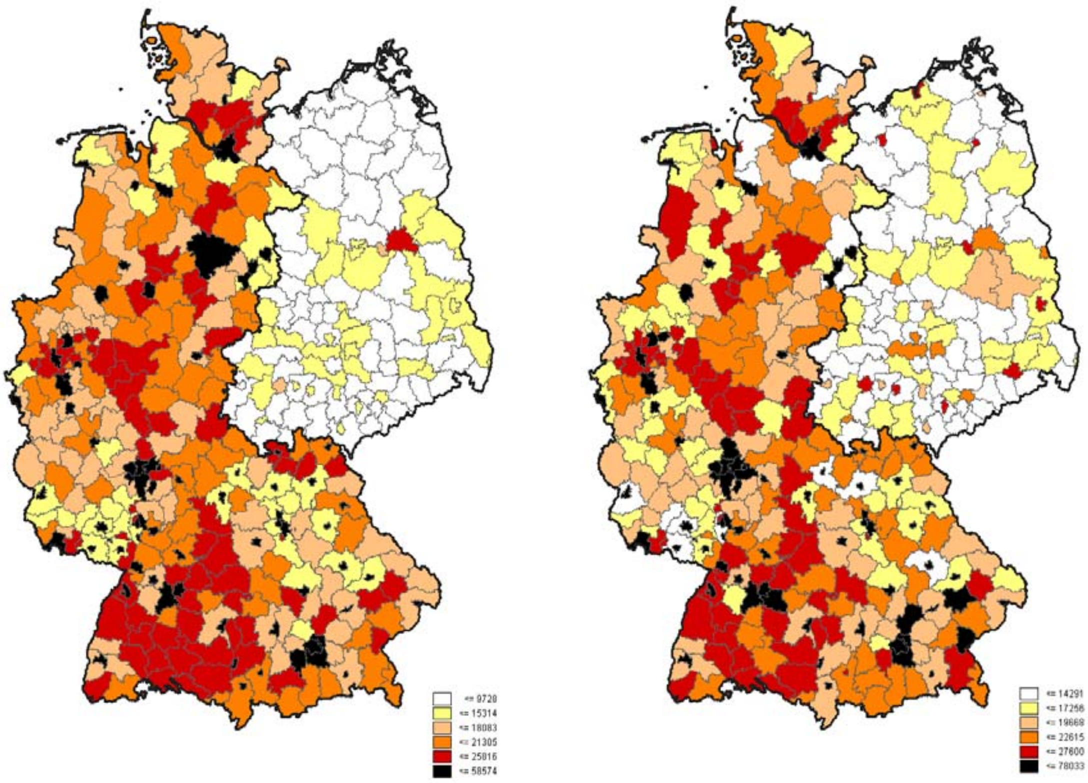
Figure 7 paints the same picture in greater detail as six different income groups are identified. This 6-category map allows a more precise view of the spatial structure of real GDP per capita across the German districts, while retaining a natural consistency with the 3-category map of Figure 6. Of the two lowest income ranges, the very poorest range is observed only in East Germany in 1992. By 2001, however, it is apparent from the more detailed 6-category maps that a number of West German states now reside in this lowest category, particularly in the north and south-west. Within the East there is also a discernible movement from the lowest income range to the second lowest, over the period in question. In the middle income ranges there has been a perceptible emergence of middle-income category districts in the East German states over the 1992-2001 time period, whereas in the West those regions residing in the middle income ranges in 1992 have broadly retained their status throughout the period. Similar to the trends observed in Figure 6, the relatively wealthier regions tend to be concentrated in the west and southern areas of the country in both 1992 and 2001. One can also discern the emergence of a sprinkling of relatively wealthy regions in the East by 2001, due perhaps to real GDP growth associated with urban, commercial areas such as Berlin and Dresden.

Taken as a whole, the visual impression created in Figures 6 and 7 of the spatial structure of the real GDP per capita leads one to conclude the following: over the period 1992-2001 there has been a noticeable catching-up process in terms of the real GDP of East German regions; West German regions that have been residing in middle income ranges tend to have retained this status throughout the period in question, though as illustrated by the 6-category maps a small number of western regions which were in the lower income categories in 1992 have fallen back somewhat by 2001; the relatively richer clusters in the western and southern areas of the country have consolidated their position over the period in question, while a sprinkling of relatively wealthy regions has also emerged in the East.

[57] In 2001 the share of western districts included in the low-income group over the total number of West German districts rose to 18 percent (58 out of a total of 326 West German districts) from 10 percent in 1992.

Figure 7: Real GDP per capita 1992 and 2001

1992 **2001**



Notes: The thin lines indicate the regional boundaries of the 439 districts, the thick lines indicate the East German states and districts.

4. A Nonparametric Test of Density Time Invariance

The visual impression from the density estimates in Figures 1 - 4 is that the per capita income densities have indeed changed across time. In order to determine whether this "eye-ball evidence" is statistically significant, Li´s (1996) nonparametric test has been carried out.[58]

Let $f(x)$ and $g(x)$ denote two bounded and continuous probability density functions observed in two different time periods. The null hypothesis of the test is H_0: $f(x) = g(x)$, against H_1: $f(x) \neq g(x)$.[59]

A conventional measure of global closeness between two functions is the integrated square difference denoted by I [see Pagan and Ullah (1999)]. Given the observations $X = (X_1, ..., X_n)$ and $Y = (Y_1, ..., Y_n)$ drawn from the unknown density functions f_X and f_Y the test is therefore defined by:

[58] Similar ideas have been developed by, inter alia, Anderson et al. (1994) and Berkowitz (2001). In general, this line of research has produced tests either to measure the discrepancies between two density functions or to test the hypothesis that the predictive density generated by a particular model is statistically different from the true density.

[59] For the sake of simplicity and for clarity of exposition, we assume the samples of observations on X and Y to be of equal sizes. The extension of the test for the case of different sample sizes is straightforward.

$$(12) \qquad I = \int \left[f_X(t) - f_Y(t) \right]^2 dt$$

$$(13) \qquad I = \int \left[f_X^2(t) + f_Y^2(t) - 2 f_X(t) f_Y(t) \right] dt$$

$$(14) \qquad I = \int f_X(t) dF_X(t) + \int f_Y(t) dF_Y(t) - 2 \int f_Y(t) dF_X(t).$$

where F_X and F_Y are the distribution functions. In our application, f_X and f_Y correspond to the distributions from different years, i.e. f_X and f_Y are the per capita *GDP* distributions in period t and $t+i$, respectively. The feasible estimator of I, denoted by I_n, can be obtained if one substitutes the density functions f_X and f_Y by their kernel estimates

$$(15) \qquad \hat{f}_X(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h} \right)$$

and

$$(16) \qquad \hat{f}_Y(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{Y_i - x}{h} \right)$$

Using these estimates and replacing F_X and F_Y by their empirical distribution functions, one can write $I_n = I_{1n} + I_{2n}$, where

$$(17) \qquad I_{1n} = \frac{2K(0)}{nh} - \frac{2}{n^2 h} \sum_{i=1}^{n} K\left(\frac{X_i - Y_i}{h} \right)$$

and

$$(18) \qquad I_{2n} = \frac{1}{n^2 h} \sum_{i=1}^{n} \sum_{\substack{i \neq j \\ j=1}}^{n} \left[K\left(\frac{X_i - X_j}{h} \right) + K\left(\frac{Y_i - Y_j}{h} \right) - K\left(\frac{Y_i - X_j}{h} \right) - K\left(\frac{X_i - Y_j}{h} \right) \right].$$

Under the null hypothesis of time invariance, Li (1996) has shown that the test statistic which is based on global closeness between two unknown density functions, is given by:

(19) $$T_n = nh^{1/2}\left(\frac{I_n - \frac{2K(0)}{nh}}{\hat{\sigma}}\right) \xrightarrow{d} N(0,1)$$

where:

(20) $$\hat{\sigma} = \frac{1}{n^2 h}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[K\left(\frac{X_i - X_j}{h}\right) + K\left(\frac{Y_i - Y_j}{h}\right) + 2K\left(\frac{X_i - Y_j}{h}\right)\right]\left[\int K^2(t)dt\right].$$

This test statistic displays several attractive properties in that it has a known limiting standard normal distribution. Furthermore, Li (1996) has shown that the test statistic has a convergence rate faster than the $\sqrt{n}$ rate. The Monte Carlo results indicate that the test performs well for sample size $n \geq 50$ when $n_1 = n_2 = n$.[60] On the consistency of the bootstrap estimates of σ in this context see Hall (1992). The results of the pairwise comparison over time are reported in Table 4 and 5.

Table 4: The Li (1996) T_n Test Statistics for n = 419

TIME COMPARISON	TEST STATISTIC
1992 versus 1994	8.68***
1994 versus 1996	0.12
1996 versus 1998	0.47
1998 versus 2000	1.40*
1992 versus 2001	9.96***

Note: (***), (**) and (*) indicates significance at the 1%, the 5% and the 10% level, respectively.

[60] Alternatively, Li (1996) has also suggested a J_n and J_{nc} test statistic, respectively. The Monte Carlo evidence indicates that J_{nc} has a significant negative bias, while J_n and T_n have similar power.

Table 5: The Li (1996) T_n Test Statistics for n = 439

TIME COMPARISON	TEST STATISTIC
1996 versus 1998	0.49
1998 versus 2000	1.38*
1996 versus 2001	3.36***

Note: (***) and (*) indicates significance at the 1% and the 10% level, respectively.

As shown in Table 4 and 5, there is one highly significant change in the distribution occurring within the first two years after unification.

5. Suggestions for future research

Thusfar our primary focus has been that of identifying both the distribution and intra-distributional dynamics of aggregate regional per capita income over the 1992-2001 period. Having established the bimodal/trimodal nature of this distribution across districts, it is understandable that one would like to uncover the underlying drivers of this distribution process. We now delve into these underlying factors in a little more detail, offering some tentative insights and pointing the way for useful future research. The approach adopted here stems from the work of Desmet and Fafchamps (2004), who, using non-parametric methods, have recently examined the spatial distribution of employment, as opposed to GDP, across US counties between 1972 and 1992. Their results point to an increase in total employment concentration, with this aggregate dynamic driven by services sector divergence outweighing opposing primary and secondary influences. This has prompted us to consider how disaggregated sectoral dynamics interact with each other in the German context.[61]

Table 6:Sectoral employment growth on initial sectoral employment (439 *Kreise*)

DEPENDENT VARIABLE: AVERAGE ANNUAL GROWTH RATE IN SECTORAL EMPLOYMENT 1992-2001				
	Total	Primary sector	Secondary Sector	Tertiary sector
constant	0.0046	0.0257	0.0578	0.0289
	[0.59]	[0.00]	[0.00]	[0.00]
$\ln(L_{1992})$	-0.0001	-0.0077	-0.0075	-0.0011
	[0.87]	[0.00]	[0.00]	[0.13]

Notes: *p*-values given in brackets; primary sector refers to “Land- und Forstwirtschaft, Fischerei”, secondary sector refers to “Produzierendes Gewerbe” and tertiary sector refers to “Dienstleistungsbereiche” as per dataset “Arbeitskreis Volkswirtschaftliche Gesamtrechnung der Länder“.

As a prelude to further research , Table 6 provides the OLS estimation results of a linear regression of average employment growth ($\Delta\%L_{1992\text{-}2001}$) on initial, 1992, log employment $\ln(L_{1992})$ – the standard test for unconditional β-convergence, with a positive coefficient on initial employment pointing to

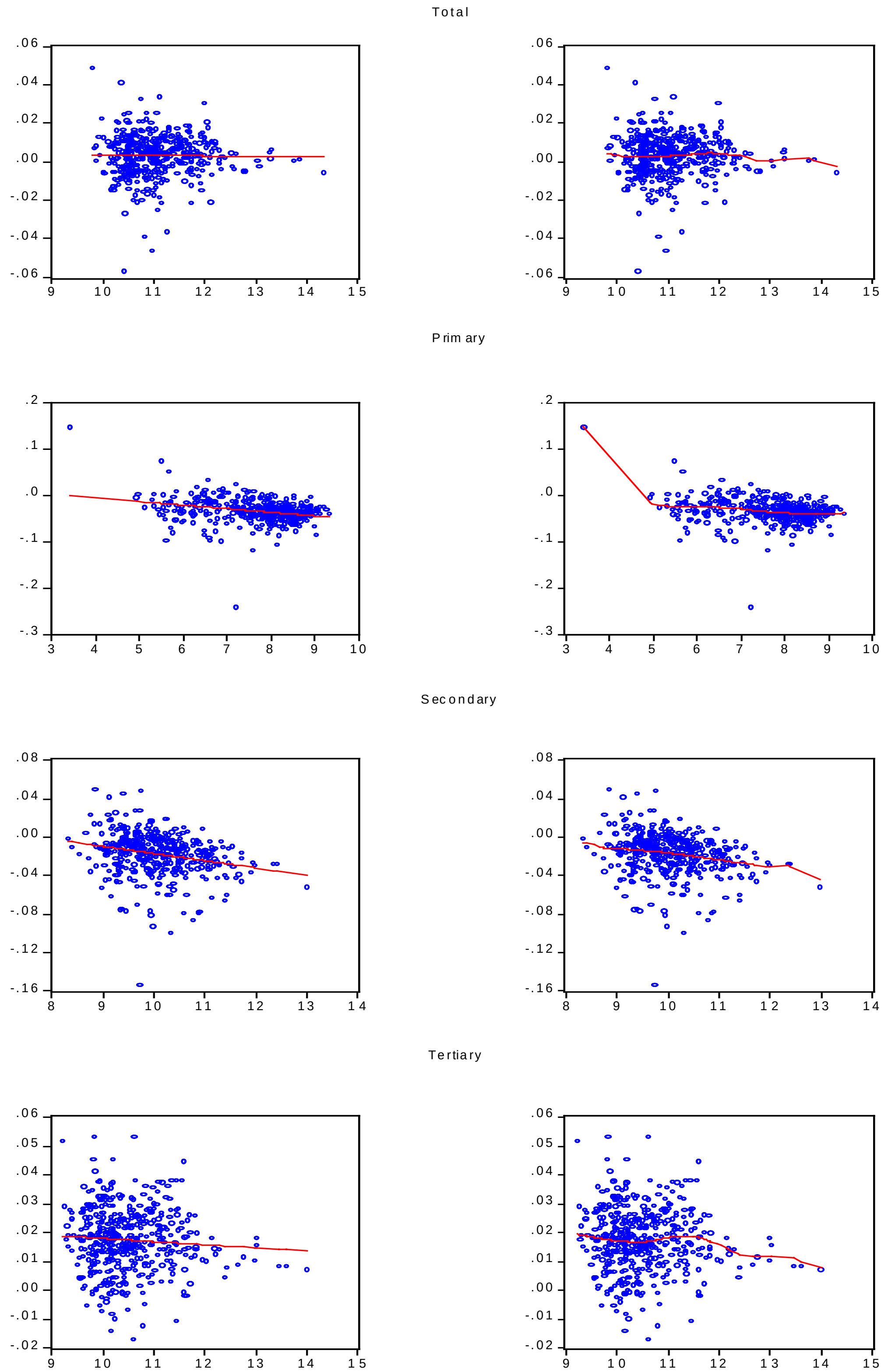
[61] As sectoral GDP data for Sachsen for 1992 is not available in the dataset “Arbeitskreis Volkswirtschaftliche Gesamtrechnung der Länder“, the following analysis is conducted using employment data from 1992 to 2001 for the 439 *Kreise*.

concentration (divergence) and a negative coefficient indicating deconcentration (convergence).[62] It is clear that for both the aggregate (total employment) and the tertiary sector (services) the coefficient on initial employment is not significantly different from zero. With respect to the agriculture and the industry sectors, negative significant coefficients on initial employment indicate convergence. For the sake of illustration, the corresponding scatter plots with linear regression line and kernel fit are also provided in Figure 8.[63] Can any one coherent insight be gleaned from these signals? Perhaps. It may well be the case that a lack of convergence in the tertiary sector is dampening the converging influence of the primary and secondary sector, culminating in a weaker convergence process in the aggregate. What this does suggest is that a more indepth analysis of the evolution towards a more services-driven economy may well represent a fruitful avenue for future research.

[62] One caveat should be borne in mind when reading the estimates in Table 6: it is conceivable that they may be sensitive to outliers or omitted variables.

[63] The results of this β-convergence check are mirrored in σ-convergence analysis, when we compare standard deviations for 1992 and 2001.

Figure 8: Sectoral employment scatter plots with OLS regression line (left column) and kernel fit (right column)



6. Conclusions and further comments

The objective of this paper is to address the question of convergence across German districts in the first decade after German unification by drawing out and emphasising some stylised facts of regional per capita income dynamics, rather than estimating any particular economic model. We achieve this by employing techniques which focus on the evolution of the entire cross-sectional income distribution. In particular, we follow a distributional approach to convergence based on non-parametric kernel density estimation and implement a number of tests to establish the statistical significance of our findings. The visual inspection of the estimated densities indicates the following: the presence of trimodality in 1992; in subsequent years less pronounced trimodality, supporting the notion of a catching-up process of eastern Germany in the early 1990s; and a tendency for the remaining two peaks to move apart, resulting in a swelling of the middle income mode and a more pronounced high income mode. This ambiguity between trimodality and bimodality over the period in question is supported by statistical tests such as the Silverman multimodality test and the BM index. Li´s (1996) nonparametric test lends further statistical support to the visual impressions. It should be noted, of course, that empirical evidence suggesting bimodality runs counter to recent theoretical views.[64]

The colour-coded maps of the German districts geographically illustrate the clusters detected in the Kernel density estimations and provide evidence that spatial clusters of income do exist over the period in question. Consistent with the density estimates there emerges a picture of East German convergence, a swelling middle-income group and a more pronounced high-income group.

An alternative approach to investigating the presence of convergence clubs would be to track in more detail the performance of each geographical unit. This may provide another dimension of disparity that is relevant for economic policy making. From a policy perspective, besides having information about the entire cross-section of observations, it is also important to know how likely is each district to improve its conditions, how many did so and what are their characteristics. In other words, whether or not districts that were rich (poor) a decade ago are the same ones that are rich (poor) now has relevant policy implications. If the poor regions are persistently poor, one may want to consider public programs aimed at enhancing the performance of these districts. On the other hand, if the incomes per capita are rotating over time, one would be less concerned about overall geographical income

[64] The exact nature of multimodality is indeed surrounded by some degree of uncertainty. At first glance, it might seem promising to consider growth model with multiple equilibria in the tradition of Aghion and Howitt (1998, Chapter 10), Azariadis (1996), Drazen and Azariades (1990) and Matsuyama (1991) when trying to explain "job-poor" versus "job-rich" growth experiences. In such models, a country may be trapped in a "job-poor" equilibrium when, in principle at least, an alternative and superior equilibrium is also feasible. However, the recent literature has cast doubts on the robustness of multiple equilibria. Frankel and Pauzner (2000) analyse a two sector model with increasing returns, based upon Matsuyama (1991). They show that if the wage is stochastic and arrives as a Poisson process, the muliplicity property may be eliminated because some of the deterministic equilibria are more robust to perturbations than others. A similar conclusion has been established by Herrendorf et al. (1999) for heterogeneous agents. They show that sufficient heterogeneity of agents will lead to a refinement in the set of observable equilibria and uniqueness in models like that of Matsuyama (1991).

distribution. Our approach has not conceptualised this alternative mixing or ranking change aspect of disparity. Further consideration should be given to such indicators in future research.

References:

Aghion, P. and P. Howitt (1998) *Endogenous Growth Theory*, Cambridge (MIT Press).

Anderson, N.H., Hall, P. and D.M. Titterington (1994) “Two-Sample Test Statistics for Measuring Discrepancies Between Two Multivariate Probability Density Functions Using Kernel-Based Density Functions”, *Journal of Multivariate Analysis* 50, 41-54.

Azariadis, C. (1996) "The Economics of Poverty Traps Part One: Complete Markets", *Journal of Economic Growth* 1, 449-486.

Azariadis, C. and A. Drazen (1990) "Threshold Externalities in Economic Development", *Quarterly Journal of Economics* 105, 501-526.

Barro, R.J. (1991) "Economic Growth in a Cross-Section of Countries", *Quarterly Journal of Economics* 106, 407-443.

Berkowitz, J. (2001) “Testing Density Forecasts, with Applications to Risk Management”, *Journal of Business and Economics Statistics* 19, 465-474.

Bianchi, M. (1997) "Testing for Convergence: Evidence from Non-Parametric Multimodality Tests", *Journal of Applied Econometrics* 12, 393-409.

Brakman, S., Garretsen, H. and M. Schramm (2004) „The Spatial Distribution of Wages: Estimating the Helpman-Hanson Model for Germany“, *Journal of Regional Science* 44, 437-466.

Bulli, S. (2001) “Distribution Dynamics and Cross-Country Convergence: A New Approach”, *Scottish Journal of Political Economy* 100, 223-251.

Corrado, L., Martin, R. and M. Weeks (2005) "Identifying and Interpreting Regional convergence Clusters Across Europe", *The Economic Journal* 115, C133-C160.

Desmet, K. and M. Fafchamps (2004) “Employment Across US Counties”, *CEPR Discussion Paper No.* 4689, London.

Efron, B. (1979) "Bootstrap Methods – Another Look at the Jack-Knife", *Annals of Statistics* 7, 1-26.

Efron, B. and R.J. Tibshirani (1993) *An Introduction to the Bootstrap*, London (Chapman and Hall).

Faini, R. (1984) “Increasing Returns, Non-Traded Inputs, and Regional Development”, *Economic Journal* 94, 308-323.

Frankel, D.M. and A. Pauzner, A. (2000) „Resolving Indeterminacy in Dynamic Settings: The Role of Shocks, *The Quarterly Journal of Economics* 115, 285-304.

Funke, M. and Niebuhr, A. (2005) “Threshold Effects and Regional Economic Growth – Evidence from West Germany”, *Economic Modelling* 22, 61-80.

Geweke, J., Marshall, R.C. and G. Zarkin (1986) “Mobility Indices in Continuous Time Markov Chains”, *Econometrica* 54, 1407-1423.

Hall, P. (1992) “Effect of Bias Estimation on Coverage Accuracy of Bootstrap Confidence Intervals for a Probability Density”, *Annals of Statistics* 20, 675-694.

Hall, P., Horowitz, J.L. and B.-Y. Jing (1995) “On Blocking Rules for the Bootstrap with Dependent Data”, *Biometrika* 82, 561-574.

Hall, P. and M. York (2001) "On the Calibration of Silverman´s Test for Multimodality", *Statistica Sinica* 11, 515-536.

Herrendorf, B., Valentinyi, A. and R. Waldmann (2000) "Ruling Out Multiplicity and Indeterminancy: The Role of Heterogeneity", *Review of Economic Studies* 67, 295-307.

Islam, N. (1998) "Growth Empirics: A Panel Data Approach – A Reply", *Quarterly Journal of Economics* 113, 325-329.

Izenman, A.J. and C.J. Sommer (1988) "Philatelic Mixtures and Multi-Modal Densities", *Journal of the American Statistical Association* 83, 941-953.

Jones, C.I. (1998) *Introduction to Economic Growth*, London (W.W. Norton & Company).

Jüßen, F. (2005) “A Distribution Dynamics Approach to Regional GDP Convergence in Reunified Germany”, Dortmund University, mimeo.

Lee, K., Pesaran, M.H. and R. Smith (1997) "Growth and Convergence in a Multi-Country Empirical Stochastic Growth Model", *Journal of Applied Econometrics* 12, 357-392.

Lee, K., Pesaran, M.H. and R. Smith (1998) "Growth Empirics: A Panel Data Approach – A Comment", *Quarterly Journal of Economics* 113, 319-323.

Li, Q. (1996) “Nonparametric Testing of Closeness Between Two Unknown Distribution Functions”, *Econometric Reviews* 15, 261-274.

Lopez-Bazo, E. Vaya, E. and M. Artis (2004) “Regional Externalities and Growth: Evidence from European Regions”, *Journal of Regional Science* 44, 43-73.

Magrini, S. (2004) “Regional (Di)Convergence”, in: Henderson, J.V. and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*, Vol. 4, Amsterdam (Elsevier), 2741-2796.

Mankiw, N.G., Romer, D. and D.N. Weil (1992) "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics* 107, 407-437.

Matsuyama, K. (1991) "Increasing Returns, Industrialization and Indeterminancy of Equilibrium", *Quarterly Journal of Economics* 106, 617-650.

Murphy, K.M., Shleifer, A. and R.W. Vishny (1989) "Industrialization and the Big Push", *Journal of Political Economy* 97, 1003-1293.

Nelson, R. (1956) “A Theory of the Low-Level Equilibrium Trap in Underdeveloped Economies”, *American Economic Review* 46, 394-408.

Paap, R. and H.K. van Dijk (1998) "Distribution and Mobility of Wealth of Nations", *European Economic Review* 42, 1269-1293.

Pagan, A. and A. Ullah (1999) *Nonparametric Econometrics*, Cambridge (Cambridge University Press).

Park, B.U. and B.A. Turlach (1992) “Practical Performance of Several Data Driven Bandwidth Selectors (with Discussion)”, *Computational Statistics* 7, 251-285.

Pittau, M.G. (2005) "Fitting Regional Income Distributions in the European Union", *Oxford Bulletin of Economics and Statistics* 67, 135-161.

Quah, D. (1996a) “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics”, *The Economic Journal* 106, 1045-1055.

Quah, D.T. (1996b) "Convergence Empirics Across Economies with (some) Capital Mobility", *Journal of Economic Growth* 1, 95-124.

Quah, D.T. (1997) "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs", *Journal of Economic Growth* 2, 27-59.

Quah, D. (2006) “Growth and Distribution”, unpublished manuscript, London School of Economics (see http://econ.lse.ac.uk/staff/dquah/p/gnd.pdf).

Rodrik, D. (1996) "Coordination Failure and Government Policy: A Model with Applications to East Asia and Eastern Europe", *Journal of International Economics* 40, 1-22.

Silverman, B.W. (1981) "Using Kernel Density Estimates to Investigate Multimodality", *Journal of the Royal Statistical Society* 43, 97-99.

Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*, London (Chapman and Hall).

Sinn, H.W. and Sinn G. (1992). *Jumpstart. The Economic Unification of Germany*, MIT Press, Cambridge, Mass.

Solow, R.M. (1956) "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics* 70, 65-94.

Timm, N.H. (2002) *Applied Multivariate Analysis*, Berlin (Springer Verlag).

Wand, M.P. and M.C. Jones (1995) *Kernel Smoothing*, London (Chapman and Hall).

Economic Growth Across Space and Time: Subprovincial Evidence from Mainland China

Declan Curran, Michael Funke, and Jue Wang

1. Introduction

China's macroeconomic growth performance over the last decade has been phenomenal, with GDP growing at a blistering pace of 8 percent per annum, on average. The expansion of China's role in world trade has been no less remarkable, with its overall share in world trade rising from less than 1 percent in 1979 to 6 percent in 2005. Furthermore, according to the World Bank, economic growth has contributed to rapidly falling poverty rates in China. From 1994 to 2004, the percentage of population living below the poverty line declined from 35% to 17% in rural China, and from 0.90 percent to 0.30 percent in Chinese cities.[65] Despite these remarkable achievements, a great deal of debate and attention has focussed on China's uneven regional developments. Urban and rural standards of living continue to be poles apart. Rural prefectures and townships still struggle to get to grips with basic healthcare and education provision. Despite commitments from the central government to implement a new medical insurance scheme and free education, outlays on health care and education as a proportion of total spending remain lower than they were a decade ago.[66]

While high-speed economic growth and dramatic social changes continue to distinguish China across the globe, the country's leadership has recently been eyeing a smoother ride on its development path by setting forth a guideline prioritising "harmony". The sustained reforms and opening-up over the past two and half decades have resulted in prosperity for many Chinese citizens, but the income gap across the country are amongst the top concerns of the Communist Party of China (CPC). Over the past three years rural income per head has risen by more than 6% annually in real terms, but this has not halted the widening of the urban-rural income gap. The CPC's uneasiness stems from the fact that China's history is littered with rebellions, uprisings, and revolutions sparked by economic inequalities. Against this historical experience, Chinese leaders have placed the concept of a "harmonious socialist society" for renewed political legitimacy and political cohesion of the country at the top of their "to do" list. It is envisaged that this harmonious society should feature democracy, the rule of law, and enable all the people to share the social wealth brought by reform and development.

In this paper we consider the process of regional economic growth across China over the period 1997-2005. We introduce a county- and city-level dataset of real GDP per capita that spans the entirety of

[65] See http://iresearch.worldbank.org/PovcalNet/jsp/index.jsp. Some optimistic observers have argued that China's GDP is likely to grow at rates of at least 8 percent per year for at least a generation, i.e. to 2030, and perhaps beyond that date [see Fogel (2006)].

[66] "Rural China: Missing the Barefoot Doctors", *The Economist*, 11 October 2007, pp. 27-29.

mainland China.[67] The main motivation for this paper is to contribute to fuller understanding of the persistent differences in economic performance across China. The paper is structured as follows: Section 2 provides a concise literature review which collects the key findings to date concerning Chinese regional growth. Section 3 introduces the county- and city-level dataset utilised in this paper, as well as illustrating with colour-coded maps the insights gained from moving from provincial-level to county- and city-level disaggregation. In Section 4 we establish the evolution of the entire cross-sectional distribution of real GDP per capita over time using non-parametric kernel density estimation and track the dynamics of individual county- and city-level districts over time using the transition probability matrix technique and the associated stochastic kernel estimator. Section 5 expands our dataset to include a set of explanatory variables and utilizes OLS, LTS, and BIF regression estimators to test for conditional β-convergence across these county- and city-level districts. Section 6 sets out conclusions, as well as implications for policymakers.

2. Literature Review

Previous literature has analysed the uneven pace of reform and growth across Chinese regions from various angles. The insights and results of existing studies can be summarised as follows:

(a) The assessment of regional inequality is not independent of the degree of disaggregation. In most papers the measurement of inequality is still based upon provincial-level data. On the contrary, Herrmann-Pillath et al. (2002) use prefecture-level data for a total of 312 prefectures in 1993 and 1998, and conclude that regional developments should be analysed on a high level of disaggregation. Jones et al. (2003) and Song et al. (2000) have used data for about 200 cities and have concluded that differences in growth rates are far more severe than indicated in studies using data at higher levels of aggregation.

(b) Unel and Zebregs (2006) have demonstrated that capital deepening has been by far the most important source of GDP per capita growth across Chinese provinces in the 1980s and 1990s.

(c) Differences in natural endowments have contributed to the divergence in economic activity and income across space [Bao et al. (2002), Demurger et al. (2002)].

(d) Uneven preferential open-door policies in the post reform period may have led to different policy-determined clubs of provinces. For example, Démurger et al. (2002) have constructed an index ranging from 0 to 3 for each province during the reform years depending upon the type and extent of favoured free trade zones that are present.

(e) Dayal-Gulati and Husain (2002) have shown that the prevalence of state-owned enterprises and a high ratio of bank loans-to-deposits – an indication of large directed lending – were associated with lower growth and centripetal forces.

[67] The quantity of real GDP generated each district, scaled by district population, is a standard proxy for the productivity in the face of data constraints at high levels of disaggregation. It is not intended to represent income per capita.

(f) One reason given for diverse regional growth patterns is an uneven influx of FDI with a high concentration in coastal areas [see Wei et al. (1999), Wen (2007)].
(g) Démurger (2001) has demonstrated that transport facilities are a key differentiating factor in explaining regional growth differentials.
(h) The coastal areas have taken advantage of their long commercial and industrial traditions and geographical and ethnic links with Hong Kong, Macao and Taiwan. Therefore they have attracted a dominant proportion of FDI before FDI started to penetrate into interior regions.

3. Regional Economic Growth Across China: Descriptive Evidence

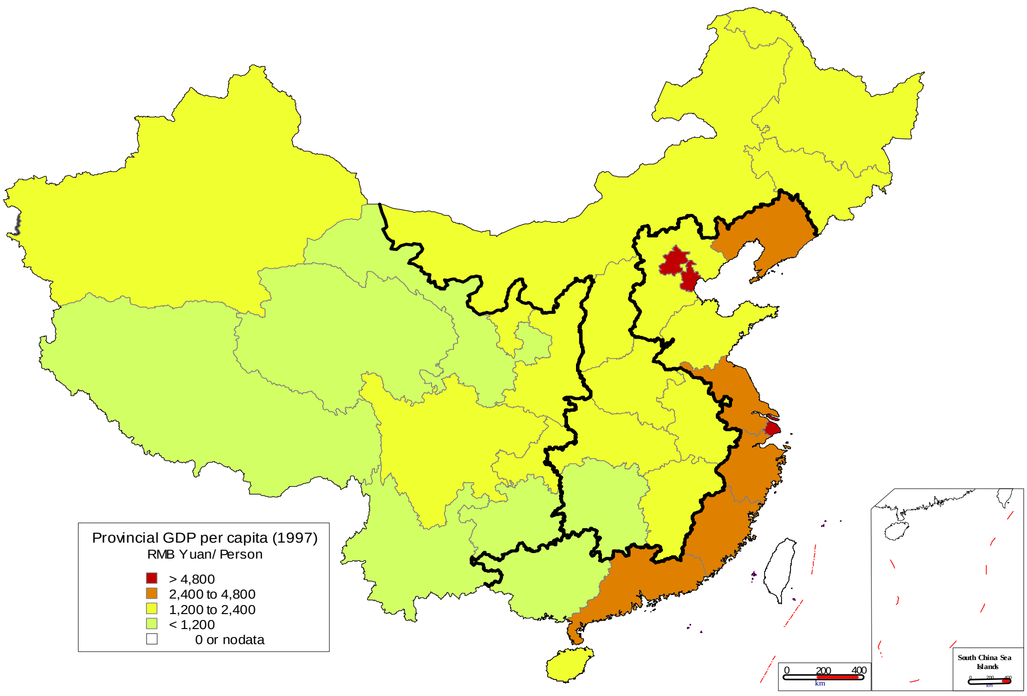
Many regional growth studies of the Chinese economy have been based upon data relating to Chinese province-level (shěngjí) divisions. However, these Chinese provinces represent very large geographical units, especially in the western and central regions – in many cases they are comparable in size to large European countries. A lower level of aggregation can therefore be regarded as a natural choice for an analysis of regional growth patterns. In this paper we utilize a dataset disaggregated to county- and city-level. As of December 31, 2005, the People's Republic of China administers 33 province-level regions, which comprise of 333 prefecture-level regions that are further divided into 2,862 county-level regions (xiànjí), 41,636 township-level regions, and several village-level regions. The thirty-three province-level (shěngjí) divisions are comprised of twenty-two provinces, five autonomous regions, four municipalities, and two special administrative regions. The dataset utilized here reports on 2,283 county- and city-level districts; which we refer to as "districts" for the remainder of the paper; and when missing values are excluded the dataset yields 2,199 observations for county- and city-level GDP per capita over the period 1997-2005. The GDP per capita data has been deflated using provincial-level GDP deflators obtained from nominal and real GDP indices available from the *CEIC* Database. Unless otherwise indicated, all other data has been derived from the China Data Centre at the University of Michigan (see http://www.umich.edu/~iinet/chinadata/). Every effort has been made to take into account changes in administrative boundaries over time, with case-by-case estimates where counties and/or cities had to be reshuffled and fitted into newly formed larger aggregates.[68]

In order to gain a more intuitive feel for the different levels of aggregation of Chinese data (provinces, county- and city-level, as well as a Western-Central-Coastal distinction that has emerged in the literature), it is useful to begin with the most aggregated view and then zoom in. The obvious starting point in such a "top-down" view is with the empirically observed belt of the three regions (western, central, and coastal) that have become the standard point of departure in the literature. Even a casual

[68] While GDP is the more common measure of national income, GNP, in cases of countries benefiting from substantial foreign direct investment inflows, is regarded the more appropriate measure, as it excludes profits and remittances repatriated by foreign multinationals to their home country. Unfortunately, in the Chinese case, data constraints dictate that we use GDP as our measure of national income.

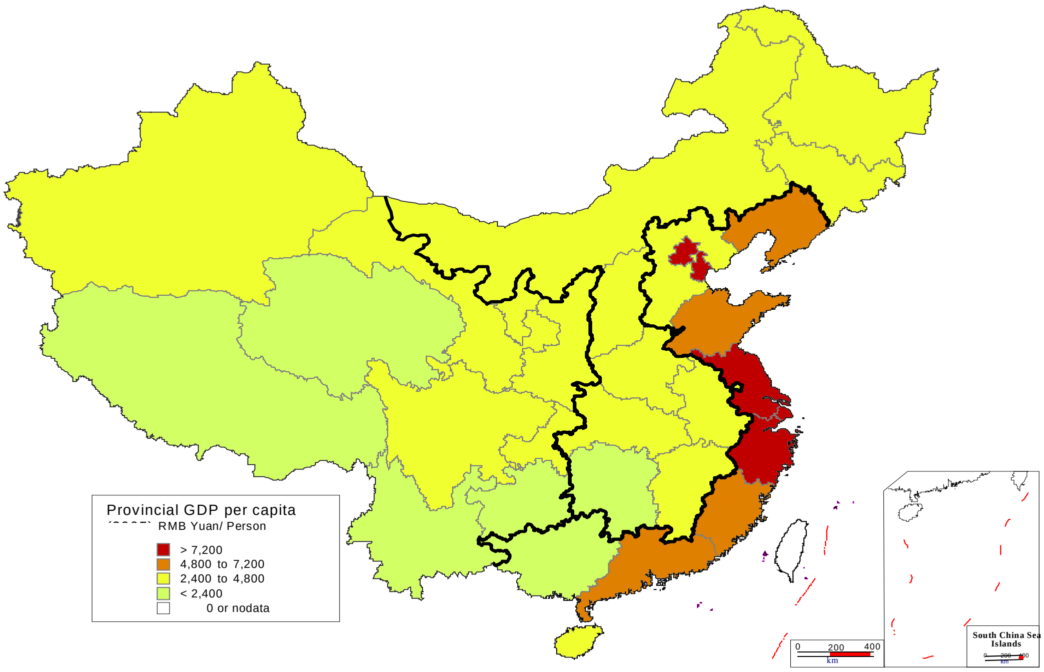
glance at Chinese national accounts data reveals the disparity existing between these regions: in 1997, for example, the real GDP per capita of the western and central regions were 71% and 82% respectively of the national per capita figure, while the coastal (eastern) region's real GDP per capita was 159% of the national per capita figure. In 2005, a similar situation was evident, with the western, central, and coastal regions now clocking in at 78%, 94%, and 185% of the national per capita figure. This reflects an annual growth rate of 8.1% in the west, 8.6% in the centre, and 8.7% on the coast over the 1997-2005 period. Taken as a whole, a comparison of these three belts paint a picture of strong growth across the board, but also one of the coastal region continuing to steam ahead while the central region is unable to close the gap between it and the coast, and the western region falls further behind. Figures 1 and 2 provide a colour-coded illustration of the Chinese West-Central-Coastal disparities in real GDP per capita as they stood at 1997 and 2005, with the three belts divided into their constituent provinces. The West-Central-Coastal distinction is clear to see from Figures 1 and 2. What is more, coastal real GDP per capita appears to perform strongly over the 1997-2005 period, in contrast with performance of the western and central regions. This impression is broadly in line with that of Yao and Zhang (2001), who have suggested that the three belts can be divided into three distinct diverging clubs. In this way, China could be characterised by a three-tiered cluster growth pattern.

Figure 1: West-Central-Coastal Disparities in Real GDP per Capita 1997



Note: The three belts consist of the following provinces: (i) *Coast*: Beijing, Tianjin, Liaoning, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Hainan, and Guangxi; (ii) *Central*: Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Inner Mongolia, and Hunan; (iii) *West*: Sichuan, Chongqing, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang. The relevant data source here is China Data Online: http://chinadataonline.org/. The nominal GDP per capita data has been deflated using provincial-level GDP deflators (base year 2000 = 100) obtained from nominal and real GDP indices available from the *CEIC* Database. To ensure the compatibility and integrity of the data on different levels of aggregation, belt-level data has been obtained by aggregating the respective provincial data. The national total is calculated directly from national data in China Data Online.

Figure 2: West-Central-Coastal Disparities in Real GDP per Capita 2005

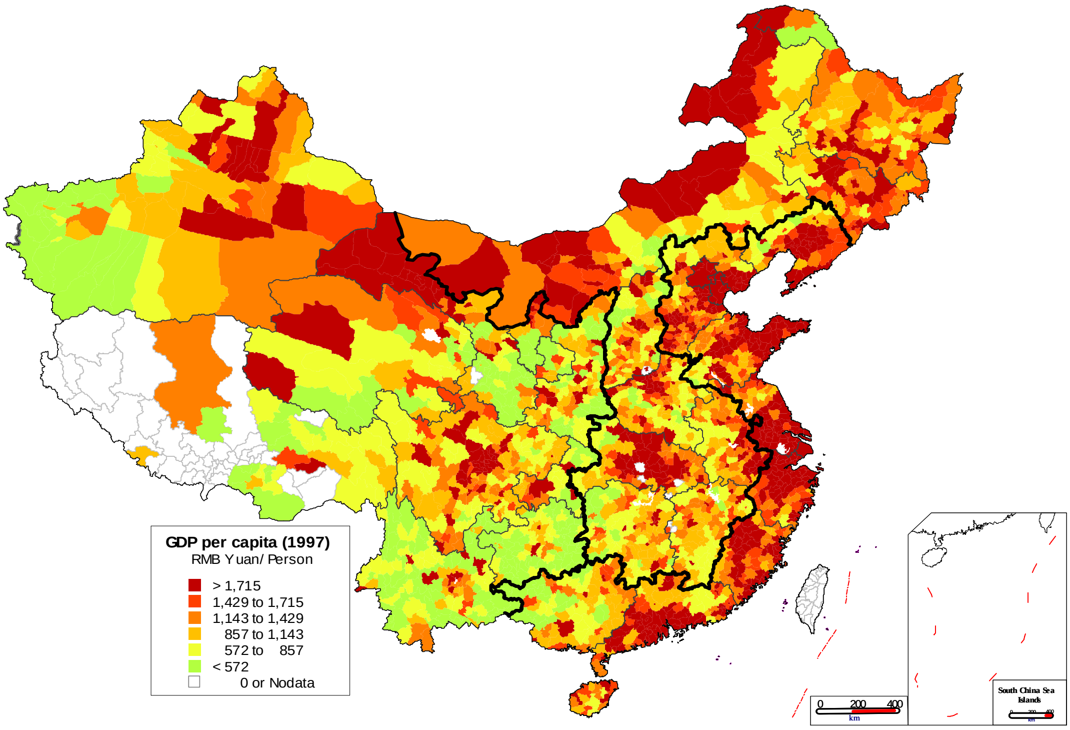


Of course, it is natural to wonder how much information can really be gleaned from such an aggregated picture, which may conceal substantial heterogeneity or smooth over the impact of important economic developments within each of the three belts. This is where the county- and city-level data can make a real contribution to understanding the facts of Chinese economic growth "on the ground", as it offers the potential of attributing economic growth (or lack of) to the specific region from which it emanates, rather than averaging it across large geographical units. Figures 3 and 4 provide colour-coded maps of this county- and city-level data for 1997 and 2005.[69]

What emerges from Figures 3 and 4 is a much more mixed picture than that suggested by the data at either provincial or "three belt" level. Firstly, it is clear that pockets of relatively high GDP per capita are dispersed across the entirety of Mainland China, rather than being confined just to the coastal areas. It seems that the relatively high GDP per capita districts in the coastal region are not as cohesive as Figures 1 and 2 might suggest. What is more, it is the western region and northernmost parts of the central region that appear to gain a greater foothold in the relatively high GDP per capita category over the 1997-2005 period.

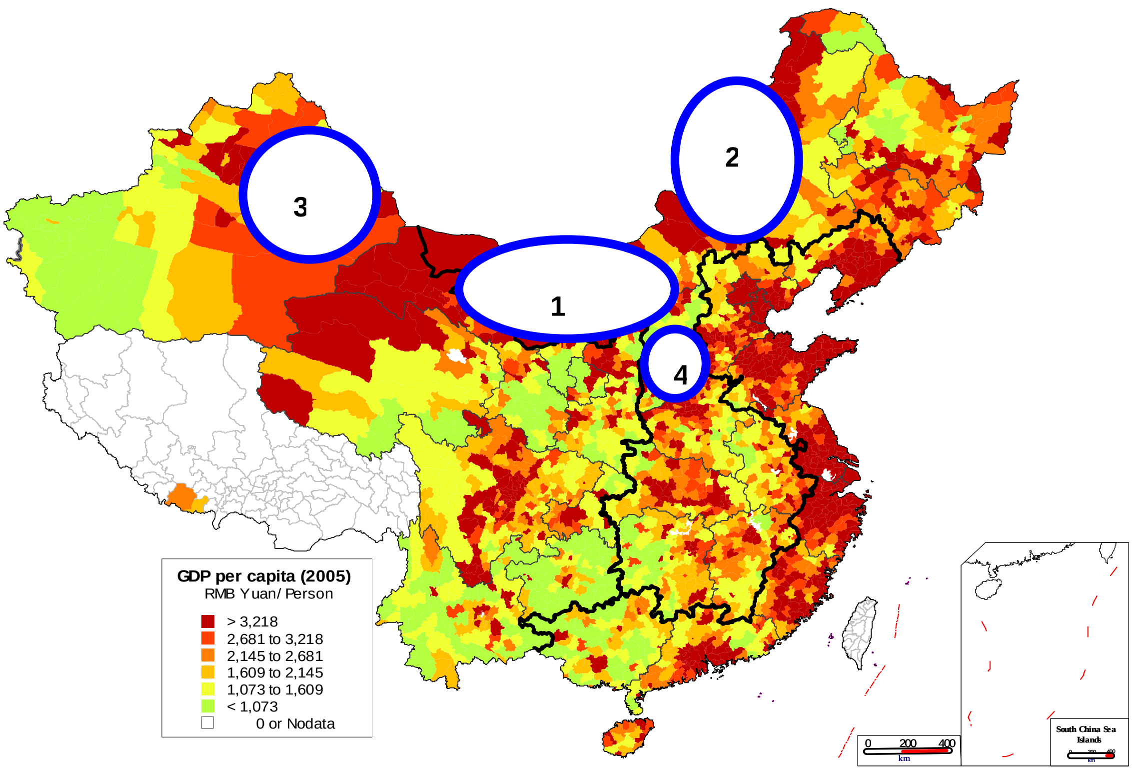
[69] Data prior to 1997 is not available at this level of disaggregation.

Figure 3: County-Level and City-Level Real GDP per Capita 1997



Note: Areas in white are those districts for which data is not available.

Figure 4: County-Level and City-Level Real GDP per Capita 2005



However, it is wise to treat Chinese national accounts data with some caution. The Chinese National Bureau of Statistics (NBS) is still in the process of fully implementing the principles laid down by the Standardised National Accounts System (SNA), as advocated by the OECD.[70] In light of this, we now consider in more detail those districts of the western and central regions which exhibit a notable change in GDP per capita over the 1997-2005 period.

The districts whose GDP per capita grew rapidly over the 1997-2005 period (circled and numbered in Figure 4, above) are mostly located in central Inner Mongolia (1 and 2) and Xinjiang province (3), as well as in the middle of Shanxi province (4). These pockets of high growth do receive support from existing literature: Gao (2004) notes that the central counties in Inner Mongolia experienced rapid economic growth of over 10% per annum over the 1995-2002 period. Luo (2004) has also pointed to Inner Mongolia and Xinjiang as being fastest growing provinces of the 12 Western provinces (including Inner Mongolia). Zhang et al (2006, 2007), using exploratory spatial data analysis methods to analyse the real GDP per capita growth of a sample of 341 districts obtained from dividing sub-provincial regions into counties and municipalities, also identify central Inner Mongolia as a region of high growth over the period 1990-2004. Table 1 presents the sectoral composition of GDP for the ten districts which generated the highest real GDP per capita in western China in 2005. Primary industry refers to farming, forestry, animal husbandry and fishing, while secondary industry includes mining, manufacturing, electricity production, and construction.

Table 1: Composition of GDP in the 10 Western Counties with Highest GDP per Capita in 2005

Area	Province	Primary Industry (%)	Secondary Industry (%)	Tertiary Industry (%)	GDP per capita (yuan)
Eji'na Qi	Inner Mongolia	4	36	60	20,624
Erenhot City	Inner Mongolia	1	18	81	16,390
Yi Jin Huo Luo Qi	Inner Mongolia	4	41	55	16,098
E Tuo Ke Qi	Inner Mongolia	5	75	19	14,935
Zhun Ge Er Qi	Inner Mongolia	3	62	34	12,425
Yanchuan County	Shannxi	3	92	5	12,151
Korla City	Xinjiang	6	79	15	11,744
Golmud City	Qinghai	1	70	29	11,644
Akesaihasake County	Gansu	6	60	34	11,391
Yu Men City	Gansu	6	77	17	10,237
National Total		13	48	40	3,686

Note: Yanchuan county data refers to 2004.

The three districts generating the highest real GDP per capita in 2005 are noticeably more tertiary-orientated than the other districts shown. Erji'na, traditionally an agricultural county, has enjoyed strong secondary and tertiary growth since 2000. The annual gross industrial output value of Erji'na has grown to nearly 600 million RMB, over 6 times that produced in 2000, while its value-added has

[70] For a detailed account of differences between existing Chinese GDP measurement techniques and 1993 SNA guidelines see Xu (2003), where he concludes that China's ongoing transition to the 1993 SNA does not detract from the international comparability of Chinese GDP estimates.

expanded more than eightfold from 2000 to 2005. Tourism has also increased in Erji'na as Dongfeng Spaceflight City, the launch site of China's first and second manned space flights, is located in this district. Since 2004, the local government has invested over 310 million RMB to build the Ceke border crossing, an outlet for cross-border trade which facilitates the export and import of goods, especially coal. Similarly, Erenhot city has recently established two important outlets for cross-border trade, which accounts for its high proportion of tertiary industry (over 80% of total GDP). From 2000 to 2005, the value of imports and exports to and from Erenhot have soared from 400 million USD to 2.2 billion USD, growing at a rate of 41.1% rate per year. Yi Jin Huo Luo Qi, on the other hand, appears to have developed a tertiary sector to complement its well-established mining industry. Yi Jin Huo Luo Qi is the main extractive area of Dongsheng coalfield, which is one of the most important coalfields in China.

Of those districts in Table 1 which are intensive in secondary industry activities, the extraction of natural resources features very strongly. Yanchuan, Korla, and Yumen are heavily dependent on petroleum extraction and refining. In Yanchaun, for example, the Yanchuan Petroleum Company is the district’s largest tax revenue contributor, accounting for 65% of local tax revenue. In 2006 Yanchuan’s gross industrial output was twice that of the previous year. A similar story is evident in Korla: its total Gross Industrial Output Value has grown by over 25% per annum between 2000 and 2005. Korla’s secondary industry share of total GDP is approximately 79%, but this figure falls to 27% when the oil sub-sector is excluded. In Yumen, 60% of local government revenue emanates from petroleum exploiting and refining, which contributes over 36,000 job to the district (61.7% of the district’s total employment). Zhungeer is a coal mining district, which in 2005 experienced 45% growth in its industrial value-added as a result of higher prices for coal products. Asbestos extraction and production is the main industry of Akesaihasake and accounts for 90% of local government revenue. Akesaihasake’s annual asbestos production is 170,000 tons, accounting for over 50% of national asbestos production. As well as asbestos, the district is rich in other minerals and metals, such as Gold, Zinc, and Crystal. Etuoke, traditionally an agricultural district, has in recent times focussed on attracting manufacturing investment. Since 2000 two industrial areas have been constructed in the district. Many industrial companies such as Mengxi Limited, Xingguang Limited have established branches or factories in these industrial areas. In 2005, Etuoke’s GDP rose to 600 million RMB, on the back of a 44% annual growth rate since 2000.

Taken as a whole, our dataset indicates that classifications based on provincial data are inadequate in that they conceal considerable heterogeneity within the provinces and may smooth the impact of important localised economic developments over larger economic units. The most important implication of this is that Chinese provinces may not be the optimal unit of regional analysis because aggregation leads to a distorted view of reality.[71] The West-Central-Coastal belts and the provinces

[71] One can draw an interesting parallel between China and Germany. While in China the discussion has been governed by the "three belt hypothesis", the discussion in Germany after unification was governed by the "two belt hypothesis" (eastern vs. western Germany). Comparable with our evidence for China, the German Council

appear to be inappropriate units for government policies of awarding preferential treatment to specific regions. Although we still have to investigate the mechanisms that underlie the observable uneven patterns of GDP per capita, we can conclude that a large regional variance below the provincial level has been averaged away in many conventional studies. This paper, representing the first attempt to focus on county- and city-level data across the entirety of Mainland China, aims to address this deficit.[72]

4. From the Bottom Up: Non-Parametric Evidence on the Distribution of County- and City-Level GDP per Capita

Nonparametric techniques, such as the Kernel density estimator, can reveal interesting features of the data and therefore help to capture the stylised facts that need explanation. Such techniques allow one to ascertain the distribution of the underlying data without imposing any parametric restrictions: "letting the data speak for itself", as the old adage goes. In the case of our Chinese real GDP per capita data, such an approach is intuitively appealing given the large amount of county- and city-level observations available and the possibility of a number of distinct distributions or patterns being present in the underlying data.

4.1. Kernel Density Estimation and Real GDP per Capita

The kernel estimator for the density function $f(x)$ at point x is

$$(1) \qquad \hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right)$$

where $x = x_1, x_2, \ldots, x_n$, is an independent and identically distributed sample of random variables from a probability density $f(x)$ and $K(\cdot)$ is the standard normal kernel with window width h. The window width essentially controls the degree to which the data are smoothed to produce the kernel estimate. The larger the value of h, the smoother the kernel distribution. A crucial issue is the selection of this smoothing parameter. Here, the two-stage direct plug-in bandwidth selection method of Sheather and Jones (1991) is employed, which has been shown to perform quite well for many density types by Park

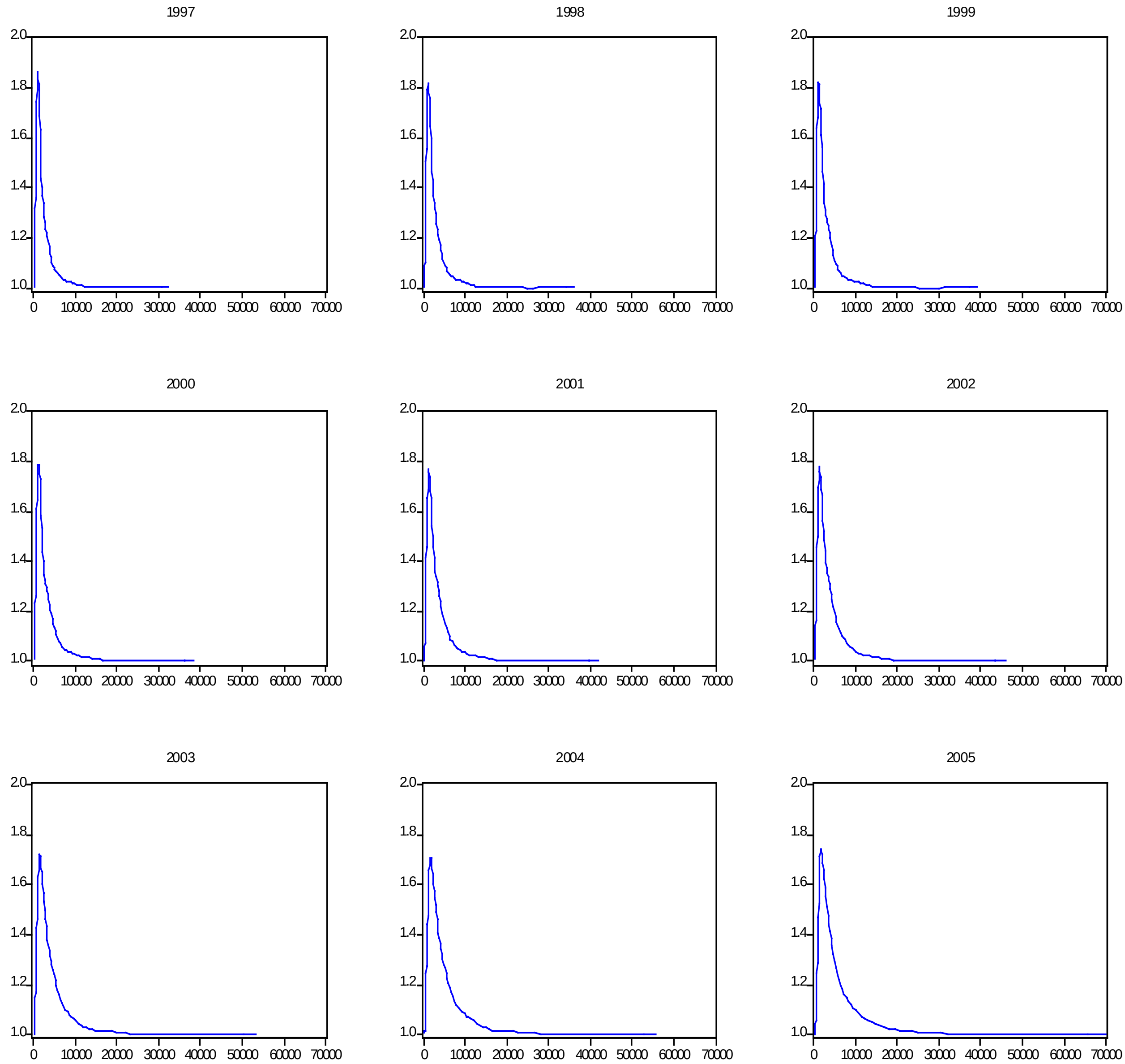
---

of Economic Experts [see Sachverständigenrat zur Begutachtung der Gesamtwirtschaftlichen Entwicklung (1999), pp. 116-133] has demonstrated that the German "belt view" is only correct superficially. On the surface it appears as if such a distinction exists, but in reality the situation is different. 10 years after German unification several prosperous counties and cities exist in eastern Germany and therefore the two belt hypothesis is inadequate as a guide for regional economic policies.

[72] The existing literature advancing to this level of disaggregation has focussed on case studies of provinces. See, for example, Lyons (1998). So far, there are no national cross-county and cross-city statistical analyses of Chinese regional development. Zhang et al. (2006, 2007) have recently attempted to disaggregate sub-provincial data in their analysis of real GDP per capita growth using a sample of 341 districts obtained from dividing sub-provincial regions into counties and municipalities

and Turlach (1992) and Wand and Jones (1995).[73] The distributions have been fitted to the logarithm of real GDP per capita. Figure 5 presents the kernel density estimations for the (log) *GDP* from 1997 to 2005 obtained using the abovementioned bandwidth selection method and by transforming the income variable to the original scale.

Figure 5: Kernel Density Estimates with Sheater and Jones Plug-in Bandwidth Selection Method



Note: The horizontal axes in the above kernel density estimates are given Rmb.

In the kernel density estimation context, a convergence process occurs if, for instance, a bimodal density is detected at the beginning of the sample period and over time there is a tendency in the

[73] Given the crucial role played by the bandwidth selection method, it is important to assess the performance of alternative bandwidth selectors. When the Silverman (1986) rule of thumb bandwidth selector has been used for the above kernel density estimation, similar trends are exhibited by the distributions. Detailed results are available from the authors on request.

distribution to move towards unimodality. Alternatively, if there already is a unimodal distribution at the beginning of the time span in question, convergence occurs when the dispersion of this density and therefore per capita income declines over time. The kernel density estimates of Figure 5 reveal a number of interesting features: firstly, there is clearly no multimodality present in the distribution – suggesting that there are not three distinct distributions or patterns in the county- and city-level Chinese real GDP per capita data. This confirms that talking about Chinese growth in terms of a Western-Central-Coastal division is overly simplistic. Secondly, the kernel density estimates are clearly skewed, with one high real GDP per capita generating region in particular, Shenzhen, visible as an outlier in the tail on the right.[74] Thirdly, the high GDP per capita regions appear to be pulling further away to the right over the 1997-2005 period.[75] It also seems that the mode of the distribution, representing the main body of the observations, has widened somewhat over time.

These visual impressions gained from the kernel density estimates also find some support from descriptive statistics of 1997 and 2005 real GDP per capita, given in Table 2. In 2005 both the maximum and minimum real GDP per capita are approximately twice their 1997 values. That said, the skewness decreases somewhat over this period – suggesting that the main body of observations move upwards in the distribution, as evidenced by the noticeable increase in the median value.

Table 2 : Summary Statistics for County- and City-level Real GDP per Capita, 1997 and 2005

	1997	2005
Mean (Rmb)	1,562.50	3,170.90
Median (Rmb)	1,142.50	2,148.00
Standard Deviation	1,555.78	3,472.48
Maximum (Rmb)	32,377.35	69,962.88
Minimum (Rmb)	158.53	310.00
Skewness	6.25	5.91
Kurtosis	87.18	77.76
Jarque-Bera	662,436.90	523,941.40

In all, the kernel density estimates and descriptive statistics convey the following message: there are some obvious movements in the tails of the distribution, with the highest and lowest GDP per capita districts exhibiting the clearest changes. But this should not overshadow the fact that the main body of

[74] Shenzhen, which forms part of the Southern China's Pearl River Delta region, has experienced phenomenal growth since being designated as a Special Economic Zone in 1979. While it was initially associated with labour-intensive industries, since the 1990s Shenzhen has focused on the manufacture of electronics, attracting substantial technology-based investment flows from Hong Kong, Taiwan, Japan, Europe and the United States. In 2002 Shenzhen accounted for around 20% of Mainland China's computer production and 15% of its semiconductor integrated circuits, according to Enright et al. (2005, pp. 47-49). Shenzhen's population rose from 321,000 in 1979 to more than seven million in 2000. Shenzhen Goverment Online reports that in 2004 average per capita GDP in Shenzhen was the highest in China, while its total import and export volume accounted for 1/7th of the country's total and ranked first in China for 12 consecutive years. Container throughput in the city ranked second in China and fourth worldwide (see: http://english.sz.gov.cn/)

[75] One data limitation arises from gaps in population data due to this data still being based upon the so-called "hukou system". This leads to distortions in regions with large inflows of workers who should be counted as part of that region's regular population from an economic point of view.

observations appear to have moved to a position of higher real GDP per capita over the 1997-2005 period. In the next sub-section, the visual impressions conveyed by the kernel density estimates are further probed using transition probability matrices and stochastic kernel estimates.

4.2. Distribution Dynamics of Chinese Real GDP per Capita

While the kernel density estimates in Figure 5 provide snapshots of the entire distribution of Chinese real GDP per capita as it evolves over time, it may well be the case that the skewness of the kernel density estimates conceals a convergence process among those central and western districts which were seen to enjoy such strong growth in Figures 3 and 4. Additional techniques are required to uncover the movements of the individual districts over time. This underlying process is further examined by considering the intra-distributional dynamics of the observations over the 1997-2005 period. This involves modelling directly the evolution of relative real GDP per capita distributions by constructing transition probability matrices that track changes over time in the relative position of districts within the distribution. This is an exercise that a number of authors have undertaken (see Quah, 1996a, 1996b). The modelling of distribution dynamics assumes that the density distribution ϕ_t has evolved in accordance with the following equation:

(1) $$\phi_{t+1} = M\,\phi_t,$$

where M is an operator that maps the transition between the income distributions for the periods t and t+1. Since the density distribution ϕ for the period t only depends on the density ϕ for the immediately previous period, this is a first-order Markov process. The controlling factor in a Markov chain is the transition probability, i.e. a conditional probability for the system to go to a particular new state, given the current state of the system. The maximum-likelihood estimate of the transition probabilities can be expressed in the form

(2) $$\hat{P}_{ij} = \frac{n_{ij}}{n_i}$$

where n_{ij} is the number of districts that were in income category i in the previous period and have migrated to income category j in the current period, and n_i is the total of districts that were in income category i in the previous period. In other words, the estimate equals the proportion of time that the process, after leaving state i, next enters state j.

The main advantage of the transition matrix is that it allows one to summarise the random ups and down of regional fortunes in a handful of numbers. The transition probability matrix in Table 3 reports transitions between the 1997 and 2005 distributions of the real GDP per capita relative to the median value. The main diagonal of the matrix gives the proportion of districts that remained in the same range of the distribution throughout the period in question, while the off-diagonal probabilities are those associated with moving from one state to another. Table 3 also provide information about *n*, the number of districts that begin their transitions in a given state. Furthermore, the classes that divide up the state space are provided.

Table 3: Transition Probability Matrix Relative to the Median Real GDP per Capita

	GDP PER CAPITA 2005							
	n		*405*	*447*	*346*	*245*	*169*	*583*
	372	[0-0.50]	77.42	18.01	3.23	1.08	0.00	0.27
GDP PER CAPITA 1997	*435*	[0.50-0.75]	19.77	51.03	21.15	4.83	0.92	2.30
	391	[0.75-1.00]	5.63	29.16	34.78	17.90	8.95	3.58
	298	[1.00-1.25]	2.01	10.74	22.48	30.87	17.79	16.11
	169	[1.25- 1.50]	0.59	4.73	12.43	19.53	21.30	41.42
	530	[1.50- ∞]	0.38	0.75	3.40	4.72	7.74	83.02
			[0-0.50]	[0.50-0.75]	[0.75-1.00]	[1.00-1.25]	[1.25- 1.50]	[1.50- ∞]

The transition probability matrix in Table 3 reveals a number of noteworthy behavioural patterns in the distribution of real GDP per capita over time. It is clear from the probabilities that lie along the diagonal that some states are more susceptible to movement than others. Districts in the lower two states and those in the highest state appear to be relatively more static, as their probability of staying put is quite high. These large diagonal entries at the beginning and end of the distribution are consistent with the Markov chain analysis in Bhalla et al. (2003), who have used provincial-level data.
However, the districts residing in the middle of the distribution appear to be far more mobile. These middle states exhibit a greater degree of shuffling between relative categories. Both the diagonal and off-diagonal probabilities for the third, fourth, and fifth states suggest large potential for movement – in both forward and backward directions.

In Table 3, the operator *M* has been constructed by assuming that the distribution ϕ_t has a finite number of states. This discrete modelling approach leads to the problem that the researcher has to determine the number of intervals and the limit values of each interval in an arbitrary and ad hoc way. Furthermore, the discretisation process may eliminate the property of Markovian dependence in the data, as Bulli (2001) has pointed out. The solution which addresses these shortcomings consists of carrying out a continuous analysis of transition, which avoids discretisation through the use of conditional densities that are estimated non-parametrically and referred to as stochastic kernels. A stochastic kernel amounts to a transition matrix with an infinite number of infinitely small ranges. The results from this tool are displayed as three-dimensional graphs in Figure 6 and a two-dimensional contour map in Figure 7.

Figure 6: Stochastic Kernel Estimates, 1997-2005

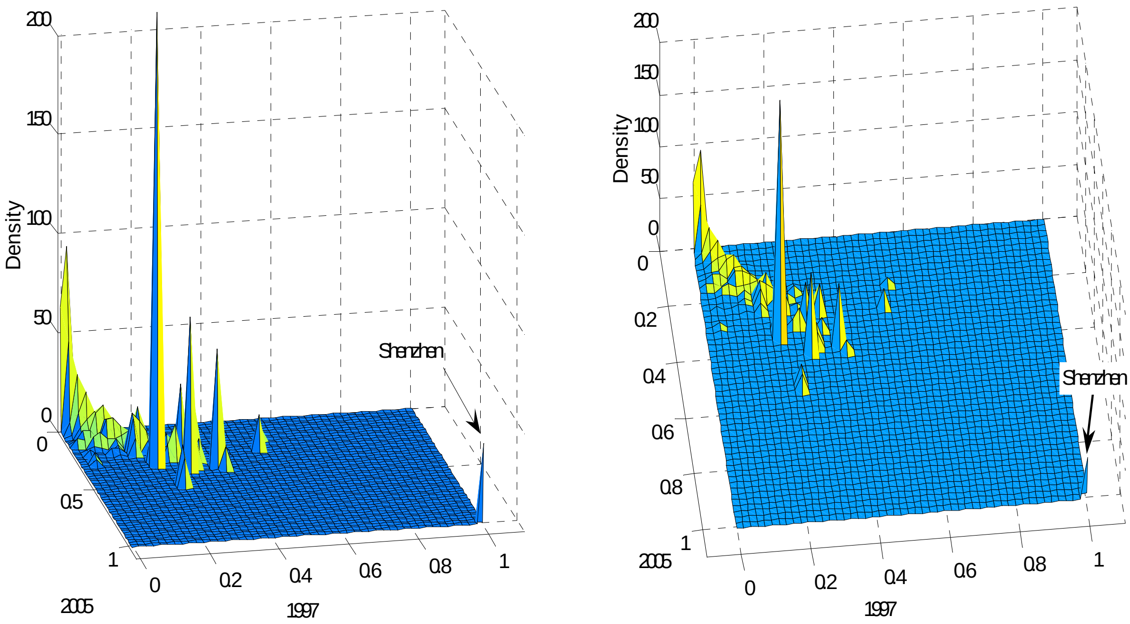
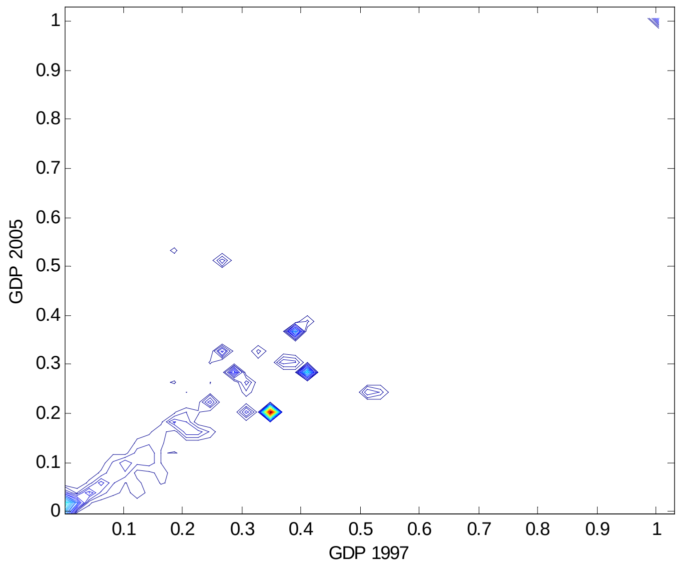


Figure 7: Stochastic Kernel Contours, 1997-2005



Note: In Figures 6 and 7 the district with the highest GDP per capita has been used as a numeraire. Scaling real GDP per capita relative to the median value has also been explored but yielded the same results. The outlier, Shenzhen, is indicated in Figure 6 with an arrow.

The three-dimensional stochastic kernel estimates of Figure 6, together with the associated stochastic kernel contour of Figure 7, tackle some of the shortcomings of the transition probability matrices as well as reiterating the main findings of the previous sections. In order to fully exploit the information content of this construct we adjust the perspective by tilting the graph downwards, as if looking down on the three dimensional distribution from above. This "aerial view" is further enhanced by means of a contour image of the distribution. The graphs in Figure 6 highlights two features: the pronounced peaks at the beginning and very end of the distribution (the outlier, Shenzhen); and middle section of the distribution which, apart from a few spikes, is relatively lower and suggests the possibility of either slippage or enhancement of one's relative position. The contour image in Figure 7, highlights the diagonal pattern of the distribution over time, illustrating the tendency of districts residing at the extremities of the distribution in 1997 to remain there in 2005. The districts in the middle, however, clearly scatter from this diagonal pattern, with this off-diagonal movement supporting the findings of the transition probability matrix which pointed to a clear tendency for movement amongst the districts residing in the middle of the distribution over the 1997-2005 period.

5. Empirical Growth Regressions

After establishing both the spatial disparity and the underlying distribution of Chinese county- and city-level real GDP per capita, we now set about identifying those factors which may explain the trends observed in Chinese regional growth over the 1997-2005 period. We expand our original dataset by introducing a wide range of explanatory variables and we then estimate a growth equation using the now-standard Barro (1991) framework, which tests for conditional β-convergence by incorporating a set of explanatory variables reflecting differences in the steady-state equilibrium. Despite constraints stemming from data availability our expanded dataset covers a broad spectrum of economic and demographic factors.

The explanatory variables introduced to our analysis address a number of key features which have emerged from the literature as being influential in the economic growth process. Foremost amongst these are education and human capital formation, which are necessary to raise productivity. Investment in education leads to the acquisition of skills that improve efficiency through the better use of technologies. Education also reduces the imitation lag. With this in mind, we include the secondary-level education enrolment rate (expressed as a percentage of population of a given district) in our growth equation.[76] Another factor widely regarded as influential in modern-day Chinese regional growth is the substantial inflow of overseas investment the country has attracted. China has attracted foreign direct investment (FDI) as part of a concerted development strategy.[77] The resulting dramatic expansion of FDI has allowed China to reap growth-enhancing benefits from FDI in several areas. First, the opening up of the economy has contributed to the acceleration of growth because increasing efficiency hinges on the implementation of new technologies, managerial skills, and labour training.

[76] China already had a high literacy rate prior to the beginning of our sample period. The 1986 Compulsory Education Law increased mandatory education from five years to nine. According to official estimates, 93% of the country had achieved nine-year basic education in 2004.

[77] The opening up of the Chinese economy began in 1979 with the promulgation of the Chinese-Foreign Joint Venture Law. A new phase in the reform process began in 1992 when FDI was allowed in all major inland cities. The open-door policy eventually led to a surge in FDI, making China the largest single FDI recipient in the world.

Second, FDI has increased China's export competitiveness. Third, FDI helped to broaden the knowledge of Chinese authorities about market mechanisms during the transition process.[78] To capture this Chinese FDI phenomenon, utilised foreign capital (expressed as a percentage of GDP in a given district) is also included in our growth equation specifications. A related issue is the extent to which GDP composition influences regional economic performance. To explore this further, the proportion of each district's agricultural GDP and secondary industry GDP (both expressed as a proportion of that district's total GDP) are included in our specifications. As in Section 3, primary industry refers to farming, forestry, animal husbandry and fishing, while secondary industry includes mining, manufacturing, electricity production, and construction.

A further factor thought to be influential in Chinese regional growth has been disparities in infrastructure networks across Chinese districts. Démurger (2001), for example, has demonstrated that transport facilities are a key differentiating factor in explaining regional growth disparities. Of course, there are many types of infrastructure and measures thereof, and any one measure can only capture part of the story. However, a measure that gets at the essence of the infrastructure problem will presumably be highly correlated with other measures. We incorporate this infrastructure disparity into our growth equation specification with the inclusion of the number of hospital beds per capita in each district. Traditionally, health care was provided by state-owned enterprises. Reform, however, has severely disrupted this system, as market pressure has led many firms to abandon their social services. This has led to the development of an uneven healthcare net across China.

Various new economic geography, new trade theory, and endogenous growth models have been applied to highlight the nexus between geographic location and economic growth. Conclusions emanating from this line of inquiry are: (i) landlocked regions and countries trade less vis-à-vis coastal regions or countries, and (ii) coastal regions and maritime countries experience on average higher growth than landlocked regions and countries.[79] In order to consider the influence of geographic location in China's regional growth, we construct a set of dummy variables which indicate whether districts lie within the western, central, or coastal belts of China, as well as dummy variables which indicates the proximity of airports and seaports to the districts under observation. Figures 8 and 9 map the location of these airports and seaports, as well as categorising each airport by the passenger flows it caters for. As illustrated in Figures 8 and 9, we have identified all the airports and seaports across mainland China, matched them to their respective county- and city-level districts, and highlighted the neighbouring regions that are likely to benefit from proximity to these transport facilities. We have also classified China's airports in terms of their passenger flow, with the number of airports in each classification provided. The airports are categorised as follows: 1 = airports with 0-50,000 passengers

[78] Wen (2007) has investigated the mechanisms whereby FDI has contributed to China´s regional development.

[79] The specifics of the trade and growth umbrella is one of the greatest puzzles in the economics profession. Studies this phenomenon include Sachs and Warner (1995) and Edwards (1992, 1998). Similarly, Vamvakidis (2002) has demonstrated in a historical context that trade is associated with growth after 1970 but not before. Another strand of the openness and growth literature seeks to improve cross-country regressions by employing panel methods, geared at controlling for time-invariant unobservable fixed effects. Wacziarg and Welch (2003) provide evidence for a strong effect of openness on growth in a panel set-up.

or no available passenger data; 2 = airport with 50,000-1,000,000 passengers; 3 = airport with 1,000,000-1,500,000 passengers; 4 = airport with over 1,500,000 passengers.

A number of caveats should be noted before we proceed with the estimation of growth equation regressions using our county- and city-level data. From a methodological perspective, one weakness of cross-region regressions is that of reverse causality and endogeneity. We have used regressors dating from 2000, due to the fact that most county and city-level data are not available for the years 1997-1999. These regressors dating from 2000 are assumed here to be weakly exogenous, thus obviating the need for instrumental variable techniques. Furthermore, the existing empirical growth literature using "Barro-regressions" has been criticized for its lack of robustness. Durlauf and Quah (1999) and Temple (1998, 2000) stress that applied macroeconomists are inclined to follow theory rather loosely and simply try variables to establish factors determining economic growth. In these empirical specification searches, econometric problems such as robustness are often ignored [Durlauf (2001)].

In order to both shed further light on the robustness issue, and to make our cross-region estimates more sensible in face of the common pitfalls stemming from OLS estimation, we use the *L*east *T*rimmed *S*quares (LTS) estimator and the *B*ounded *I*nfluence *F*unction Regression (BIF) estimator as specification devices and diagnostic tools.[80] The LTS estimator is very similar to OLS, the only difference being that the largest squared residuals are not used in the summation, thereby allowing the fit to avoid the outliers. In other words, the LTS estimator searches for a core subset of data that follows best a certain model without taking into account the rest of observations. The LTS estimator is $\sqrt{n}$-consistent and asymptotically normal. With k unknown parameters the LTS method attains the highest possible breakdown value, namely $\{[(n-k)/2]+1\}/n$ which asymptotically equals 50 percent, i.e. it can withstand a lot of bad leverage points occurring anywhere in the data.[81]

The BIF estimator is a robust estimator proposed by Krasker et al. (1983).[82] The purpose of the method is to attribute a lower weight to the impact of potentially influential observations using a chosen weighting function. The estimator is constructed by means of the so-called influence function measuring the impact of outlying observations. Optimal choices for the weighting function have been suggested by Hampel et al. (1986). Rousseeuw and Leroy (1987) have demonstrated in experiments that the BIF estimator achieves a breakdown point of slightly above 30 percent.

[80] For an excellent survey of robust estimation methods and applications, see Rousseeuw and LeRoy (1987).

[81] In order to obtain the LTS regression a large number of subsamples, each of size k (the number of regression coeffcients, including the constant term) has to be drawn and evaluated. In this paper 3000 subsamples have been drawn.

[82] The BIF estimator is also referred to as the generalised M-estimator (GM-estimator) in the literature.

Figure 8: Airports in Mainland China, Categorised by Passenger Flows

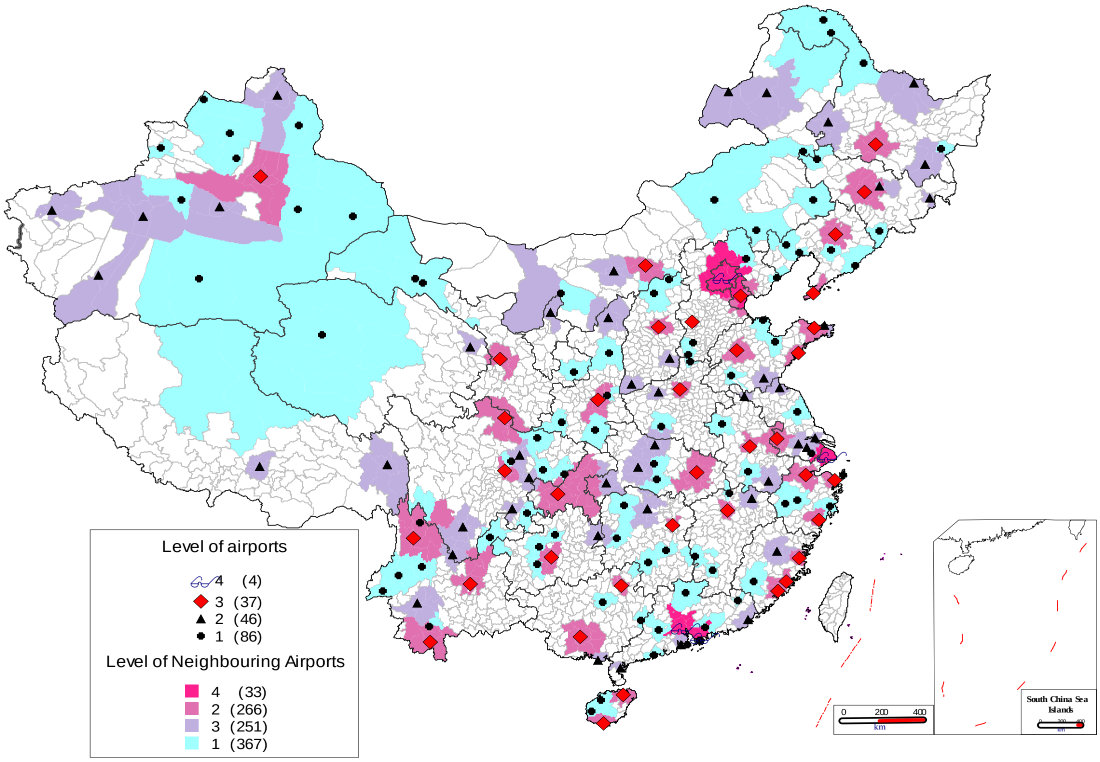
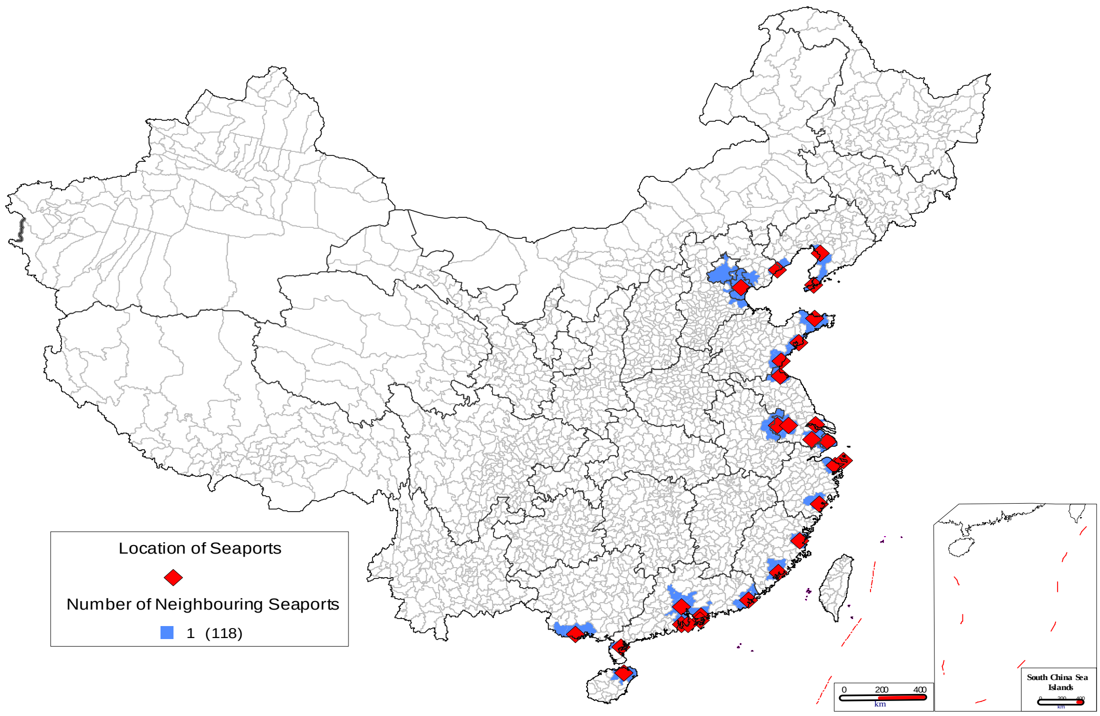


Figure 9: Location of Seaports in Mainland China



Tables 4-7 below present OLS, LTS, and BIF linear regressions of average county- and city-level GDP per capita growth on initial, 1997, log GDP per capita *(lnGDP per capita$_{1997}$)* and the selection of explanatory variables detailed above – the standard test for conditional β-convergence, where a negative significant coefficient on initial log GDP indicates convergence and a positive significant coefficient indicates divergence. Tables 4 and 5 use the full 2,199 observations of the dataset, while Tables 6 and 7 are based on a sample of 1,150 observations due to constraints on the availability of utilised foreign capital data.[83] Tables 5 and 7 introduce the dummy variables into our specifications.[84]

Table 4: OLS, LTS, and BIF Estimation of Growth Equations

Dependent Variable: Growth of real GDP per Capita (1997-2005)			
	OLS	LTS	BIF
Constant	0.228	0.197	0.202
	(0.00)	(0.00)	(0.00)
lnGDP per capita$_{1997}$	-0.023	-0.021	-0.023
	(0.00)	(0.00)	(0.00)
Hospital Beds/ pop (%)	0.036	0.027	0.035
	(0.00)	(0.00)	(0.00)
Secondary Enrolment / pop (%)	0.000	-0.001	0.000
	(0.57)	(0.15)	(0.76)
Primary GDP/total GDP (%)	-0.001	-0.0003	-0.0002
	(0.00)	(0.00)	(0.03)
Secondary GDP/total GDP (%)	0.001	0.001	0.001
	(0.00)	(0.00)	(0.00)
Adjusted R^2	0.13	0.18	0.16
Total no. of observations	2,199	2,112	2,199
No. of outliers	--	87	109

Note: prob-values are given in brackets.

[83] A breakdown of districts bearing missing values for utilised foreign capital is as follows (total number of districts in each province is provided in brackets): *Coast*: Beijing 0 (3), Tianjin 0 (4), Liaoning 8 (58), Hebei 29 (147), Shanghai 0 (2), Jiangsu 1 (65), Zhejiang 8 (69), Fujian 2 (67), Shandong 4 (108), Guangdong 10 (88), Hainan 12 (18), and Guangxi 56 (89); *Central*: Shanxi 64 (93), Jilin 20 (49), Heilongjiang 42 (77), Anhui 30 (78), Jiangxi 16 (89), Henan 64 (126), Hubei 24 (76), Inner Mongolia 54 (88), and Hunan 21 (100); *West*: Sichuan 99 (105), Chongqing 15 (26), Guizhou 70 (82), Yunnan 103 (123), Tibet 73 (73), Shaanxi 64 (93), Gansu 71 (81), Qinghai 39 (40), Ningxia 9 (16), and Xinjiang 73 (86).

[84] Dummy variables cannot be included as regressors in the BIF estimation procedure.

Table 5: OLS and LTS, and BIF Estimation of Growth Equations

Dependent Variable: Growth of real GDP per Capita (1997-2005)			
	OLS	LTS	BIF
Constant	0.239	0.182	0.202
	(0.00)	(0.00)	(0.00)
lnGDP per $capita_{1997}$	-0.024	-0.015	-0.023
	(0.00)	(0.00)	(0.00)
Hospital Beds / pop	0.018	0.001	0.022
	(0.02)	(0.58)	(0.01)
Enrolment Sec/ pop (%)	0.000	-0.001	0.001
	(0.79)	(0.19)	(0.11)
Utilised Foreign Cap/ GDP (%)	0.022	0.082	0.067
	(0.09)	(0.00)	(0.02)
Primary GDP/total GDP (%)	-0.001	-0.001	-0.0003
	(0.00)	(0.00)	(0.04)
Secondary GDP/total GDP (%)	0.001	0.0003	0.001
	(0.00)	(0.00)	(0.00)
Adjusted R^2	0.11	0.11	0.18
Total no. of observations	1,150	1,084	1,150
No. of outliers	--	66	58

Note: prob-values are given in brackets.

Table 6: OLS and LTS Estimation of Growth Equations Including Foreign Capital

Dependent Variable: Growth of real GDP per Capita (1997-2005)				
	OLS		LTS	
	(1)	(2)	(3)	(4)
Constant	0.229	0.229	0.207	0.199
	(0.00)	(0.00)	(0.00)	(0.00)
lnGDP per $capita_{1997}$	-0.023	-0.023	-0.021	-0.023
	(0.00)	(0.00)	(0.00)	(0.00)
Hospital Beds/ pop (%)	0.036	0.036	0.018	0.042
	(0.00)	(0.00)	(0.00)	(0.00)
Secondary Enrolment / pop (%)	0.001	0.000	0.001	0.001
	(0.58)	(0.58)	(0.05)	(0.16)
Primary GDP/total GDP (%)	-0.001	-0.001	-0.001	-0.0002
	(0.00)	(0.00)	(0.00)	(0.01)
Secondary GDP/total GDP (%)	0.001	0.001	0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Coast	0.001	0.001	-0.002	0.001
	(0.71)	(0.70)	(0.39)	(0.75)
Central	-0.001	-0.001	-0.003	-0.004
	(0.61)	(0.61)	(0.05)	(0.03)
Seaports		-0.002		0.018
		(0.81)		(0.01)
Airports		0.000		0.001
		(0.91)		(0.80)
Adjusted R^2	0.12	0.12	0.17	0.17
Total no. of observations	2,199	2,199	2,105	2,106
No. of outliers	--	--	94	93

Notes: prob-values are given in brackets. Western China is the base region for the "three-belt" geography dummies.

Table 7: OLS and LTS Estimation of Growth Equations Including Foreign Capital

Dependent Variable: Growth of real GDP per Capita (1997-2005)		
	OLS	LTS
Constant	0.240	0.177
	(0.00)	(0.00)
lnGDP per capita$_{1997}$	-0.024	-0.020
	(0.00)	(0.00)
Hospital Beds / pop	0.015	0.006
	(0.07)	(0.39)
Enrolment Sec/ pop (%)	0.000	0.001
	(0.76)	(0.27)
Utilised Foreign Cap/ GDP (%)	0.022	0.136
	(0.08)	(0.00)
Primary GDP/total GDP (%)	-0.001	0.000
	(0.00)	(0.13)
Secondary GDP/total GDP (%)	0.001	0.001
	(0.00)	(0.00)
Seaports	-0.002	0.015
	(0.79)	(0.02)
Airports	0.003	0.001
	(0.47)	(0.77)
Adjusted R^2	0.11	0.17
Total no. of observations	1,150	1,074
No. dropped due to LTS	--	76

Note: prob-values are given in brackets.

The econometric evidence provided in Tables 4-7 offers a number of insights into the determinants of county- and city-level regional growth over the 1997-2005 period. First, there is unanimous evidence across all the estimators and specifications utilised that a convergence process has occurred over the 1997-2005 period.[85] The OLS, LTS, and BIF estimates of the *lnGDP per capita$_{1997}$* coefficient lies between -0.015 and -0.023 in each specification, indicating a speed of convergence of approximately 2% per annum.[86] The R^2 for our specifications is quite low, ranging from 0.11-0.18. This, however, is common amongst studies estimating cross-sectional growth equations for Chinese regions. Jones et al. (2003, pp. 194-197), for example, report R^2 values ranging from 0.12-0.22. Second, hospital beds per capita appears to be positively significant in the majority of specifications - reiterating the contribution of infrastructure networks in the regional growth process. The secondary-level education enrolment rate, however, does not yield a significant coefficient in the majority of specifications. This suggests that, in the Chinese case, further analysis may be required in order to identify the channels through which secondary-level education contribute to the economic performance of that district where the secondary-level education is actually obtained. The positive significant estimates for utilised foreign

[85] When using the LTS estimator, it is important to examine the observations identified as outliers. Inspection of the outliers reveals that they are not concentrated in one region. The western provinces of Shaanxi, Gansu, and Xinjiang do contain outliers, but no more than the central regions of Inner Mongolia, Shanxi, and Heilongjiang. The coastal provinces of Jiangsu, Guangdong, and Hebei also produce outliers. A complete list of outliers for each LTS specification is available from the authors on request.

[86] The LTS and BIF estimators identify a similar set of outliers, as evidenced by their correlation coefficient of 0.7. The BIF outliers are detected using the Studentized Residuals method. See Judge et al. (1988) for a summary of outlier detection methods.

capital are as one would expect, reflecting the investment-led nature of Chinese regional economic growth over the 1997-2005 period. Regarding the issue of industry composition, it appears that a propensity for secondary GDP production exerts a positive influence on the growth of county- and city-level districts, while being an intensive producer of primary GDP exerts a negative influence.

The dummy variables included in our regression specifications offer insights into the influence of geographic location and the importance of proximity to transport facilities in China's regional growth process. The inclusion of airport and seaport dummies allows us to distinguish between passenger flows, on one hand, and the import and export of raw materials and finished goods on the other. As illustrated in Figures 8 and 9, China's geography dictates that its seaports are located exclusively on the eastern and south eastern coast. Its airports appear to be spread across the entire mainland, though there is clearly a higher concentration in the coastal and central regions. All the airports catering for passenger flows greater than 1.5 million passengers are located close to major seaports on the coast, while all airports catering for 1-1.5 million passengers are located either in the coastal or western provinces. According to the coefficients estimated in Tables 6 and 7, it is the seaports dummy variable that is the influential one in terms of regional growth over the 1997-2005 period, at least in our LTS regressions. This dummy variable may be capturing the role of China's massive intake of raw materials, coupled with its surging export outflow, over the last decade as a key driver of its economic success. The insignificance of the airports dummy variable in Tables 6 and 7 may be a related issue. The inflow and outflow of passengers through regional airports may not have fuelled China's regional growth to anything near the extent of its seaports.

The dummy variables indicating whether districts lie within the western, central, or coastal belts of China appear to be broadly indicative of a convergence process. Relative to the western belt (the base region), estimated coefficients for the central belt are negatively significant, while estimated coefficients for the coast do not appear to be significant. This suggests that the western regions, perceived generally as lagging behind in terms of growth, experienced faster growth than their central counterparts and a rate of growth not significantly different to coastal districts. This western catch-up occurs in spite of the coast's many economic advantages, such as preferential economic policies enjoyed by the coast's Special Economic Zones and its prime location for international trade. This finding lends further support to the impression created by Figures 3 and 4 that pockets of high growth are not confined to coastal region, but permeate the entirety of mainland China.

6. Conclusions

Having employed an array of complementary techniques to analyse the development of Chinese county- and city-level real GDP per capita over the 1997-2005 period, it now remains to collect the various findings and identify the coherent trends which emerge. The opening salvo of this paper is the observation that much of the existing literature investigating Chinese regional growth has focused on large geographic units which are unsuitable for that purpose. This paper introduces a new dataset comprising of county- and city-level data that spans the entirety of mainland China over the 1997-2005 period. In this way we hope to uncover the stylised facts of Chinese regional growth dynamics, which will be of use to both policymakers and academics alike. The colour-coded maps of Section 3 provide a vivid illustration of the enhanced level of detail available when one moves from provincial-level data to county- and city-level data: when one compares the growth performance of large geographical units, such as provinces, to that of the districts which form these units, the full extent of GDP per capita disparities within provinces becomes patently clear. What is more, it becomes apparent that pockets of high GDP per capita districts permeate the entirety of mainland China rather than being confined to a certain "belt " or province. We delve further into this county- and city-level disaggregation by ascertaining the evolution of the entire cross-sectional GDP per capita distribution using non-parametric kernel density estimation. Visual inspection indicates that the distribution is characterised by outliers in the upper tail and a main body of districts which exhibit real GDP per capita growth over the 1997-2005 period. In order to track the dynamics of each individual district over time we rely on the transition probability matrix technique and the associated stochastic kernel estimation. What emerges from this exercise is a picture of relatively static districts at both tails of the distribution, but large potential for movement amongst the districts in the middle of the distribution. In light of this fluidity amongst the districts in the middle of the distribution, we ask the question: is there evidence of a convergence process at work across Chinese county- and city-level districts? This question is answered unequivocally in the affirmative with the conditional β-convergence regressions of Section 5. Our OLS, LTS, and BIF linear regressions of average county- and city-level GDP per capita growth on initial log GDP per capita, and a set of explanatory variables, yield estimates of the log GDP per capita coefficient that lie between -0.015 and -0.023 in each specification. This indicates a speed of convergence of approximately 2% per annum.

The explanatory variables included in our conditional β-convergence regressions offer an opportunity to pinpoint influential factors in the regional growth process. The significance of hospital beds per capital, capturing disparities in the infrastructure network at district level, is supportive of findings emanating from provincial-level studies. The insignificance of secondary-level education enrolment, on the other hand, may come as a surprise but does raise questions regarding the ability of a given district to capture the benefits of the secondary-level education provided by that district. The significance of utilised foreign capital in the regional growth process comes as no surprise given the concerted efforts of China's policymakers to attract FDI inflows. The fact that the proportion of secondary industry GDP that a district generates positively influences its growth rate may indicate that secondary industry-intensive districts have proved to be the more fertile locations for these investment inflows, while districts predominantly dependent on the primary sector (represented by primary GDP in our regression specifications) get left behind. Geographic location has long been regarded as a key factor in the regional growth process. The set of dummy variables we have constructed appears to confirm this finding in the case of Chinese county- and city-level districts. The inclusion of airport and seaport dummies allows us to distinguish between passenger flows, on one hand, and the import and export of raw materials and finished goods on the other. Given China's massive intake of raw materials, coupled with its surging export outflow, it is understandable that the seaports dummy appears to be the more influential one over the 1997-2005 period. The West-Central-Coastal dummy variables appear to be broadly indicative of a convergence process. Relative to the western belt (the base region), estimated coefficients for the central belt are negatively significant, while estimated coefficients for the coast do not appear to be significant. This suggests that the western regions, perceived generally as lagging behind in terms of growth, experienced faster growth than their central counterparts and a rate of growth not significantly different to coastal districts.

Taken as a whole, these findings provide much food for thought for Chinese policymakers. At the time of writing, there appears to be a cohort of districts that are persistently poor, a core for districts that whose ability to generate GDP per capita growth seems to be variable, and a clique of affluent districts. That said, there does appear to be a catching-up process at work across districts. What is more, pockets of high growth districts appear to be spread across the entirety of mainland China. Further good news is the strong performance of some western and central districts over the 1997-2005 period. Having established current state of affairs facing Chinese districts, the gauntlet must now be thrown down before the policymakers: after years of preferential economic policies for the chosen regions and neglect for the rest, can the fledgling catch-up process identified in this paper be properly cultivated?

References:

Bao, S., Chang, G.H., Sachs, J.D. and W.T. Woo (2002) "Geographic Factors and China´s Regional Development Under Market Reforms, 1978-1998", *China Economic Review* 13, 89-111.

Barro, R.J. (1991) "Economic Growth in a Cross-Section of Countries", *Quarterly Journal of Economics* 106, 407-443.

Bhalla, A., Yao, S. and Z. Zhang (2003) "Regional Economic Performance in China", *Economics of Transition* 11, 25-39.

Bulli, S. (2001) "Distribution Dynamics and Cross-Country Convergence: A New Approach", *Scottish Journal of Political Economy* 100, 223-251.

Dayal-Gulati, A. and A.M. Husain (2002) "Centripetal Forces in China´s Economic Takeoff', *IMF Staff Papers* 49, 364-394.

Démurger, S. (2001) "Infrastucture Development and Economic Growth: An Explanation for Regional Disparities in China?", *Journal of Comparative Economics* 29, 95-117.

Démurger, S., Sachs, J., Woo, T.W., Bao, S., Chang, G. and A. Mellinger (2002) "Geography, Economic Policy and Regional Development in China", *Asian Economic Papers* 1, 146-197.

Durlauf, S.N. (2001) "Manifesto for Growth Econometrics", *Journal of Econometrics* 100, 65-69.

Durlauf, S.N. and D. Quah (1999) "The New Empirics of Economic Growth", in: Taylor, J.B. and M. Woodford (eds.) *Handbook of Macroeconomics* Volume 1A, Amsterdam (North-Holland), 235-308.

Edwards, S. (1992) "Trade Orientation, Distortions, and Growth in Developing Countries", *Journal of Development Economics* 39, 31-57.

Edwards, S. (1998) "Openness, Productivity, and Growth: What do We Really Know?", *Economic Journal* 108, 383-398.

Enright,M. J., Scott, E. E., and K. Chang (2005) *Regional Powerhouse: The Greater Pearl River Delta and the Rise of China*, Singapore (John Wiley and Sons).

Fogel, R.W. (2006) "Why China is Likely to Achieve its Growth Objectives", *NBER Working Paper No.* 12122, Washington.

Gao, F. (2004) "The inequality of the county economic development of Inner Mongolian and the contribution to the regional economic growth", *Journal of the Inner Mongolian Finance and Economics College*, 2004(3), pp115-117

Hampel. F.R., Ronchetti, R., Rousseaw, P.J. and W. Stahl (1986) *Robust Statistics: The Approach Based on Influence Functions*, New York (Wiley).

Herrmann-Pillath, C., Kirchert, D. and J. Pan (2002) "Disparities in Chinese Economic Development: Approaches on Different Levels of Aggregation", *Economic Systems* 26, 31-54.

Jones, D.C., Li, C. and A.L. Owen (2003) "Growth and Regional Inequality in China During the Reform Era", *China Economic Review* 14, 186-200.

Judge, G. G., Hill, R. C., Griffiths, W. E., Lütkepohl H., and T. C. Lee (1988) *Introduction to the Theory and Practice of Econometrics*, Wiley Series in Probability and Mathematical Statistics, Second Edition. pp. 887-896.
Krasker, W. S., Kuh, E. and R.E. Welch (1983) "Estimation For Dirty Data and Flawed Models", in: Griliches, Z. and M.D. Intriligator (Eds) *Handbook of Econometrics, Vol. 1*, Amsterdam (North-Holland), 652-698.
Lyons, T.P. (1998) "Intraprovincial Disparities in China: Fujian Province," 1978-1995, *Economic Geography* 74, 405-432.

Luo, J. (2004) "The Spatial Difference in Economic Development Level among the West Regions and Rapid Development", *Journal of Chongqing Technology and Business University (West Economic Forum)*, 2004, 6-9.

Park, B.U. and B.A. Turlach (1992) "Practical Performance of Several Data Driven Bandwidth Selectors", *Computational Statistics* 7, 251-285.

Quah, D. (1996a) "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics", *The Economic Journal* 106, 1045-1055.

Quah, D.T. (1996b) "Convergence Empirics Across Economies with (some) Capital Mobility", *Journal of Economic Growth* 1, 95-124.

Rousseeuw, P.J. and Leroy, A.M. (1987) Robust Regression and Outlier Detection. Wiley, New York.

Sachs, J.D. and A. Warner (1995) "Economic Reform and the Process of Global Integration", *Brookings Papers in Economic Activity*, Issue 1, 1-118.

Sachverständigenrat zur Begutachtung der Gesamtwirtschaftlichen Entwicklung (1999) Jahres Gutachten 1999/2000 - Wirtschaftspolitik unter Reformdruck, Wiesbaden.

Sheather, S. J. and M. C. Jones (1991) "A Reliable Data-based Bandwidth Selection Method for Kernel Density Estimation", *Journal of the Royal Statistical Society* B 53, 683-690.

Song, S., Chu, G.S.F. and R. Cao (2000) "Intercity Regional Disparity in China", *China Economic Review* 11, 246-261.

Silverman, B.W. (1981) "Using Kernel Density Estimates to Investigate Multimodality", *Journal of the Royal Statistical Society* 43, 97-99.

Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*, London (Chapman and Hall).

Temple, J. (1998) "Robustness Tests of the Augmented Solow Model", *Journal of Applied Econometrics* 13, 361-375.

Temple, J. (2000) "Growth Regressions and What the Textbooks Don´t Tell You", *Bulletin of Economic Research* 52, 181-205.

Unel, B. and H. Zebregs (2006) "The Dynamics of Provincial Growth in China: A Nonparametric Approach", *IMF Discussion Paper No.* WP/06/55, Washington.

Vamvakidis, A. (2002) "How Robust Is the Growth Openess Connection? Historical Evidence", *Journal of Economic Growth* 7, 177-194.

Wacziarg, R. and K.H. Welch (2003) "Trade Liberalization and Growth: New Evidence", *NBER Working Paper No.* 10152, Cambridge (Mass.).

Wand, M.P. and M.C. Jones (1995) *Kernel Smoothing*, London (Chapman and Hall).

Wei, Y., Liu, X., Parker, D. and K. Vaidya (1999) "The Regional Distribution of Foreign Direct Investment in China", *Regional Studies* 33, 857-867.

Wen, M. (2007) "Foreign Direct Investment, Regional Market Conditions and Regional Development – A Panel Study on China", *Economics of Transition* 15, 125-151.

Yao, S. and Z. Zhang (2001) "On Regional Inequality and Diverging Clubs: A Case Study of Contemporary China", *Journal of Contemporary Economics* 29, 466-484.

Xu, X. (2003) "Some Differences in GDP Measurement between China's practice and 1993 SNA Guidelines", *Research of Methodological Issues on National Accounts (No.6)*, Department of National Accounts, National Bureau of Statistics and National Accounts Division, OECD, Beijing.

Zhang, X., and J. He (2006) "Spatial Dependence Analysis of Chinese Regional Economy Growth", *Soft Science Journal* 20, 29-32.

Zhang, X., and J. He (2007) "Spatial Analysis for the Disparities if Chinese Regional Economy Development", *Areal Research and Development* 26, 35-39.