GROWTH DYNAMICS: AN EMPIRICAL INVESTIGATION OF OUTPUT GROWTH USING INTERNATIONAL DATA.

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Mónica Hernández Alava, BSc. MSc.

Department of Economics

University of Leicester

July 2002

To my mother María del Carmen,

my brother Jose María,

and

in memory of my father Jose María.

GROWTH DYNAMICS: AN EMPIRICAL INVESTIGATION OF OUTPUT GROWTH USING INTERNATIONAL DATA.

by

Mónica Hernández Alava

ABSTRACT

The rates of growth of output per head vary across countries. Despite the fact that these differences are of a small order of magnitude, they would translate into large differences in the average living standards of the countries if they were to persist over the years. It is therefore very important to understand the process of long run growth and as a consequence many recent studies concentrate on the issue of cross country convergence.

The aim of this thesis is to investigate the process of growth across countries and the possibility of inter-relationships of these processes across countries. To this avail, an empirical analysis of per capita output across countries is carried out first using the exact continuous time version of two neoclassical growth models, the Solow growth model and The Ramsey-Cass-Koopmans model. Results show that when these models are estimated consistently countries do not seem to be converging in the sense typically used in the literature. The rest of the thesis aims to investigate in more detail the processes by which growth in different countries might be related. Based on extensions of another neoclassical model, the Overlapping Generations model, and using a nonlinear switching regime model for estimation, two empirical analyses are carried out. The first one examines the role of balance of payments constraints in cross country growth determination. The second studies the extent of technology spillovers across countries and their effect on the process of growth. On one hand, results reveal little evidence of current account deficits constraining growth in the long run in the G7 countries although there is ample evidence of an influence in the short run dynamics of growth. On the other hand, spillovers of technology across the G7 countries are found to be of importance in the process of growth.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Professor Kevin Lee. I am grateful to him for what I have learnt.

I am indebted to many people in the economics department at the University of Leicester who have helped me over the years both before and during my thesis. In particular to Dr. Ali al-Nowaihi, Mr. Ian Bradley, Professor Woiciech Charemza, Mr. Derek Deadman and Professor Steve Pudney for their invaluable help.

I would also like to express my gratitude to Allan Wailoo for his support and Adriana Agapie, Rachel Carman, Daniela Hristova and Ewa Majerowska for their friendship.

I am grateful to all the people who have supported me over the years, both in Spain and in the UK. I could not possibly name them all in this space, but they know who they are.

Many thanks to all my family for their unconditional love, financial support and their encouragement beyond belief throughout all my studies.

TABLE OF CONTENTS

Length: approximately 72,000 words.

Chapter 1

Introduction.

Measures of output per head remain the primary means of comparing living standards in a country over time and of comparing standards across countries at any instant. Understanding the processes by which per capita income evolves over time is of enormous practical importance and is fundamental to an understanding of the determination of the wealth of nations and global inequality across countries since persistently different growth rates of output across countries would lead to large differences in standards of living across countries in the long run.

In the late 1960's, many economists lost interest in the theories of growth of which the traditional neoclassical model (Solow (1956), Ramsey(1928), Cass (1965), Koopmans (1965)) was one of the most influential. Over recent years, however, there has been a growing interest in international growth and convergence, following publication of, and stimulated by, the theoretical work of Romer (1986) and Lucas (1988). This "new" theoretical literature which has become known as "endogenous growth" literature, has concentrated on trying to explain growth with models for which long run growth is determined within the model itself. Complementing the revival in interest in theoretical growth models, there has also been considerable interest shown in growth empirics. The majority of the empirical literature has continued to employ neoclassical models, which assumes an exogenous technological change to explain the growth of output per capita in the long run. A prominent issue in the empirical literature has been that of international convergence of per capita income[1] although the results from the empirical studies have given conflicting evidence. Cross section studies have gener-

[1]The notion of convergence will be developed fully in Chapter 2. For the moment, it can be taken to mean that poor countries tend to catch up with rich ones.

ally found evidence of convergence taking place but the approach has been criticized on both econometric and theoretical grounds. These criticisms have prompted a number of other empirical studies following various alternative methods of estimation which find no evidence of convergence taking place.

This thesis is concerned with the study of the process of growth and convergence across countries. It revolves around three main empirical analyses found in chapters 3, 5 and 6. To provide a context within which to consider these three empirical exercises, the thesis also describes some of the relevant recent literature. This is carried out in chapters 2 and 4.

Chapter 2 reviews the background literature, both theoretical and empirical, on the issue of growth and convergence across countries. The chapter is split into three parts. The first part is concerned with the theoretical models of growth. Here, having briefly discussed the development of the theoretical models of growth, it presents a more detailed account of three theoretical neoclassical growth models which have dominated the discussion in this field; namely, the Solow growth model, the Ramsey-Cass-Koopmans model and the Overlapping Generations model. In the second part of the chapter, the most important empirical findings on growth and convergence of output per capita are reviewed focusing mainly on linear studies. The literature on nonlinear econometric papers is left to chapter 4. The linear empirical literature is divided between those who advocate the use of cross section estimation, those who advocate the use of time series, those who advocate the use of panel data and those who advocate other methods to try to obtain a picture of the evolution of the intra-distribution dynamics. In reviewing this literature, the advantages and disadvantages of each approach are drawn out. Additionally, since the majority of studies in this field use slightly different datasets, some of the more influential studies are reproduced using a common dataset[2] to assess the true differences among these studies. The majority of the empirical studies in this area have concentrated mainly on convergence of output per capita across countries. However, there has been a strand of the literature which has emphasized the importance of convergence of technology levels due to spillovers across countries. If technology across countries converges, then, as a consequence, per capita output will also converge. This literature on technology convergence is also reviewed here.

Finally, in the third part of the chapter, two simple theoretical extensions of the Overlapping Generations model are presented. These two simple extensions highlight the importance of analysing growth within a more realistic framework, including other macroeconomic variables that are relevant for the process of growth. They illustrate how the patterns of growth might be affected when interactions between countries are taken into account. Two explicit forms of interaction are considered. First, the basic model is extended to an open economy setup with imperfect capital markets following Obstfeld and Rogoff (1996). In this case, the interactions between countries arise directly through the balance of payments. Second, the basic model is extended to allow for technology spillovers across countries. This model captures the idea of "technology catch up" by characterizing cross country technology differences via a simple probability distribution. These two models motivate the empirical analysis carried out in chapters 5 and 6.

Chapter 3 deals with the issue of convergence of per capita output. This chapter takes the standard discrete time regression model based on Solow and found frequently

[2]The data comes from the Penn World Tables and it includes a group of 81 countries and two subsamples of this group (see Chapter 2 for a more detailed explanation of these data).

in the empirical literature and extends it in two directions. It considers a more elaborated linear system of equations based on the intertemporal optimising framework of Ramsey-Cass-Koopmans and compares it with the linear equation arising from the Solow growth model. In addition, exact discrete econometric models are derived from the theoretical continuous time Solow and Ramsey-Cass-Koopmans models. Using the well known Penn World Tables data for 81 countries, this chapter shows the systematic differences arising in some parameters of interest when these models are estimated consistently. Furthermore, these differences make convergence across countries in the sense used in the literature even more unlikely.

The remaining of the thesis aims to investigate the growth processes in different countries in more detail by allowing for inter-relationships across countries. For this purpose, the rest of the empirical work on this thesis concentrates on a smaller group of countries, namely, the G7 countries. Specifically, two types of interactions are considered, both based on the extended Overlapping Generations models presented in chapter 2. These models require the use of a more flexible econometric approach: the switching regressions model whereby output growth is generated by one of two different processes with imperfect sample separation. Changes between these two processes over time are then governed by the outcome of a third regression. Chapter 4 reviews the relevant literature on nonlinear switching regime models. It also presents the econometric issues relating to the estimation of the specific switching regime model used in this thesis; the switching regression model. These issues are not always addressed in the empirical literature. To tackle them, a preferred method of estimation is suggested. In addition post-estimation model evaluation tests and computation of impulse responses are discussed.

Having examined the econometric issues, chapter 5 investigates the effects of current account deficits on the dynamics of growth and on long run growth. Chapter 6 studies the importance of technology spillovers across countries. In the countries studied here, there seems to be little evidence of current account deficits constraining growth in the long run with the exception of Canada. However, there is evidence that past accumulations of current account deficits or surpluses influence the short run dynamics of the growth equation. However, for the same group of countries, there seems to be supporting evidence of spillover effects across countries influencing the growth process.

Chapter 2

The background to the convergence debate.

2.1 Introduction.

Growth rates differ substantially across countries. Figure 2.1 shows a histogram of the growth rates of real per capita GDP for a group of 102 countries over the period 1960 to 1989[3]. Although these differences seem small, they would imply large differences in the average living standards of the countries if they were to be maintained over a long period of time. For instance, in this data sample, the average growth for the lowest quintile is -0.16% and the average growth for the highest quintile is 2.78%. It is quite unlikely that these differences will continue forever but assuming that they do so, say, for the next 30 years, real per capita GDP will fall by almost 5% for the average country in the lowest quintile and it will increase by more than 127% for the average country in the highest quintile. As an example, Singapore, a country in the highest quintile, increased its real per capita GDP from $1,626 in 1960 to $11,062 in 1989, while at the other extreme, per capita GDP in Madagascar fell from $1,187 to $680. All this shows the importance of understanding long term growth since it has large repercussions on the standards of living. However, short and medium run fluctuations should not be disregarded on the grounds of being transitory and therefore,not important in comparison to long run growth since it is difficult to believe that large fluctuations do not affect long run growth.

One of the most influential theoretical growth models is the neoclassical growth model (Solow (1956), Ramsey(1928), Cass (1965), Koopmans (1965)). In the late 1960's, many economists appeared to lose interest in theories relating to long run eco-

[3]See Section 2.3 for a description of the data

nomic growth. The theoretical work of Romer (1986) and Lucas (1988) stimulated new research that came to be known as endogenous growth whereby growth is determined within the model in contrast with neoclassical models for which the growth rate is exogenous. However, although there is a consensus that theoretical models of growth should move in this direction, that is, trying to endogenize the long run growth rate, it has been shown (see Solow (2000)) that some of the most influential models in the area of endogenous growth (Romer (1986, 1990), Lucas (1988), Grossman and Helpman (1991a), Aghion and Howitt (1992)) rely on very strong assumptions which are not substantiated in the derivation of these models. Relaxing these assumptions leads to either no endogenous growth or even worse, output becoming infinite in finite time.

At the same time that these new theories were being developed, a fresh set of scaled data became available for a large number of countries which made comparisons across different nations and growth models feasible. One issue that has been particularly prominent in empirical growth studies is that of international convergence of per capita income. While there have been many studies in the area, results have conflicted. In general, cross section and some forms of panel data estimation have found evidence to support convergence (see, for example, the influential papers of Barro (1991,1997), Barro and Sala-i-Martin (1992, 1995), Mankiw, Romer and Weil (1992), Islam (1995) and Sala-i-Martin (1996b, c and 1997)). However, these approaches have been criticized for a variety of both econometric and theoretical reasons. Such criticisms have prompted a number of studies, using alternative estimation methods which have usually found no evidence of convergence in growth rates (see for example Lee, Pesaran and Smith (1995, 1997, 1998), Evans and Karras (1996a), Quah (1993a, b and 1996b), Maddala and Wu (2000)). Overviews of some parts of the recent empirical research have appeared in the literature, for example, de la Fuente (1997), Klenow and Rodriguez-Clare (1997b) and Durlauf and Quah (1999).

The main aim of this chapter is to review and reassess some of the most influential empirical results in the growth literature in the light of the recent econometric advances. Section 2.2 provides a brief overview of the developments of the theoretical growth literature and the background to the convergence debate. It also presents in some detail three neoclassical models of growth which have been used extensively in the growth literature and reviews some of the more important empirical findings on convergence. There has been a part of the literature which has maintained the importance of focusing on technological convergence using Total Factor Productivity as opposed to convergence of output per capita. The reason being that if technologies across countries are not converging, then per capita output will not converge either. This section also reviews this strand of the literature. Section 2.3 provides some empirical results to help assess this empirical evidence on per capita output growth and convergence, replicating some of the studies using three samples of countries. The purpose of this exercise is to judge whether the disparate results found using different estimation techniques can be attributed to the different datasets employed in these studies. In section 2.4 some extensions to the Overlapping Generations model are presented. These simple theoretical models motivate the empirical analysis of chapters 5 and 6. Finally section 2.5 concludes.

2.2 Growth and Convergence.

One of the most influential models in growth theory which is still very much in use today in empirical studies is Solow's (1956) growth model. Solow tried to improve on the Harrod-Domar model (Harrod (1939), Domar (1946)) by relaxing some of its assumptions. In the Harrod-Domar model, growth in the long run is only possible if and only if the savings rate of a country is equal to the product of the capital-output ratio and the labour growth rate.[4] It is only in this case that a country's supply of labour will be in balance with the stock of plant and equipment and there will be no scarce or unused resources anywhere in the economy. In this case, the growth of the country will be equal to the savings rate divided by the capital-output ratio. One of the problems of this model is that sustained growth was reached by chance and it could not be considered a highly likely outcome. In addition, one implication of the model which did not seem very realistic was that the savings rate affects the rate of growth in the long run. However, in this model, the three main components, the savings rate, the capital-output ratio and the labour growth rate were assumed to be three given constants. The neoclassical Solow growth model relaxes the assumption of fixed proportions of the Harrod-Domar model and allows for a positive elasticity of substitution between the inputs in the production function. This modification of the model has two main consequences. First, a positive growth rate in the long run is no longer a rare event, there is a stable path along which a country would eventually reach its steady state. Second, although the savings rate affects the level of output, it has no influence on the long run growth rate whatsoever. In this model, the rate of growth of output in the long run is given by the sum of

[4]Labour in this model is measured in efficiency units to allow for technical progress, that is, labour is multiplied by the level of technology.

the growth rate of population and the growth rate of technological progress. Therefore, it is straightforward to work out that output per capita grows at the same rate as technological progress in this model. Using Ramsey's (1928) optimization analysis, Cass (1965) and Koopmans (1965) refined the basic neoclassical model introducing endogeneity of the saving rate. In this model, infinitely lived households make intertemporal decisions about consumption and savings. Although the transitional dynamics are different to those of the Solow model, once the economy has converged to the steady state its behaviour is the same. A second model which also introduces endogeneity of the saving rate is the Overlapping Generations model developed by Diamond (1965). It is also an optimizing model, but it differs from the Ramsey-Cass-Koopmans model in that there is turnover of the population. Again, the behaviour of the economy once it reaches the steady state is the same as in the Solow model even tough the transitional dynamics are also different.

In the late 1960's, the interest in growth theory died down and it was not until the eighties that the revival of the growth theory started. This renewed interest in the growth theory has continued since then and it was stimulated by the theoretical work of Romer (1986) and Lucas (1988). This new strand of the theory was later knows as "endogenous growth theory". In the neoclassical growth model the growth rate of output per capita in the long run is equal to the rate of technological progress which is taken as an exogenous constant. Thus, growth is determined by something that is exogenous to the model and for this reason, neoclassical growth models are also sometimes known as "exogenous growth models". This exogeneity of the growth rate is what endogenous growth theorists criticize in the neoclassical model. Consequently, their aim is to develop models in which the growth rate is endogenous, hence the name of "endogenous growth theory". However, in words of Solow (2000) "The way to think about exogenous growth theory is that it means to show how the path of aggregate output adjusts to the rate of population growth and the rate of technological progress, whatever they happen to be and for however long they persist". There has been a wealth of models following the idea of an endogenous growth rate. Most of these models try to incorporate in the model a theory of technological progress. Some of the most influential of these models are those of Romer (1986, 1990), Lucas (1988), Grossman and Helpman (1991), Aghion and Howitt (1992) amongst others. Solow (2000) critically reviews the most discussed endogenous growth models. He shows that the assumptions embedded in these models which might appear innocuous at first sight are usually very strong. Thus, the endogenous growth literature trades off the exogeneity of the growth rate for the presence of very precise key assumptions which are not usually justified at all. The role of these assumptions is to ensure that technology (or its equivalent in the model) grows exponentially in the steady state, so that then, growth can be taken as being explained by the model. For this reason, in this thesis attention is restricted to neoclassical growth models.

Section 2.2.1 explains what it is meant by convergence across countries and presents the different definitions that have been used in the literature. In the light of Solow's (2000) criticisms of the strong assumptions embedded in endogenous growth models, section 2.2.2 presents the three neoclassical models mentioned above, namely, the Solow growth model, the Ramsey-Cass-Koopmans growth model and the Overlapping Generations growth model in more detail. These are the three theoretical models which will be used as the basis for the different empirical chapters in this thesis. This section also outlines their similarities and differences. In section 2.2.3 the empirical findings on growth and convergence are described together with a critical assessment of the different approaches taken to investigate this issue. In all the literature reviewed in section 2.2.3, technology is assumed to be exogenous to the growth process, and for most of the empirical analyses it is also assumed to grow at the same rate in all economies. Section 2.2.4 concentrates on a part of the literature on convergence which focuses on technology transfers or spillovers as a major force behind convergence.

2.2.1 The convergence hypothesis.

One of the issues that has attracted a lot of attention recently is the convergence hypothesis. In "traditional" neoclassical models the growth of income per capita in the long run is given by the exogenous technological growth rate. There exists a well defined steady state level of income and differences in income per capita between countries persist only if countries have different savings, population growth rates or different rates of technological change. This convergence property derives from the diminishing returns to capital in the production function. If an economy has low capital per worker relative to its long run level, it will have a high rate of return on capital and consequently a high growth rate. However, while it might be reasonable to assume that two equally developed economies have similar savings rate, population and technological growth rates, it would seem excessive to assume the same for two economies in very different development stages.

There exist several alternative definitions of convergence and some discussions on this topic are provided in Bernard and Durlauf (1996), Bernard and Jones (1996b), Galor (1996), Quah (1996a, b) and Sala-i-Martin (1996a), amongst others. The most commonly used definition of convergence is *beta convergence* which examines how quickly the logarithm of output per capita moves towards its steady state value from some initial conditions. Beta convergence can be either *absolute* (poorer economies "catch up" with richer economies by growing at a faster rate at the beginning and subsequently slow down to a common rate of technological progress) or *conditional* (economies grow faster the further they are from their steady state positions).[5] However, beta convergence represents only average behaviour (see Quah (1993a)), so that it does not provide any information on the behaviour of the dispersion of the logarithm of output per capita in the long run. This relates to the second concept of growth convergence; namely, if the cross section dispersion diminishes over time, economies are said to exhibit *sigma convergence.* As noted by Friedman (1992), Quah (1993b) and Sala-i-Martin (1996a) inter alia, beta convergence is a necessary condition for sigma convergence but it is not sufficient. Therefore, sigma convergence captures aspects of convergence not adequately captured by beta convergence. However, estimating sigma on its own does not capture the whole story since it does not convey any information on intra distribution dynamics. Even with constant dispersion, there might be persistence in the inequality of the logarithm of output or, alternatively, there might be significant intra distribution dynamics. In this case, none of these aspects are picked up by this convergence definition.

A third interpretation of the convergence idea found in the literature is more concerned with the absence of divergence. In this interpretation, countries have their own steady state, but they are said to converge if they have either the same stochastic or

[5] In the latter case, one economy with low per capita output might grow more slowly than another economy with higher per capita output simply because the first economy is close to its steady state level of per capita output, while the other economy is far away from its (higher) steady state level.

deterministic long run trends (see for example Bernard and Durlauf (1995, 1996) and Evans and Karras (1996a,b)). Consequently, these tests look at the long run behaviour of the differences in per capita output across countries. The definition of convergence used by Evans and Karras (1996a, b) states that the deviations of individual series from their cross country average, conditional on all available information, can be expected to approach a constant. If this constant is zero, the convergence is absolute. Alternatively, if this constant is different from zero, the convergence is conditional. This implies that for convergence to take place, each series has to be nonstationary but deviations from the cross country average need to be stationary. Bernard and Durlauf (1995) present a similar but alternative definition based on Johansen's cointegration procedure to test whether there is a unique common trend across countries.

To test convergence, empirical studies have used cross section, time series and panel data estimation methods. Time series and cross section estimation have usually given conflicting empirical results. On the one hand, cross section tests have usually supported the convergence hypothesis, while time series have generally supported the hypothesis of no convergence. Bernard and Durlauf (1996) showed that cross section tests and time series tests place different assumptions on the data and that time series tests are associated with a stronger definition of convergence than cross section tests. In cross section tests, it is assumed that economies are in the transition towards the steady state and convergence is interpreted as *catching up* (i.e. differences in output between countries decrease over time). In time series tests, however, it is assumed that economies are near their limiting distributions and convergence implies that the initial conditions have no significant effect on the expected value of output differences. Bernard and Durlauf (1996) also showed that if the data are generated by processes with multiple long run equilibria, cross section tests can reject the hypothesis of no convergence. Moreover, if the data are highly influenced by transitional dynamics, time series may accept the no convergence null.

2.2.2 Three Neoclassical growth models.

This subsection presents some of the growth models that have been used in the neoclassical growth literature. Three closed economy models will be reviewed: the Solow growth model, the Ramsey-Cass-Koopmans (RCK) model and the Overlapping Generations (OG) model. Most of the recent empirical work in the area of growth and convergence has concentrated on the Solow growth model. This has focused mostly on its deterministic version (which will be presented here), although there has been some part of the literature dealing with the stochastic version of this model (see, for example, Binder and Pesaran (1999)).[6] This model however assumes a constant savings rate which is exogenously given. The second theoretical model presented here, the RCK model, introduces endogeneity of the savings rate by using optimisation analysis.[7] The OG model treats the savings rate as endogenous like the RCK model. The difference between the two models is that the RCK model assumes households which are infinitely lived, whereas in the OG model there is turnover in the population; that is, new individuals are perpetually being born while old individuals are perpetually dying.

[6]The deterministic version of the Solow growth model will be used in Section 2.3 to reassess the empirical evidence on growth and convergence using a common dataset.

[7]This model will be used in Chapter 3 to examine whether the more sophisticated transitional dynamics, relative to those of the Solow model, provide additional insights.

2.2.2.1 The standard Solow growth model.

The Solow growth model is the basic reference point for most growth analyses. This model has been derived in both continuous time (see, for example, Romer (1996)) and discrete time (see, for example, Obstfeld and Rogoff (1996)). In this section, the derivation of the Solow growth model will be presented in discrete time. However, in chapter 3, the implications of using this version of the model rather than the continuous time version, will be outlined explicitly.

The discrete time derivation of the Solow model presented here is along the lines of that in Mankiw *et al* (1992, Lee, Pesaran and Smith (1995) and Obstfeld and Rogoff (1996). In what follows, output Y_t is produced by physical capital K_t and labour L_t. The variable A_t represents technology and endowment,[8] so effective labour input is measured by A_tL_t. Assuming a Cobb Douglas production function, we have

$$Y_t = K_t^{\alpha}\left[A_tL_t\right]^{1-\alpha} \qquad 0<\alpha<1 \tag{2.1}$$

Technology and population are assumed to grow at fixed rates, g and n, respectively and capital stock depreciates at δ. There is a closed economy, so investment, I, equals savings, where it is assumed that a fixed proportion of income, s, is saved.[9] Lower case letters will be used to denote effective measures (units per effective labour), while "$\sim$" over lower case letters will denote per capita variables, e.g. $\tilde{y}_t = Y_t/L_t$. The

[8] The assumption of labour augmenting technological progress in the Solow model is not needed to develop the theory but it is introduced because of the attraction of exponential steady states. Labour augmenting technology is also referred to as "Harrod neutral". This kind of technology has the property that the relative input shares $(KF_K)\,/\,(LF_L)$ remain unchanged.

[9] If there are several forms of capital, each one would have its corresponding investment rate.

production function can be written in terms of effective units as

$$\frac{Y_t}{A_t L_t} = y_t = f(k_t) = k_t^{\alpha} \tag{2.2}$$

The evolution of the effective capital stock over time is given by the following nonlinear difference equation:

$$(1+n)(1+g)k_t = sk_{t-1}^{\alpha} + (1-\delta)k_{t-1} \tag{2.3}$$

where δ is the rate of depreciation. In steady state, the effective output is constant at

$$y^* = \left(\frac{s}{(1+n)(1+g)-(1-\delta)}\right)^{\frac{\alpha}{1-\alpha}}$$

Thus, a higher saving rate raises the steady state level of output per effective worker y^*. A higher rate of depreciation or a higher population or technological rate of growth have a detrimental effect on y^*, either because more of the economy's investment has to be used to replace the capital that is depreciating or because capital is provided for new effective workers. This steady state is stable in the sense that no matter where the economy starts (apart from zero), it will eventually end up in the steady state.

Output per worker ($\tilde{y}_t$) is given by $\tilde{y}_t = A_t y_t$. Taking natural logarithms, differentiating and using a first order Taylor's approximation around the steady state of equation (2.3), the following equation is obtained

$$\ln \tilde{y}_t = \mu + (1-\lambda) gt + \lambda \ln \tilde{y}_{t-1} \tag{2.4}$$

where

$$\lambda = 1-(1-\alpha)(n+g+\delta) \tag{2.5}$$

$$\mu = \lambda g+(1-\lambda)\left[\ln(A_0)-\frac{\alpha}{1-\alpha}\ln(n+g+\delta)+\frac{\alpha}{1-\alpha}\ln(s)\right] \tag{2.6}$$

and $(1-\lambda)=(1-\alpha)(n+g+\delta)$ is the speed of convergence to equilibrium. In cross country analyses of the neoclassical growth model, it is usually assumed that there is common technology across countries but different tastes, leaving, α, g and δ the same across countries but allowing for different values of s_i, n_i and A_{i0} (the savings rate, the rate of population growth and initial endowments). For estimation purposes, a serially uncorrelated disturbance, ε_{it}, with mean zero and variance σ_i^2, independently distributed of s_i and n_i and independent across countries is typically added to the equation. These disturbances are often interpreted as technological or productivity shocks, although this is an *ad hoc* interpretation not necessarily compatible with the neoclassical model in its stochastic variant (see Binder and Pesaran (1999)). With these extensions, equation (2.4) becomes

$$\ln(\tilde{y}_{it})=\mu_i+(1-\lambda_i)\,gt+\lambda_i\ln(\tilde{y}_{i,t-1})+\varepsilon_{it} \qquad \begin{array}{l} i=1,...,N \\ t=0,1,2,...,T \end{array} \tag{2.7}$$

where

$$\begin{aligned}\lambda_i &= 1-(1-\alpha)(n_i+g+\delta)=1+\beta_i,\\ \mu_i &= \lambda_i g+(1-\lambda_i)\left[\ln(A_{i0})-\frac{\alpha}{1-\alpha}\ln(n_i+g+\delta)+\frac{\alpha}{1-\alpha}\ln(s_i)\right]\end{aligned}$$

If this is the case, the β_i's should be negative and differ across countries depending on the rate of growth of population, n_i, and g should be positive and constant across countries. Lee, Pesaran and Smith (1995) estimated this model but allowing for different growth rates of technology, g, for each country in their three samples. That is, they used the following equation

$$\ln(\tilde{y}_{it}) = \mu_i + (1 - \lambda_i) g_i t + \lambda_i \ln(\tilde{y}_{i,t-1}) + \varepsilon_{it} \tag{2.8}$$

with $i = 1, ..., N$ and $t = 0, 1, 2, ..., T$. They found that both, the estimates of the speed of convergence and the technology growth rates differ across countries. This in turn implies that if countries in their three samples are converging, they are converging to their own steady state, not to a common one as it has been sometimes suggested in the empirical literature.

2.2.2.2 The standard Ramsey-Cass-Koopmans model.

Cass (1965) and Koopmans (1965) introduced endogeneity of the saving rate in the model by bringing in Ramsey's (1928) optimization analysis. The Ramsey-Cass-Koopmans model is a continuous time model, although it can also be derived using discrete time (see for example Obstfeld and Rogoff (1996)). Saving and consumption decisions are made by infinitely-lived households who are maximizing lifetime utility in a closed economy. Firms produce only one type of output according to a constant returns to scale production function. Technology, as in the Solow model, is assumed to grow exogenously at a rate g. Factor and output markets are competitive so each factor of production is paid its marginal product. There are a large number (H) of identical households which grow at a rate n.[10] In every period, each individual supplies one unit of labour and rents the capital s/he owns to the firm. Households divide income between consumption and savings to maximize lifetime utility subject to a budget constraint. Assuming a constant risk aversion utility function with parameter θ, the households lifetime utility can be written as

$$U = A(0)^{(1-\theta)} \frac{L(0)}{H} \int_0^{\infty} e^{-(\rho-n-(1-\theta)g)t} \frac{c(t)^{(1-\theta)}}{1-\theta} dt \tag{2.9}$$

where ρ is the discount rate and $c(t)$ is consumption per unit of effective labour. To avoid problems with infinite lifetime utilities, it is assumed that $\rho - n - (1-\theta)g > 0$. The household budget constraint states that the present value of lifetime consumption cannot be higher than the initial wealth plus the present value of lifetime income which in terms of effective units can be written as

$$\int_0^{\infty} e^{-R(t)} c(t) e^{(n+g)t} dt \leq k(0) + \int_0^{\infty} e^{-R(t)} w(t) e^{(n+g)t} dt \tag{2.10}$$

In a closed economy, any change in capital in the firms must come from households savings which, in terms of effective units, can be written as

$$\dot{k}(t) = f(k(t)) - c(t) - (n+g+\delta) k(t) \tag{2.11}$$

Note that this equation corresponds to equation (2.3) of the Solow model. The household maximises lifetime utility (2.9) subject to a budget constraint (2.10). Using Lagrange, taking logarithms, differentiating with respect to t and substituting the marginal

[10]The number of households stay the same because births are considered additions to the family.

product of capital we get the following differential equation

$$\frac{\dot{c}(t)}{c(t)} = \frac{f'(k(t)) - (\rho + \theta g + \delta)}{\theta} \tag{2.12}$$

In what follows, it is understood that k, c, y and s are functions of time. Assuming a Cobb-Douglas production function ($y = k^{\alpha}$) and using the fact that the saving rate is given by $s = (y - c)/y$, we can transform equations (2.11) and (2.12) into the following system of differential equations in y and s

$$\dot{y} = \alpha y \left[y^{\frac{\alpha-1}{\alpha}} s - \Gamma \right] \tag{2.13}$$

$$\dot{s} = (1 - s) \left[\alpha y^{\frac{\alpha-1}{\alpha}} \left(s - \frac{1}{\theta} \right) - \alpha\Gamma + \frac{\Phi}{\theta} \right] \tag{2.14}$$

where $\Gamma = n + g + \delta$ and $\Phi = \rho + \theta g + \delta$.The steady state values of y and s are:

$$y^* = \left(\frac{\alpha}{\Phi} \right)^{\frac{\alpha}{1-\alpha}} \tag{2.15}$$

and

$$s^* = \frac{\alpha\Gamma}{\Phi} \tag{2.16}$$

In this model, y is predetermined and s is a jump variable and the economy exhibits saddle path stability.[11] Once the economy has reached the steady state, its behaviour is identical to that of the Solow model on the balanced growth path.[12] Capital, output and

[11]Rodriguez and Sachs (1999) showed that in a Ramsey growth model augmented with natural resources, overshooting of the steady state equilibrium can be optimal. The rationale behind this overshooting is that resource abundant countries are likely to be living beyond their means. If this is the case, countries will converge towards the steady state from above and they will, therefore, exhibit negative growth rates during the transition period.

[12]Note that this model implies a slightly different speed of convergence to the steady state than the Solow model because the savings rate is allowed to change over time. The speed of convergence in

consumption per unit of effective labour are constant and the saving rate is also constant. All per capita variables grow at the rate of growth of technology, g, and variables in levels grow at the sum of technology and population rates of growth, $n + g$.

2.2.2.3 The standard Overlapping Generations model.

The RCK model used in the previous section is that of a representative agent who lives forever and maximizes lifetime utility. The main difference between the RCK model and the Overlapping Generations (OG) model, first introduced by Samuelson (1958) and Diamond (1965), is that there is turnover in the population; that is, new individuals are continuously being born while old individuals are continuously dying. A relatively simple version of the OG model, based on Romer (1996), is provided below. In contrast to the RCK model, time here is assumed to be discrete simply because the derivations turn out to be more straightforward. Individuals live two periods. They work and save in the first period and in the second, they live off the benefits of their first period savings. The Cobb-Douglas production function,

$$Y_t = K_t^{\alpha} (A_t L_t)^{(1-\alpha)} \qquad 0 < \alpha < 1 \tag{2.17}$$

is assumed, where Y_t is output at time t, K_t is the capital stock, A_t is the level of technology and L_t is the labour force which is the same as the cohort born at time t.

Ramsey's model is given by the negative eigenvalue of the system of differential equations (2.13) and (2.14)

$$\Psi = \frac{\Phi - \Gamma - \sqrt{(\Phi - \Gamma)^2 - \frac{4(1-\alpha)(\alpha\Gamma - \Phi)\Phi}{\alpha\theta}}}{2}$$

and not only depends on n, g, δ and α as in the Solow model, it also depends on the household willingness to shift consumption between periods (θ) and the rate at which households discount future utility (ρ).

This can be written in terms of effective units as

$$y_t = \frac{Y_t}{A_t L_t} = \left(\frac{K_t}{A_t L_t}\right)^{\alpha} = k_t^{\alpha} \tag{2.18}$$

The number of the young, and, therefore, the population grows at rate n and technology grows at g:

$$L_t = (1+n) L_{t-1} \tag{2.19}$$

$$A_t = (1+g) A_{t-1} \tag{2.20}$$

The domestic markets are perfectly competitive and capital depreciates at a constant rate δ. Thus, in equilibrium, the net marginal return on each factor of production equals its price

$$(1-\alpha) k_t^{\alpha} = w_t \tag{2.21}$$

where w_t is the wage rate per unit of effective labour and

$$\alpha k_t^{\alpha-1} = r_t + \delta \tag{2.22}$$

where r_t is the real interest rate.

For expositional ease, it is assumed that a young individual's utility function is given by

$$U_t = \ln\left(\tilde{c}_t^{Y}\right) + \beta \ln\left(\tilde{c}_{t+1}^{O}\right) \tag{2.23}$$

where $\tilde{c}_t^Y$ stands for per capita consumption during the youth of someone born at time t, $\tilde{c}_{t+1}^O$ denotes the consumption of the same person while old in period $t+1$ and β is the discount rate. Note that the logarithmic utility function of equation (2.23) is a special case of a constant relative risk aversion utility function[13]; that is, a utility function of the form $U = C^{1-\theta}/(1-\theta)$, with the parameter θ set to 1. The budget constraints for the young and the old are given respectively by

$$K_{t+1} = A_t L_t w_t - L_t \tilde{c}_t^Y \tag{2.24}$$

and

$$L_t \tilde{c}_{t+1}^O = (1 + r_{t+1}) K_{t+1} \tag{2.25}$$

These can be combined, by eliminating K_{t+1}, to obtain the following

$$\tilde{c}_t^Y + \frac{\tilde{c}_{t+1}^O}{1 + r_{t+1}} = A_t w_t \tag{2.26}$$

That is, the combined budget constraint states that the present value of a young person's lifetime consumption equals the present value of his lifetime income. Hence a young individual maximizes the utility function (2.23) subject to the budget constraint (2.26) and thus

$$\mathcal{L} = \ln\left(\tilde{c}_t^Y\right) + \beta \ln\left(\tilde{c}_{t+1}^O\right) + \lambda \left[A_t w_t - \tilde{c}_t^Y - \frac{\tilde{c}_{t+1}^O}{1 + r_{t+1}}\right]$$

[13]It is important to mention here that this particular form of the utility function (i.e. constant relative risk aversion) is needed in the overlapping generations model to obtain balanced growth.

The first order necessary conditions are then

$$\frac{\partial \mathcal{L}}{\partial \tilde{c}_t^Y} = \frac{1}{\tilde{c}_t^Y} - \lambda = 0 \tag{2.27}$$

and

$$\frac{\partial \mathcal{L}}{\partial \tilde{c}_{t+1}^O} = \frac{\beta}{\tilde{c}_{t+1}^O} - \frac{\lambda}{1 + r_{t+1}} = 0 \tag{2.28}$$

together with the budget constraint (2.26). From this point onwards, it is also assumed that depreciation, δ, is equal to 1 which is not such a strong assumption considering that each period in this model lasts half a lifetime. Setting $\delta = 1$ does not change the general conclusions, but it simplifies the analytical derivation of the model. Eliminating λ from equations (2.27) and (2.28), and then substituting into the budget constraint (2.26), consumption per effective units of the young in equilibrium is given by

$$c_t^Y = \frac{(1-\alpha)}{(1+\beta)} k_t^\alpha \tag{2.29}$$

It is worth mentioning at this point that this last equation (2.29) can be written in terms of output per effective worker using equation (2.18) as

$$c_t^Y = \frac{(1-\alpha)}{(1+\beta)} y_t$$

Thus, the consumption of the young is a constant fraction of output and, consequently, so are the savings of the young in the sense that they are independent of the interest rate.[14] This, however, is not a feature inherent to the model, but rather of the utility

[14]The consumption of the old is also a constant fraction of output in this special case in which depreciation is assumed to be equal to one. Therefore total consumption is also a constant fraction of output as in the Solow growth model of chapter 2.

function chosen (see equation (2.23)). As it was stated before, the logarithmic utility function is a special case of a constant relative risk aversion utility function in which the parameter θ is set to 1. Any other value of θ in the utility function will make consumption a function of the interest rate.

The dynamics of the economy can be characterized simply by aggregating the behaviour of individuals. Dividing equation (2.24) by $A_{t+1}L_{t+1}$, substituting equations (2.21) and (2.29) and using equations (2.19) and (2.20), the following equation of motion for k is found:

$$k_{t+1} = \frac{\beta\,(1-\alpha)}{(1+\beta)\,(1+n)\,(1+g)} k_t^{\alpha} \tag{2.30}$$

Thus, there is a unique equilibrium or steady state level of capital per effective worker, k^*, which is constant and given by

$$k^* = \left(\frac{\beta\,(1-\alpha)}{(1+\beta)\,(1+n)\,(1+g)}\right)^{\frac{1}{1-\alpha}} \tag{2.31}$$

Note that this can also be expressed from equation (2.18) as

$$k^* = \left(\frac{\alpha}{1+r}\right)^{\frac{1}{1-\alpha}} \tag{2.32}$$

where r stands for the equilibrium interest rate. The relevance of this point will be apparent in the section 2.4 where this closed economy model will be compared to an open economy model. The steady state value of capital per effective units, k^*, depicted in either equation (2.31) or equation (2.32) is globally stable; that is, no matter where the economy starts (apart from zero) it will always converge to the steady state. Using equation (2.18) together with (2.31), the effective output in steady state is given by

$$y^* = \left(\frac{\beta (1 - \alpha)}{(1 + \beta)(1 + n)(1 + g)} \right)^{\frac{\alpha}{1-\alpha}}$$

Thus, we get the familiar result that variables measured in effective units are constant in the steady state, implying that the variables in levels are growing at $n + g$; that is, the population plus the technology growth rates (see equations (2.19) and (2.20)). Therefore, once the economy has converged to the steady state (or balanced growth path), this model shares the properties of the Solow or RCK models with respect to their steady state.

To assess the speed of convergence of this economy to the steady state, equation (2.30) needs to be linearized around the steady state k^*. Using a first order Taylor approximation around k^*, we get

$$k_{t+1} - k^* \simeq \alpha (k_t - k^*)$$

Since α is between 0 and 1, the economy converges smoothly towards the steady state and α is the speed of convergence. Although at first sight this speed of convergence seems to be very high compared to the results obtained with the Solow or RCK models, it is important to see that they are not directly comparable because, in the OG model, each period lasts half a lifetime.

2.2.3 Empirical findings on growth and convergence.

The issue of convergence has been studied profusely in empirical papers. Cross section and some forms of panel data estimation have generally found evidence on convergence. This approach has been criticized in a number of grounds and has given rise to a bunch of studies which find divergence using alternative estimation methods. In what follows, the main findings and criticisms of the different approaches are described briefly.

2.2.3.1 Cross section studies.

Cross section analysis involves estimating cross country regressions of the logarithm of output growth per capita on the logarithm of per capita output at some initial point. That is, a regression of the form[15]

$$\ln(\tilde{y}_{i,T}) - \ln(\tilde{y}_{i,0}) = \alpha + \beta \ln(\tilde{y}_{i,0}) + \varepsilon_i$$

A negative value of β is taken as evidence of *absolute β convergence*. Baumol (1986) finds evidence of absolute convergence in a group of sixteen industrialized countries during the 18th century using Maddison's data.[16] However, De Long (1988) criticizes this evidence on the grounds of both a sample selection bias and the absence of any corrections for measurement errors (this last problem had already been mentioned by Baumol, but no attempt was made to take into account measurement errors). Baumol and Wolff (1988) acknowledge the sample selection bias but maintain that the conclusions of Baumol's (1986) paper are still supported by the evidence, although the results are not as strong.

[15]The approach of using per capita output has been critized by, for example, Sachs and Warner (1997). Their argument is that output per worker is the variable explained in production theory as opposed to output per capita. During transitions, as changes in the age structure of the population take place, there might be a gap between the growth of the population and the growth of the workforce.

[16]Maddison (1982, 1987 and1991) constructed a very long dataset (starting from the 1820's for some countries) which covers a small group of industrialized countries.

Absolute convergence is ruled out by all empirical studies nowadays. However it is clear that neoclassical growth models do not predict absolute convergence unless countries share common saving and population growth rates and have access to the same technology, what they do predict is conditional convergence. There are two ways in which the hypothesis of conditional convergence can be tested: one can choose a homogeneous group of countries with potentially the same steady state, or one could hold the steady state constant by including other explanatory variables in the regression to take into account heterogeneity across countries (different levels of technology, saving rates, tastes, tax rates, etc.). Many cross section studies have found that countries converge conditionally at a remarkably similar speed, roughly of 2% per year.[17] This rate of convergence is very slow and it would take approximately 35 years to make up for half of the gap between two countries. Even more astonishing is the finding of similar speeds of convergence in regional data sets such as the U.S. states, Japanese prefectures and regions in 8 Western European countries (see Barro and Sala-i-Martin (1995) and Sala-i-Martin (1996b)).[18] This apparent regularity takes Barro (1997) to state: "*Basically, 2 percent per capita growth seems to be about as good as it gets in the long run for a country that is already rich* ".

2.2.3.2 Criticisms of the cross section studies.

The cross section approach has been criticized in the context of both the econometrics and the economic theory. It has been argued that in most empirical studies, the

---

[17]See Barro (1991), Barro and Sala-i-Martin (1992, 1995), Makiw, Romer and Weil (1992), Sala-i-Martin (1996a, b and c, 1997) for some recent cross section empirical studies and Barro (1997) for related panel estimation.

[18]These 8 countries are Germany, United Kingdom, Italy, France, Netherlands, Belgium, Denmark and Spain.

choice of control variables is ad hoc[19]. Sachs and Warner (1997) argue that the variables that should be included in this regressions are measures of geography, measures of resource endowment and measures of economic policy. Additionally, these regressions have also been criticized in the grounds of "robustness". Levine and Renelt (1992) applied an extreme-bounds test to identify the "robust" relations in economic growth. They found few or no robust variables at all. On the other hand, Sala-i-Martin (1996c, 1997) argued that the extreme bounds test is too strong. He called "robust" those variables which are found to be significant in 95% of regressions run using a wide range of combinations of control variables. Out of a set of 62 variables, Sala-i-Martin uses three of them in all regressions: the level of income in 1960, the life expectancy in 1960 and the primary school enrolment rate in 1960. Then, for each variable tested, he combines the remaining 58 variables in sets of three. The same exercise is then repeated with the investment rate included as a fixed variable. Using this definition, Sala-i-Martin finds that there is a large number of variables that are "robust".[20]

2.2.3.3 The use of Panel data econometric methods.

It has also been argued that if the conditioning variables used to hold the steady state constant are imperfect, the estimated speed of convergence will be biased downwards. In particular, if the omitted determinants of the steady state are positively correlated with output after holding fixed the rest of the conditioning variables then the estimated speed of convergence will be understated. These arguments have led to the

[19]The number of explanatory variables that at some point have been found to be significant in this kind of regressions now exceeds fifty.

[20]This exercise, however, has been criticized by Durlauf and Quah (1999). They argue that it is unclear whether it uncovers something about the robustness of this variable or whether it just shows the covariance structure of the control variables. They also argue for the introduction of control variables which are economically interesting.

extensive use of panel data econometric methods to study the issue of convergence across countries. Using Lest Squares Dummy Variable (estimation with "fixed effects") and Minimum Distance (estimation with "correlated effects") over five five-yearly time periods in three different samples of countries,[21] Islam (1995) obtains somewhat higher convergence speeds, between 4.5 and 10 percent. A problem common to both the cross section and the panel data estimation used by Islam is that they restrict the technological growth rate to be the same across countries. Lee, Pesaran and Smith (1995, 1997, 1998) showed that the estimates obtained will not only be inconsistent but also biased if this assumption is violated. Lee *et al* estimated the coefficients of individual time series regressions and examined their distribution over three different groups of countries. They found that the estimated speeds of convergence and technological growth rates varied across countries so that the restriction of common speed of convergence and technology growth is rejected by the data.[22] Allowing for this heterogeneity and assuming that output is trend stationary, the mean speed of convergence was found to be approximately 20%. Imposing homogeneity of the convergence speed and the technological growth rate causes the estimated speed of convergence to drop to roughly 4% giving a clear indication of the size of the bias.

Miller (1996) also uses panel data methods and he finds support for the convergence hypothesis in a group of 22 OECD countries over the period 1960-1988. However, his approach has been criticized on the grounds that his estimation procedure does not take into account the dynamic structure of the panel. Lee, Longmire, Matyas and Harris (1998) use a panel of, again, 22 OECD economies during the period 1950-1990.

[21] This estimation allows for individual production functions which differ in unobservable individual "country effects".

[22] The speed of convergence estimated in this case is the speed towards each country *own* steady state, not towards a common one as is assumed in cross section studies.

They use more than 30 different methods to estimate the Solow model and conclude that these countries converge towards the steady state at a rate of 2 to 4% per annum. Maddala and Wu (2000) discuss the problem of biased estimates when using pooled regressions in the presence of heterogeneity. This paper is along the lines of Lee, Pesaran and Smith, but uses a Bayesian approach using shrinkage estimation. Four groups of countries were used in this study: a group of 98 countries; two subgroups of this, similar to those used by Mankiw, Romer and Weil (1992), for the period 1960-1989; and a fourth group covering 17 European countries during the period 1950-1990. Under heterogeneity, they find a faster convergence rate than under homogeneity, in line with the studies of Lee, Pesaran and Smith. Additionally, Maddala and Wu also show that the convergence rates vary with time. Bernard (2001) argues that for a group of 22 OECD countries, the estimates of a common long run growth rate are zero or significantly negative. He also finds that the hypothesis of homogeneity is rejected in favour of heterogeneity in the long run growth rates across this group of countries. A good discussion about the different methods used to estimate panels can be found in Nerlove (1999, 2000). He concludes that the method of estimation matters in that estimates of the speed of convergence are significantly different when using alternative estimation methods.[23]

2.2.3.4 Alternative statistical approaches.

Alternative approaches have also been suggested to study convergence. For example, Quah in a series of papers (1993a,b and 1996b) argues that cross section looks only at average behaviour of the economies and that it is more sensible to look at how

[23]With all the these estimation problems in mind, Temple (1998) argues the importance of using robustness tests when studying convergence across countries.

the distribution of income of the economies evolves over time. He suggests ways of studying this using stochastic kernels and points towards club convergence as a competing theory to conditional convergence. In this case, economies have multiple locally stable steady state equilibria and, therefore, transitory shocks may affect the long run equilibrium the economy tends to. He concludes that there is cross country divergence, with countries tending towards the very rich or the very poor and the gap between them expanding. Durlauf and Johnson (1995) find evidence of multiple regimes using regression tree analysis which involves estimating different cross section regressions and allowing the data to classify the countries into different regimes (these different regimes are shown to have different parameters in their production functions). Bernard and Durlauf (1995) present evidence that there is no cross country convergence although they do find evidence of common trends among countries suggesting that cross country growth cannot be explained only by country specific factors. Evans and Karras (1996a, b) find evidence of conditional convergence using time series tests. However, they also find, as many other studies, that the parameters of technology differ across economies. Feve and Le Pen (2000) also try to model convergence clubs for a group of 92 countries during the period 1960-1989. To classify countries into different clubs, they use an endogenous switching regression model with imperfect sample separation. This sample separation depends on initial per capita output. Using this method, no evidence of convergence is found in the wealthiest club of countries.

2.2.4 The role of technology spillovers.

The majority of the empirical literature reviewed in the last section has concentrated on the convergence, or lack of, using a Solow type growth model. In this type of model technology plays a very significant role in the sense that in the long run it determines the growth rate of the economies. Therefore, it is very important to study the process by which technology evolves. There has been a strand of the literature of growth and convergence which has focused on the idea of transfers of technology both across countries and across industries.

The idea of technology transfers had already been put forward by Abramovitz (1986) who, however, emphasizes that spillovers across countries will only have an effect if the receiving country possesses the technical and social capability to absorb and implement the new ideas. This idea was later developed further by Grossman and Helpman (1991a,b) and Aghion and Howitt (1992). These models endogenize the rate of technological change and consequently the rate of growth in the long run becomes an endogenous function of the model parameters (and disturbances).

There has been other models which emphasize the role of flows of ideas across countries. In this models, the process of long run growth is driven by the growth rate of the world knowledge. The diffusion of technology may take a long time, but, it will eventually translate into all countries having the same long run growth rate. Thus, growth rates differ across countries while the technologies in the different economies are catching up. Along these lines is the "idea gaps" model of Romer (1993) and also the models of Grossman and Helpman (1991a), Parente and Prescott (1994), Barro and Sala-i-Martin (1997) and Eaton and Kortum (1995). The role of technology spillover in these models is to prevent countries from diverging forever and are based on the assumption that imitation is easier and cheaper than innovation[24].

To be more specific, the role of technology spillovers in convergence can be illustrated with reference to a simple model along the lines of Bernard and Jones (1996a,b). In this simple model, it is assumed that the growth of technology in country i is a consequence of two different forces and also some random element. These two distinct forces are a consequence of, on one hand, the country's own efforts and on the other hand, the diffusion of technology from the most advanced country to the backward economy. Formally, then, technology in country i at time t, A_{it}, evolves according to the following structure

$$\Delta \ln A_{it} = g_i + \lambda_i \ln \left(\frac{A_{1t}}{A_{it}} \right) + \varepsilon_{it}$$

A_{1t} stands for technology in the most advanced country, and therefore, $\ln\left(\frac{A_{1t}}{A_{it}}\right)$ is the technological gap between the most advanced country and country i. The $\left(\frac{A_{1t}}{A_{it}}\right)$ieter λ_i shows the amount of technology that can be transferred from the most advanced country each time period and hence, it depends on the country's ability to imitate the technology developed somewhere else. The parameter g_i is the asymptotic rate of technology growth in country i and ε_{it} represents technology shocks. It is easily seen from this simple model, that if $\lambda_i > 0$, a gap in technology between two countries will temporarily raise technology growth in the backward economy. For a given λ_i, the bigger the gap, the higher the growth rate of technology will be. Additionally, for a given gap, the higher λ_i, the higher the growth of technology. This spillover of technology will

[24]Lucas (1993) discusses some models of growth and trade. He argues that the type of models consistent with episodes of very rapid growth are models of technology adoption and learning.

make the technology gap shrink. However, in this model, the two countries will only converge in the sense of the same growth rate of technology in the long run if both $\lambda_i > 0$ and $g_1 = g_i = g$. In this case, the gap between the levels of technology in both countries will eventually disappear and both countries will grow at the same common rate, g, in the long run. If $\lambda_i > 0$ but $g_1 \neq g_i$, countries will not share the same growth rate in the long run. Nevertheless, their levels of technology will not diverge indefinitely due to the presence of technology spillovers. However, if $\lambda_i = 0$, there will be no tendency for the gap in technology to disappear. If the asymptotic rates of growth in both countries are the same, $g_1 = g_i = g$, there would be no tendency for the levels of technology to converge and the technology gap will remain constant on average. Nevertheless, both countries will grow at the same rate, g, in the long run. If, however, the asymptotic rates of growth are different, both countries will grow at a different rate forever and the levels of technology in those countries will diverge.

There have been several empirical studies which support this idea of technology catch up. Most of them use Total Factor Productivity (TFP) as a measure of technology. Assuming a Cobb-Douglas production function of the form

$$Y(t) = A(t)\, L(t)^{(1-\alpha)}\, K(t)^{\alpha}$$

TFP is measured as[25]

$$TFP = (1-\alpha)\ln\left(\frac{Y(t)}{L(t)}\right) + \alpha\ln\left(\frac{Y(t)}{K(t)}\right) \tag{2.33}$$

[25]Note that if we assume that technology is labour augmenting, the measure of TFP will be slightly different. In Chapter 6, the measure of TFP employed comes from a Cobb-Douglas production function with labour-augmenting technology.

The choice of the proper factor shares depends on the assumptions used. If technology is assumed to be the same across countries and that the prices of the factors of production are equalized (as in the Heckscher-Ohlin model), factors shares should be the same across countries, so that the average factor shares across countries provide a good approximation to the value of α. However, if the technology of production differs among countries, one should use the factor shares specific to each country. The empirical results point to the latter.

Dowrick and Nguyen (1989) found that there is a tendency for the levels of TFP to catch up for a group of OECD countries using data from 1950 to 1985. They point out that if the levels of TFP tend to converge, the income levels will also tend to converge. However, this tendency of the income levels to converge may either not appear very clearly or appear exaggerated if the growth of factor intensity varies systematically with income. Wolff (1991) also studies the issue of catching up in TFP amongst other related issues. He restricts attention to the G7 countries covering the years 1870 up to 1979. Wolff found that TFP levels converged amongst the G7 countries, but convergence appeared to be much stronger after World War II than before. Dowrick (1992) studies a group of 113 countries. He finds evidence of productivity catch-up arguing that the most logical explanation for this is the presence of technology spillovers across countries. Dollar and Wolff (1994) find convergence of TFP in industrial sectors during the years 1963-1985. This phenomenon was found to be strong during the years 1963 to 1972 but became weaker after 1973.

Bernard and Jones (1996b, c and d) found that although at a country level, the levels of TFP are converging for a group of 14 OECD countries between 1970 and 1987, this is not the case at a more disaggregated level. They found no evidence of TFP convergence in the manufacturing sector. However, they found convergence in other sectors, in particular, the service sector. Thus, the convergence result found at an aggregate level seems to be a direct consequence of the sectors which are converging. However, Bernard and Jones were worried about the robustness of the results, particularly in relation to different measures of what they termed multifactor productivity levels (or technology). The different measures emerge due to different assumptions employed to identify the levels of productivity across countries in the base year, in their case 1970. The first measure is the standard TFP measure of equation (2.33). It was explained before, that the value of $(1 - \alpha)$ is computed as the average of the labour share across all years and countries. However, if the parameter α differs across countries, this measure will be misleading, since, in this case, countries with the same amount of inputs and the same technology will produce different output. Another problem relates to the production function being Hicks-neutral. In this case, changing the units of measurement for one of the inputs, might change the ranking of the country's productivity level, if the measures of α show very small differences across countries. Nevertheless, this problem is easily solved by adopting a Harrod-neutral (labour augmenting) production function. Bernard and Jones showed empirically, how the labour shares vary considerably not only across countries, but also across time. Having this in mind, they defined a new measure of technology, Total Technological Productivity (TTP). This measure ranks the countries according to which one would produce more output when countries use the same quantities of inputs. This measure however, is not without problems either.

There are other studies using alternative measures of technology. Maudos, Pastor and Serrano (2000) use a nonparametric frontier approach to calculate the Malmquist productivity index since they argue that TFP measures are misleading under the presence of inefficiencies. This approach breaks down changes in labour productivity into changes in efficiency and changes in technology. They found that there is a high level of inefficiency in the OECD economies. They also found that the main source of convergence across countries has been the higher rate of capital accumulation present in the poorer countries. Tsionas (2000) examines technology convergence in 15 European countries from 1960 to 1997. He measures productivity growth using what he termed country specific technical change index (CSTCI). These estimates of technical change are calculated by using a modified translog production function under nonconstant returns to scale. Additionally, the production functions have three inputs as opposed as the two normally used: capital, labour and imports. He justifies including the latter by arguing that the main bulk of imports are in fact intermediate goods which will need to be processed further. Therefore, if omitted from the production function, the estimates of technology will be biased. To test convergence across this group of European countries, Tsionas uses stationarity tests around a common long run trend. These include unit root tests, KPSS, Bayesian tests and a test based on fractional differencing parameter estimation for long run memory models. He finds that the results are sensitive to the type of test used. On one hand, the Phillips-Perron tests reject convergence, whereas the Bayesian approach and KPSS tests favour it. The results of long memory testing lie in the middle with different groups of countries for which convergence is found. In fact, the fractional differencing parameter is only less than a half (condition required for convergence) in only four of the countries. In addition, the asymptotic standard error associated with this parameter is large.

2.3 Reassessing the empirical evidence on growth and convergence; some illustrative econometric analysis of Penn World Tables data.

This section replicates some of the empirical studies that have been done in the past based on Lee *et al* (1995). They showed how the empirical regularity of a 2% convergence speed found by cross section studies is just due to, first, imposing restrictions that are not accepted by the data and second, to a bias in cross section estimation. They also showed how time series estimation implies higher speeds of convergence but towards each country own steady state, not towards a common one as imposed in cross section estimation. In addition they presented evidence of a unit root in output, pointing once more towards the rejection of convergence towards the same steady state across countries.

For the purpose of this empirical exercise, the data used are from the Penn World Tables (PWT) version 5.6 (20 November 1994).[26] They come from national accounts and have been scaled using ICP (International Comparison Programmes) benchmark studies which makes possible comparisons among countries. The measure of output per capita used in the empirical analysis is labelled RGDPL in these tables and corresponds to real GDP per capita in 1985 international prices. The savings rate is taken as the share of real private and public investment in real GDP (labelled I). Population growth rates are calculated from the total population figures (POP) given in this dataset. We use the same measure of human capital accumulation as Mankiw, Romer and Weil (1992), that is the percentage of working-age population that is in secondary school. The data are

[26]An older version of the data is explained in detail in Summers and Heston 1991.

annual and cover the period 1960 to 1989. A sample of 81 countries and two different subsamples of it are selected.[27]

1. a set of 81 non-oil producing countries
2. an intermediate group of 49 countries which excludes those countries whose population in 1960 was less than one million and countries which are thought to have poor data estimates
3. a group of 18 OECD countries with populations over 1 million.

These samples follow closely those of Mankiw *et al* (1992).[28] Sections 2.3.1 to 2.3.4 replicate some of the influential empirical studies on convergence with the above mentioned dataset and groups of countries.

2.3.1 Barro cross section regressions.

As noted earlier, the notion of capital in the neoclassical growth model can be easily extended from physical capital to include other forms of capital. Following Mankiw *et al* (1992), the following version of the Solow model will be estimated which allows for two types of capital, physical and human capital

$$\ln y_t - \ln y_0 = \left(1 - e^{-(1-\lambda)t}\right) \frac{\alpha}{1-\alpha-\beta} \ln(s_k) + \left(1 - e^{-(1-\lambda)t}\right) \frac{\beta}{1-\alpha-\beta} \ln(s_h) -$$

[27]Data on a further 21 countries was available over the sample period but excluded on the grounds of potential "unreliability" since in subsequent analysis (chapter 3) the estimated equations were unstable or the associated likelihoods were uninformative.

[28]The countries included in the OECD group are: Canada, USA, Japan, Austria, Belgium, Denmark, France, West Germany, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, U.K. and New Zealand. In addition to these countries, the intermediate group includes Cameroon, Kenya, Morocco, South Africa, Zimbabwe, Costa Rica, Dominican Republic, El Salvador, Guatemala, Honduras, Jamaica, Mexico, Nicaragua, Argentina, Bolivia, Brazil, Chile, Colombia, Paraguay, Peru, Uruguay, Venezuela, Bangladesh, Hong Kong, India, Indonesia, Korea Rep., Pakistan, Singapore, Sri Lanka and Syria. The non-oil producing group includes as well Angola, Benin, Bostwana, Burkina Faso, Burundi, Central African Republic, Congo, Egypt, Gambia, Ghana, Madagascar, Mali, Mauritania, Mauritius, Mozambique, Niger, Nigeria, Swaziland, Togo, Uganda, Zaire, Zambia, Barbados, Haiti, Trinidad & Tobago, Suriname, Jordan, Myanmar, Cyprus, Luxembourg, Malta and Papua N. Guinea.

$$-\left(1-e^{-(1-\lambda)t}\right)\frac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta)-\left(1-e^{-(1-\lambda)t}\right)\ln y_0 \quad (2.34)$$

where y_0 is income per capita in 1960, y_t is income per capita in 1989, $1-\lambda=(n+g+\delta)(1-\alpha-\beta)$ is the speed of convergence, s_k is the fraction of income invested in physical capital, s_h is the fraction of income invested in human capital. Thus, countries which invest a higher proportion of their output in physical and human capital will grow quicker.[29] This equation comes from solving equation (2.7) and then substituting the steady state value of output. The inclusion of human capital in the model is justified on account of α, the share of capital in income. It is generally agreed that α should be around a third. However, the estimated α from the regressions without human capital is almost two thirds. Including human capital in the model makes the share of total capital in output much larger and the value of α implied by the estimated parameters closer to a third. Tables 2.1, 2.2 and 2.3 show the results from running regressions of the form of equation (2.34) for the three samples described before and correspond to Table VI of Mankiw *et al* (1992). Here, the OLS estimation method is employed, imposing the restriction that the coefficients on $\ln(s_k)$, $\ln(s_h)$ and $\ln(n+g+\delta)$ sum up to zero.[30] It is assumed here that $g+\delta$ is 0.05, n is measured as the average rate of growth of population, s_k is the average share of investment in GDP, s_h is proxied by the percentage of the working age population that is in secondary school. The implied speeds of convergence for the regressions including human capital are (standard errors between brackets), 0.0139 (0.0033), 0.0196 (0.0057) and 0.0207 (0.0048) for the non-

[29] Hall and Jones (1997) argued that the reason why some countries invest different proportions of their output than others is related to the institutions and government policies of each country.

[30] The restriction is not rejected in any of the samples (see Tables 2.1, 2.2 and 2.3). The LM test statistics of the restrictions and their associated significance are 1.8562[0.17], 0.7893[0.37] and 0.0106[0.92] for the non-oil, intermediate and OECD samples respectively.

oil, intermediate and OECD countries which are in line with the speeds found usually in cross section studies. The estimates of α (the estimated share of capital in output) are 0.3874, 0.4598 and 0.4736 for the three groups, which are lower than without the inclusion of human capital and more in line with the value that was expected a priori. In section 2.3.3 however we shall show that these estimates are badly biased.

2.3.2 Pooled "fixed effects" cross section.

Following Islam (1995), the following version of the Solow model can be estimated

$$y_{i,t} = \gamma y_{i,t-1} + \sum_{j=1}^{2} \beta_j x_{it}^j + \eta_t + \mu_i + v_{it}$$

where

$$\begin{array}{ll}
y_{i,t} = \ln \tilde{y}(t_2) & y_{i,t-1} = \ln \tilde{y}(t_1) \\
\gamma = e^{-(1-\lambda)\tau} & \\
\beta_1 = \left(1 - e^{-(1-\lambda)\tau}\right) \frac{\alpha}{1-\alpha} & \beta_2 = -\left(1 - e^{-(1-\lambda)\tau}\right) \frac{\alpha}{1-\alpha} \\
x_{it}^1 = \ln(s) & x_{it}^2 = \ln(n + g + \delta) \\
\eta_t = g\left(t_2 - e^{-(1-\lambda)\tau} t_1\right) & \\
\mu_i = \left(1 - e^{-(1-\lambda)\tau}\right) \ln A(0) & \\
\tau = t_2 - t_1 & \lambda = 1 - (1-\alpha)(n + g + \delta)
\end{array}$$

using Least Squares Variable Estimation. The time series dimension of the data is divided here in six groups of 5 year intervals. Thus, $\tilde{y}(t_1)$ and $\tilde{y}(t_2)$ stand for output per capita at the beginning and at the end of each time interval and the rest of the variables are defined as in section 2.3.1. Tables 2.4, 2.5 and 2.6 show the results from this estimation with the datasets used here imposing the restriction that $\beta_1 = -\beta_2$.[31] The esti-

[31] The restriction cannot be rejected in any of the three samples of countries used here (see Tables 2.4, 2.5 and 2.6). The LM test statistics of the restrictions and their associated significance levels are 2.2552[0.13], 1.0397[0.31] and 1.7576[0.18] respectively for the non-oil, intermediate and OECD group of countries.

mated speeds of convergence are (standard errors between brackets) 0.0401 (0.0065), 0.0415 (0.0089) and 0.0749 (0.0142) for the non-oil, intermediate and OECD samples respectively. Note that these speeds of convergence are higher that the ones found in section 2.3.1. The estimates of α are 0.3870, 0.4431 and 0.1580, in line with the value expected a priori without the need for human capital although the estimated α for the OECD countries seems to be quite small. However Lee *et al* (1995) showed that if the assumption of a common growth rate imposed in this type of estimation is not true, then the fixed effect estimator will give inconsistent estimates of λ and that the

$$P \lim_{N,T \to \infty} \left(\hat{\lambda} \right) = 1$$

irrespective of the true value of λ.

2.3.3 Heterogeneous panel.

Lee *et al* (1995) estimated equation (2.7) using exact maximum likelihood which restricts the coefficient of the lagged dependent variable to be less than 1, i.e. imposes stationarity. Table 2.7 shows the results of estimating the same equation but using stacked OLS.[32] This involves stacking the observations for all the countries and using a dummy for each country. The results obtained are quite similar to those found by Lee *et al* (1995). The estimated speeds of convergence are 19.82%, 17.53% and 16.18% for the non-oil, intermediate and OECD group of countries respectively, which are a lot higher than the 2% implied by cross section estimation or the 4-7% found using "fixed effects" cross section. However, as noted earlier, we have to take into account that these

[32] Although the estimation is done using all the countries in the sample, the table is calculated with the exclusion of countries whose implied g differs more than 3.5 standard deviations of the mean value.

speeds of convergence refer to countries adjusting towards their own steady state, not towards a common one as it is implied by cross section regressions and that the time series estimates of the speed of convergence are probably biased upwards due to the small sample bias. The restriction of a common λ and g across countries as implied by cross section estimation is rejected in all the samples except from the OECD group. The LM test statistics of the restrictions and their significance levels are 341.3813[0.32], 209.5021[0.00] and 37.3813[0.32] for the non-oil, intermediate and OECD groups of countries respectively. Imposing this restriction (see Table 2.8) causes the estimated speed of convergence to drop to 4.00%, 2.85% and 8.18% for the non-oil, intermediate and OECD groups respectively.

There is a second problem with cross section estimation. Lee *et al* (1995) showed that the estimated coefficients of the cross section regression are biased and that under the assumptions of a common and stationary λ the asymptotic bias can be written as

$$\begin{aligned} \underset{N \to \infty}{P\lim} \hat{\lambda}^T &= 1 - \frac{q}{1+q}\left(1 - \lambda^T\right) \\ &= \left(\frac{1}{1+q}\right) + \left(\frac{q}{1+q}\right)\lambda^T \end{aligned} \tag{2.35}$$

where

$$q = \frac{\tau^2/\left(1 - \lambda^2\right)}{\sigma_a^2 + \phi'\Sigma_w\phi}$$

Thus $\hat{\lambda}^T$ is a weighted average of the true λ^T and 1, with weights $\frac{q}{1+q}$ and $\frac{1}{1+q}$ respectively where q is the ratio of the long run variance of the time series shocks, $\tau^2/(1-\lambda^2)$, to the cross section variance, $\sigma_a^2+\phi'\Sigma_w\phi$. Typically the long run variance is small compared to the cross section variance, therefore q will be small and the bias large. To illustrate explicitly the size of the bias in cross section "Barro" regressions we shall work here with the group of 22 OECD countries. Typically, this type of studies regress the difference of the logarithm of per capita GDP between the end and the beginning of the sample on the logarithm of per capita GDP at the beginning. The OECD is a homogeneous group for which cross section studies find convergence. The scatter plot of the points in our sample (see Figure 2.4) suggests a negative relationship between growth and the level of initial income which seems to indicate strong evidence of convergence. However this is just a representation of the regression line, and therefore it suffers from the same problems. Using the results from the OLS estimation we calculate the size of the bias (vertical distance between the two lines in Figure 2.5). It is evident from this plot that no matter what the true value of λ is, cross section "Barro" regressions will always give a value of λ close to unity.

2.3.4 Unit roots.

Hitherto, it has been assumed that the logarithm of output per capita is trend stationary. This assumption, however, can be tested directly using the following augmented Dickey-Fuller version of equation (2.7)

$$\Delta \ln(\tilde{y}_{it}) = \mu_i + \theta_i t - (1 - \lambda_i) \ln(\tilde{y}_{i,t-1}) + \sum_{j=1}^{p_i} \rho_{ij} \Delta \ln(\tilde{y}_{i,t-j}) + \varepsilon_{it} \tag{2.36}$$

for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. Following Lee *et al* (1995) the data is demeaned to get rid of any possible common trends affecting all countries in our sample. This is accomplished by subtracting from the logarithm of per capita output the average across countries at each time period. Using a lag length (p_i) of 4 for all the countries, the hypothesis of a unit root in the logarithm of output per capita is rejected for 4 countries: Zimbabwe, Dominican Republic, Guatemala and Japan with t-values -4.0247, -3.7786, -4.3396 and -3.8382 respectively (critical value: -3.6119). Using the Akaike Information Criterion (AIC) to choose the lag length of the augmentation, the null hypothesis of a unit root in the logarithm of per capita output is rejected for 8 countries: Kenya, Uganda, Zimbabwe, Dominican Republic, El Salvador, Guatemala Peru and Turkey. The t-values (the number of augmented lags p_i selected by AIC are between brackets) are -3.8223 ($p_i = 0$), -3.9276 ($p_i = 2$), -4.0247 ($p_i = 4$), -3.7786 ($p_i = 4$), -4.3760 ($p_i = 1$), -4.3396 ($p_i = 4$), -4.6972 ($p_i = 1$) and -4.1393 ($p_i = 3$) respectively. Using the Swartz-Bayesian criterion to choose the lag length of the augmentation the null hypothesis of a unit root in the logarithm of output per capita is rejected for 7 countries, with the same t-values and number of augmentations as with the AIC, the only country which differs is Zimbabwe for which the SBC chooses no augmentations. This gives a t-value of -2.4148 and therefore, a unit root in the logarithm of output per capita cannot be rejected for this country.

The Dickey-Fuller test however has low power but we can make use of the panel structure of the data and apply the "t-bar test" proposed by Im, Pesaran and Shin (1995) to test for unit roots in panels. This test is based on the average value of the augmented Dickey-Fuller statistics obtained across countries. The t-bar statistic is calculated as

$$\bar{t}_{NT} = \frac{\frac{1}{N}\sum_{i=1}^{N} t_{iT}\left(p_i, \hat{\rho}_i\right) - \frac{1}{N}\sum_{i=1}^{N} E\left[t_T\left(p_i, 0\right)\right]}{\sqrt{\frac{1}{N^2}\sum_{i=1}^{N} V\left[t_T\left(p_i, 0\right)\right]}}$$

where the values of $E\left[t_T\left(p_i,0\right)\right]$ and $V\left[t_T\left(p_i,0\right)\right]$ are tabulated in Im, Pesaran and Smith (1995). Under the null hypothesis, when N and T are large and N/T goes to zero, this statistic has a normal distribution. Fixing the length at $p_i = 4$ the t-bar calculated statistics for output are -0.1909, 0.8382 and 0.4597 for the non oil, intermediate and OECD group of countries respectively with critical values; somewhere between -2.31 and -2.37 for the non-oil group, -2.37 for the intermediate group and -2.48 for the OECD group.[33] It is clear then, that the hypothesis of a unit root in the logarithm of per capita output cannot be rejected for any of the groups. Using AIC to choose the lag length of the augmentation, the t-bar calculated statistics are -2.461, -1.7640 and -0.3948 for the non-oil, intermediate and OECD groups respectively with the same critical values as before. In this case, the hypothesis of a unit root is rejected for the non-oil group, but, it cannot be rejected for the intermediate and the OECD groups. Using SBC to choose the lag length of the augmentation the t-bar calculated statistics for output are -1.070, -0.7610 and -0.1255 for the non oil, intermediate and OECD group of countries respectively. Thus, the hypothesis of a unit root in the logarithm of output per capita cannot be rejected for any of the groups. Therefore, this seems to indicate the presence of a unit root in the logarithm of output per capita.

2.4 Some extensions to the Overlapping Generations model.

The drawback of the models presented in section 2.2.2 is that they are closed economy models. In a closed economy model, all inter-relations between countries are as-

[33] The critical values of this statistic are calculated depending on the number of observations, T, and the number of countries in the panel, N. The critical value -2.31 corresponds to $N = 100$, whereas the critical value -2.37 corresponds to $N = 50$. Unfortunately, there are no critical values tabulated between these two.

sumed away, but it is difficult to think of a world of countries existing in isolation without influencing the rest of the countries. It is clear that any analysis of growth would benefit greatly from a more realistic theoretical model which takes into account these interactions. Economic theory suggests that open economies are characterized by a shorter transition to the steady state because of the international mobility of capital and technology transfers. Taylor (1999) argues that for a group of seven countries during the period 1870-1914, the standard empirical model with physical and human capital is unable to explain the observed convergence pattern. He proposes an alternative open-economy neoclassical model with capital and labour migration. There have been attempts to extend the RCK model to an open economy setup. If perfect capital mobility is assumed, extending the model leads to the result that the most patient country will eventually own everything and consume nearly all of the world's output (see Barro and Sala-i-Martin (1995) for example). This clearly contradicts the actual facts and comes about because of the assumption of infinitely lived households. The OG model, however, does not exhibit this unpleasant feature. Barro, Mankiw and Sala-i-Martin (1995) developed an open economy model of partial capital mobility and their main conclusion was that under certain assumptions the speed of convergence is similar to that predicted for a closed economy.

Sections 2.4.1 and 2.4.2 present two possible extensions to the closed economy OG model. These extensions illustrate how the patterns of growth differ from those of a closed economy once the interactions between countries are taken into account. This minor extensions will motivate the empirical work in chapters 5 and 6. First, the basic model is extended to an open economy setup with imperfect capital markets (Obstfeld and Rogoff (1996)). In these circumstances, the interactions among countries arise directly through the balance of payments. The difference between this open economy setting and the closed economy setting is that countries are not restricted to use their own capital. Rather, they are allowed to borrow from abroad. Intuitively, this increases the speed of convergence to the equilibrium compared to the very slow speed of convergence obtained in theoretical closed economy models with reasonable parameter values. Second, and as an alternative, the basic model is extended to allow for technology spillovers across countries; that is, this second model is based on a closed economy, but technology is allowed to flow from the most advanced countries to the more backward ones. This model captures the idea of "technology catch up" by characterizing cross country technology differences via a simple probability distribution. This distribution evolves over time as technology improves in all countries and as the less advanced countries (that is, those in the lower tail of the distribution) increase more quickly their level of technology, through imitation, for example. Consequently, technological advances in one economy spill over the rest of the countries. Intuitively, if the asymptotic rates of technology growth in all the countries are the same, countries will exhibit convergence in their levels of technology. Countries with relatively low levels of technology will be able to increase their level of technology relatively quickly due to the spillover effect, and since the asymptotic rates of growth are the same for all the countries, technology across countries will reach the same level eventually. In these circumstances, the level of technology (and therefore the levels of capital and output) will be higher for the countries which benefited from the technology spillovers than it would have been otherwise. If the asymptotic rates of technology growth are different across countries, again, the levels of technology across countries will be pulled together by the spillover effects. Obviously, in this case, there would not be convergence to a single common level of technology but, because of the spillover effects, technologies across countries will not diverge forever either.

2.4.1 Extension of the OG model to incorporate a balance of payments constraint.

In an open economy, much of a country's macroeconomic activity is connected in one way or another with the country's intertemporal trade which is measured by the current account of the balance of payments. A country's trade deficit is balanced by a foreign accumulation of domestic assets; that is, external liabilities that will have to be serviced at some point in the future by means of a future current account surplus. With this in mind, balance of payments constraints can be interpreted as the constraints arising out of the ability of a country to borrow in the international market to cover a current account deficit. This extension intends to consider the role of cross country interdependence on growth as exerted by a balance of payments constraint of this sort.

In this dynamic context, the intertemporal optimizing approach to macroeconomics, sometimes called the "New Open Macroeconomics" approach, provides a suitable framework for studying the effects of different types of shocks. In this setup, the function of the current account of the balance of payments is to smooth consumption in the presence of shocks to, for example, output (see Sachs (1981), Bean (1991), Frenkel and Razin (1996) and Obstfeld and Rogoff (1996)). If this is the case, as we shall see, current account deficits will not constrain growth in the long run. If the international capital market is perfect, then the balance of payments constraint would be at most just an intertemporal solvency condition. However, in a situation in which the international capital market is imperfect, there might be additional borrowing constraints each period. Although in these circumstances long run growth will not be affected, there might be implications for cross country convergence. Whether convergence across countries is achieved or not in the long run will depend on whether savings together with the maximum borrowing amount are sufficient to reach the efficient capital stock.

Extending the intertemporal optimizing RCK model to an open economy setup is not straightforward in terms of its economic implications since it leads to the counterfactual result that the most patient country eventually owns everything and consumes nearly all of the world's output.[34] This is shown in section 2.4.1.1which generalizes the RCK model of section 2.2.2.2 to the case of an open economy and discusses the unrealistic implications of this model. To generate more realistic economic outcomes, the model requires further adjustment and section 2.4.1.2, therefore, generalizes the OG model of section 2.2.2.3 to the case of an open economy with imperfect capital markets which provides more realistic conclusions.

2.4.1.1 The open economy RCK model.

Generalising the RCK model of section 2.2.2.2 to the case of an open economy is straightforward from a mathematical point of view. Under the assumption of a small open economy, the differential equation for the capital per effective units in equation (2.11) becomes

$$\overset{\bullet}{(k(t) - f(t))} = k(t)^{\alpha} - (r^{*} + \delta) f(t) - (n + g + \delta)(k(t) - f(t)) - c(t) \quad (2.37)$$

where $f(t)$ is the country's net debt to foreigners per effective units at time t and r^{*} is the world interest rate. If $f(t) = 0$, then equation (2.37) simplifies to the closed economy

[34] See, for example, Barro and Sala-i-Martin (1995) or Obstfeld and Rogoff (1996).

equation (2.11). The differential equation for consumption per effective units, equation (2.12), remains however unchanged and is reproduced below for illustrative purposes

$$\frac{\dot{c}(t)}{c(t)} = \frac{\alpha k(t)^{\alpha-1} - (\rho + \theta g + \delta)}{\theta}$$

There are several problematic features of this model which make its predictions unrealistic. Firstly, because the economy is open and there is no constraint on the amount of borrowing, the small country is able to borrow enough capital to get to the steady state in just one period and the country's interest rate and the world interest rate equalize immediately. This implies an infinite speed of convergence to the steady state for capital and therefore for output, which is obviously problematic in the light of the actual facts. Secondly, in the closed economy model the country's interest rate adjusts to equal the effective rate of time preference, so that $r = \rho + \theta g$. In this case, consumption per effective units tends to a constant in the steady state. However, in the open economy setup, the country's interest rate is pegged to the world interest rate. Assuming that all the countries have the same population and technology growth rates $(n + g)$, then consumption per effective units will tend to zero if the country is impatient, so that $r < \rho + \theta g$.[35] Consequently, an impatient country borrows early so it can benefit from high consumption early on. However, it pays the price later in the form of low consumption growth. This is another feature of the model that is problematic. At the same time, if assets per effective units $(k(t) - f(t))$, start positive they fall to zero and then become negative overtime. Because the country is impatient it mortgages all of its capital and labour income asymptotically. Again a problematic feature. The key fea-

[35] If $r > \rho + \theta g$ the economy will keep on accumulating assets until it obviously violates the small economy assumption.

ture for the behaviour of consumption and assets in this infinite horizons model is that assets are implicitly inherited by the descendants, therefore people are able to borrow against this future income. In a finite horizons model, such as the OG model of the section 2.2.2.3, people give no weight to their descendants either in the utility function or in their budget constraints. This is crucial for the behaviour of consumption and assets per effective units. In a finite horizon model the speed of convergence to the equilibrium is still infinite, but the behaviour of consumption and assets per effective units is more reasonable than in the infinite horizons model. The key property in this model is that the effective rate of time preference, that is $\rho + \theta g$, is an increasing function of the assets over consumption. The following section presents an open economy OG model with imperfect capital markets.

2.4.1.2 The open economy OG model with imperfect international capital markets.

The intertemporal model presented here is based on the OG model in Obstfeld and Rogoff (1996) to which population and technology growth is added. A small open economy facing a fixed world interest rate r^w is assumed. Individuals live two periods but they only work in the first. Their savings are either invested in the domestic market or lent in the world market. They can also borrow from the world market, although borrowing is limited to a fraction of current output. Countries may have different autarky interest rates, perhaps because of differences in preferences, although they share the same technology. Production is carried out following a Cobb-Douglas production function, population grows at rate n and technology grows at g; see equations (2.17) to (2.20). The domestic markets are perfectly competitive and capital depreciates at a constant rate δ which is set equal to 1 as before. Thus, in equilibrium, the two equations obtained for the marginal products of labour and capital are as before equations (2.21) and (2.22) respectively.

In this model, borrowing and, equivalently, the net stock of foreign assets, F_t, is limited to a fraction ξ of total output; that is, $F_t \geq -\xi Y_t$. This can be written in per capita terms as

$$\tilde{f}_t \geq -\frac{\xi}{(1-\alpha)} A_t w_t \tag{2.38}$$

where $\tilde{f}_t$ stands for per capita stock of foreign assets. When this borrowing constraint is binding, the domestic interest rate, r_t, can exceed the world interest rate, r^w.

A young individual's utility function is given again by equation (2.23). The budget constraints for the young and the old are now given by

$$K_{t+1} + F_t = A_t L_t w_t - L_t \tilde{c}_t^Y \tag{2.39}$$

and

$$L_t \tilde{c}_{t+1}^O = (1 + r_{t+1}) K_{t+1} + (1 + r^w) F_t \tag{2.40}$$

These can be combined by eliminating K_{t+1} to get the following

$$\tilde{c}_t^Y + \frac{\tilde{c}_{t+1}^O}{1 + r_{t+1}} = A_t w_t + \frac{(r^w - r_{t+1})}{1 + r_{t+1}} \tilde{f}_t \tag{2.41}$$

That is, the budget constraint again states that the present value of a young person's lifetime consumption equals the present value of his lifetime income. This lifetime income now includes any extra profit that can be made because of the interest rate differential between the domestic and foreign economies. Hence, a young individual maximizes the utility function (2.23) subject to the budget constraint (2.41) and the "slackness condition" (2.38)

$$
\begin{aligned}
\mathcal{L} &= \ln\left(\tilde{c}_t^Y\right) + \beta \ln\left(\tilde{c}_{t+1}^O\right) + \lambda_1 \left[A_t w_t + \frac{(r^w - r_{t+1})}{1 + r_{t+1}} \tilde{f}_t - \tilde{c}_t^Y - \frac{\tilde{c}_{t+1}^O}{1 + r_{t+1}}\right] \\
&\quad + \lambda_2 \left[\frac{\xi}{(1-\alpha)} A_t w_t + \tilde{f}_t\right]
\end{aligned}
$$

- **The model solution.**

The Kuhn-Tucker necessary conditions for maximization of the lagrangian are

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \tilde{c}_t^Y} &= \frac{1}{\tilde{c}_t^Y} - \lambda_1 = 0 \\
\frac{\partial \mathcal{L}}{\partial \tilde{c}_{t+1}^O} &= \frac{\beta}{\tilde{c}_{t+1}^O} - \frac{\lambda_1}{1 + r_{t+1}} = 0 \\
\frac{\partial \mathcal{L}}{\partial \tilde{f}_t} &= \lambda_1 \frac{(r^w - r_{t+1})}{1 + r_{t+1}} + \lambda_2 = 0
\end{aligned}
\tag{2.42}
$$

together with the budget constraint (2.41) and the slackness condition (2.38). This last condition states that either the borrowing constraint is binding or $\lambda_2 = 0$, in which case equation (2.42) implies that $r_{t+1} = r^w$. Hence, there are three cases to be considered depending on how long it takes for the country's autarky interest rate to adjust to the world interest rate, if it does at all. *Case 1* portrays a situation in which there is equalisation of interest rates in just one period. If $r_{t+1} < r^w$ the country will be a creditor in the steady state. Even in the case in which the country needs to borrow to attain the efficient level of capital, this might not be enough to hit the borrowing constraint. Hence the interest rates would equalize in just one period. *Case 2* and *Case 3* illustrate the case in which $r_{t+1} > r^w$ initially. However, in *Case 2,* the economy cannot reach the unconstrained efficient level of capital per effective units because savings from wages plus the maximum borrowing amount are not sufficient to attain the unconstrained level. In these circumstances the country's interest rate will always be higher than the world interest rate. In *Case 3*, however, savings from wages plus the maximum borrowing amount are sufficient to reach the unconstrained level of effective capital stock and, therefore, there would be equalisation of the interest rates. Nevertheless, the equalisation can take one or more periods depending on the initial level of wages.

Case 1: r_{t+1} is equal to r^w after one period.

If, at the country's autarky steady state, its net marginal product of capital is less than the world interest rate, the country will be a net creditor. Even if it needs to borrow to reach the efficient capital stock to equate its net marginal product of capital to r^w, the amount might be low enough not to hit the borrowing constraint. In this case, the domestic interest rate moves to r^w in one period (see *Case 3* for a more detailed exposition of this point). Therefore, in this case, the economy adjusts to the long run equilibrium in half the lifetime of an individual. In this case, the steady state level of capital per effective worker is given from equation (2.22) by

$$k_U^* = \left(\frac{\alpha}{1+r^w}\right)^{\frac{1}{1-\alpha}} \tag{2.43}$$

It is important to compare this steady state level of capital per effective units with the steady state value of the closed economy model in section 2.2.2.3. Whether the open economy unconstrained steady state k_U^* is higher or lower than the steady state value of k^* of the closed economy in equation (2.32) depends on whether the equilibrium interest rate in the closed economy, r, is higher or lower respectively than the world interest rate, r^w.

Solving the maximization problem and using equation (2.21), consumption per effective units of the young in equilibrium is

$$c_t^Y = \frac{(1-\alpha)}{(1+\beta)} k_t^\alpha \tag{2.44}$$

which will adjust to its steady state value

$$\left(c^Y\right)^* = \frac{(1-\alpha)}{(1+\beta)} \left(\frac{\alpha}{1+r^w}\right)^{\frac{\alpha}{1-\alpha}}$$

in just one period. Thus we get the familiar result that variables measured in effective units are constant in the steady state, implying that the variables in levels are growing at $n+g$; that is, the population plus the technology growth rates. Using equations (2.21), (2.39) and (2.44) and the fact that the production function in equation (2.17) can be written as $Y_t = A_t L_t k_t^\alpha$, we get the steady state debt-output ratio

$$\left(\frac{F}{Y}\right)^* = \frac{\beta(1-\alpha)}{1+\beta} - (1+g)(1+n)\frac{\alpha}{1+r^w} \tag{2.45}$$

which depends on β, that is, it depends on how impatient the country is.[36] The steady state current account-output ratio is given by

$$\left(\frac{CA}{Y}\right)^* = \left(\frac{F}{Y}\right)^* - \frac{A_{t-1}L_{t-1}}{A_t L_t}\left(\frac{F}{Y}\right)^*$$

[36]Note that the ratio of any two variables, and in particular the ratio of any variable to output, will be constant in the steady state because all the variables are growing at the same rate, $n+g$, once they have reached the steady state.

Therefore, substituting equation (2.45) the following steady state current account output ratio is obtained

$$\left(\frac{CA}{Y}\right)^{*} = (n+g+ng)\left[\frac{\beta(1-\alpha)}{(1+\beta)(1+g)(1+n)} - \frac{\alpha}{1+r^{w}}\right]$$

The country will then be a creditor in the steady state if the country's autarky steady state interest rate is below r^w, or a debtor if the country's autarky steady state interest rate is above r^w. In the latter case, the lower β, that is, the more impatient a country is, the higher the debt-output ratio.

Case 2: r_{t+1} is permanently higher than r^w.

In this case, individuals will always choose to borrow the maximum amount so that the borrowing constraint is always binding and therefore $\tilde{f}_t = -\frac{\xi}{(1-\alpha)}A_t w_t$. Solving the maximization problem, the consumption per effective units of the young is given by

$$c_t^Y = \frac{1}{1+\beta}\left[w_t + \frac{(r^w - r_{t+1})}{1+r_{t+1}}f_t\right]$$

This equation, together with the binding borrowing constraint and equation (2.21), gives the following equilibrium relationship

$$c_t^Y = \frac{(1-\alpha)}{(1+\beta)}\left[1 - \frac{\xi}{(1-\alpha)}\frac{(r^w - r_{t+1})}{1+r_{t+1}}\right]k_t^{\alpha} \tag{2.46}$$

The extra term on the right hand side of this equation compared to equation (2.44) reflects the extra profit that can be made when the borrowing constraint is binding by borrowing in the world market and investing at home. Using equations (2.39), (2.21), (2.46) and the binding borrowing constraint, capital per effective labour can be written as

$$k_{t+1} = \frac{1}{(1+g)(1+n)(1+\beta)} \left[\beta(1-\alpha+\xi) + \frac{\xi}{\alpha}(1+r^w) k_{t+1}^{1-\alpha} \right] k_t^\alpha \tag{2.47}$$

It is easy to see from this equation that an easing of the borrowing constraint increases the growth rate of the capital stock to the extent that ξ is higher (assuming that k_t lies below k_U^*) because it lowers the domestic interest rate. Setting $k_{t+1} = k_t = k_C^*$ in equation (2.47), the steady state capital stock per unit of effective labour with constrained borrowing is given by

$$k_C^* = \left[\frac{\alpha\beta(1-\alpha+\xi)}{\alpha(1+\beta)(1+g)(1+n) - \xi(1+r^w)} \right]^{\frac{1}{1-\alpha}} \tag{2.48}$$

Again, as in *Case 1*, variables measured in effective units are constant in the steady state, so that the variables in levels are growing at the population plus the technology growth rates, $n + g$. However, even though the actual growth rates of the variables are the same in the long run, the levels of capital and therefore of output, might not be the same. Using the facts that $w_U^* = (1-\alpha)(k_U^*)^\alpha$ and $(1+r^w) = \alpha(k_U^*)^{\alpha-1}$, the condition $k_C^* < k_U^*$ can be written as

$$\frac{1}{(1+g)(1+n)} \left[\frac{\beta w_U^*}{1+\beta} + \frac{\xi}{(1-\alpha)} w_U^* \right] < k_U^* \tag{2.49}$$

This equation states that the economy cannot reach the unconstrained steady state if, given w_U^*, savings from wages plus the maximum borrowing amount per effective worker (scaled down by the gross output growth rate) are not sufficient to reach the efficient capital stock. Obviously, in this case, the lower β, the more likely that the nonconvergence condition depicted in equation (2.49) is fulfilled.

Obviously, from the constraint on capital flows, the steady state debt-output ratio will be

$$\left(\frac{F}{Y}\right)^{*} = -\xi$$

and the current account output ratio will be

$$\left(\frac{CA}{Y}\right)^{*} = \left(\frac{F}{Y}\right)^{*} - \frac{A_{t-1}L_{t-1}}{A_t L_t}\left(\frac{F}{Y}\right)^{*} = -\frac{(n+g+ng)\,\xi}{(1+g)\,(1+n)}$$

The country will then be a debtor in the steady state.

Case 3: r_{t+1} is initially higher than r^{w} but then converges towards it.

If the non convergence condition in equation (2.49) is not satisfied, the economy will converge to k_U^* in the long run. If the initial wage rate per effective labour w_0 satisfies the condition

$$\frac{1}{(1+g)\,(1+n)}\left[\frac{\beta w_0}{1+\beta} + \frac{\xi}{(1-\alpha)}w_0\right] \geq k_U^* \tag{2.50}$$

then convergence will occur in one period. This is because savings, together with the allowable borrowing amount, are enough to finance the efficient capital stock per effective labour; otherwise it will take more than one period.

• The dynamics of the extended OG model.

The model presented in *Cases 1* to *3* above is deterministic, but it can be thought of as a simplification of a stochastic model. Intuitively, a positive and permanent shock to technology in the domestic economy will lead to an increase in investment due to the increase in the productivity of capital. At the same time, the shock will lift the constraint, and, therefore, this will bring about an increase in the capital flows to finance this investment. In this model, consumption is a fixed proportion of output, since a positive shock to technology increases current output, current output and current consumption will increase by the same amount. Also, the permanent increase in the level of technology will increase the expected future investment path, which in turn will lead to a larger future capital stock and a larger future income. Thus permanent income and, along with it, current consumption should rise by more than the current income. The current account is the change in capital flows. Hence, it worsens but, since output increases, the current account output-ratio might rise or fall. Thus, in an stochastic model, countries will switch from being constrained and not being constrained depending on the shocks to technology.

If, however, the shock to technology is transitory, current income and current consumption will increase but only for one period, since it will not change either current or future investment. Therefore, it will have some effect on current consumption but the effect will be weak. Hence, there will also be a very weak effect on the current account which will last only one period.

The differences of the adjustment paths of the capital per effective units, k_t, between the stochastic and deterministic model, in the open economy are illustrated in Figure 2.6. In order to generate these simulations, it is noted that the OG model, each period lasts around 25 years. Therefore, if we assume that the asymptotic growth rate of technology is around 2% per year, the actual value of g should be around 0.64; that is, the growth compounded over the 25 years. However, the OG model can be extended to more than two periods with the main conclusions remaining unchanged. Since quarterly data is used in chapters 5 and 6, it is important to see what type of dynamics this model generates with reasonable parameter values if each time period represented a quarter.

The adjustment paths for the deterministic series (the solid line) are simulated for four different cases. The following parameter values are common for the four cases, $g = 0.005$, $n = 0.005$, $\alpha = 0.35$ and $r^w = 0.03$. With these parameter values, the world interest rate is lower than the country's autarky interest rate in all three cases. The rest of the parameters are as follows

	Case A	*Case B*	*Case C*	*Case D*
β	0.7	0.7	0.6	0.5
ξ	0.3	0.1	0.1	0.1

The parameters in *Case A* do not satisfy the non-convergence condition in equation (2.49) but they meet the condition in equation (2.50), therefore the economy converges to the unconstrained steady state in only one period (see Figure 2.6, *Case A*). In *Case B*, the borrowing constraint is tightened. With these parameter values the nonconvergence condition in equation (2.49) is still not satisfied so that the economy converges again to the unconstrained steady state. However, the condition in equation (2.50) is not satisfied this time, and therefore, it takes longer than one period (see Figure 2.6, *Case B*). *Case C* is the same as *Case B* but with a lower value of β. This case, therefore, represents a more impatient country, but not impatient enough to satisfy the nonconvergence condition. In this case, the economy, converges again to the unconstrained steady state, but, it takes longer than in *Case B* (see Figure 2.6, *Case C*). *Case D* shows the consequences of an even lower value of β. Since the country is now more impatient, it tries to borrow a higher amount than before, however, this time, the nonconvergence condition in equation (2.49) is satisfied. Therefore, the economy is not able to converge to the unconstrained level of capital per effective units (see Figure 2.6, *Case D*).

The stochastic series (the dashed lines) are simulated using the same parameter values but instead of generating technology with the deterministic process in equation (2.20), it is generated using the following stochastic process[37]

$$\ln A_t = \ln A_0 + gt + u_t$$

where u_t follows an AR(1) process

$$u_t = \rho u_{t-1} + \varepsilon_t$$

In all three cases ρ is set equal to 1 and ε_t are drawn from a normal distribution with mean zero and standard deviation σ equal to 0.1. The simulated paths for the different scenarios can be seen in Figure 2.6. One important difference between the stochastic and the deterministic model is that in the deterministic model, if an economy is constrained, in the sense that the nonconvergence condition (2.49) is satisfied, it will always be constrained. However, in the stochastic model, a positive shock to technology in the domestic economy will lift the constraint as it was explained later. Therefore, in the stochastic model there are changes between the constrained regime and the unconstrained regime. These changes will be more frequent the higher the size of the shocks, σ.

[37] Taking logarithms and assuming that g is a small number, the deterministic process for technology in equation (2.20) can also be written as $\ln A_t = \ln A_0 + gt$.

Summing up, in the open economy OG model, the growth rate of output in the long run will be the same even if the economy is constrained in the world capital market. However the dynamics of adjustment towards the steady state equilibrium will be different because of the period by period borrowing constraint in the world market.

2.4.2 Extension of the OG model to incorporate technology spillovers.

The OG model of section 2.2.2.3 assumes that technology is exogenously determined and that it grows at a constant rate, g, per period in each country. It is likely, however, that economies for which the level of technology is low compared to the rest of the economies can benefit from technology spillovers; that is, they can grow temporarily quicker by adopting technologies already discovered in the most advanced economies (see section 2.2.4). In this section, the OG model of section 2.2.2.3 will be extended to allow for technology transfers between countries and, therefore, for technology in any one country to catch up with the technology in the most advanced countries if it falls excessively behind them. The model of technology catch up is constructed along the lines of Bernard and Jones (1996). They built a model in which the technology gap between a country and the most advanced country is a function of the lagged gap in productivity. This technology gap will increase the relative growth rate of the country in possession of the lower level of technology. Therefore, provided that the asymptotic rates of technology growth in both countries are the same, there will be convergence. In this section, the same idea is incorporated with the difference that it is not only the technology in the most advanced country that matters; the whole distribution of technology across countries is of importance.

2.4.2.1 Modelling technological spillovers.

It is assumed that the logarithm of technology ($\mathbf{a}_t$) across countries at time t follows a certain distribution which evolves over time.[38] A model of technology catch-up for country i can be written as

$$\triangle a_{i,t+1} = g_i + \lambda_i SPILL_{i,t} + \varepsilon_{i,t+1} \tag{2.51}$$

where g_i is the asymptotic rate of technology growth of country i; that is, it represents the technology growth of country i that arises as a result of innovation in that particular country. $SPILL_{i,t}$ measures the amount of technology that a country can adopt by imitation of other countries' technologies and, therefore, it will always be positive (or zero in the particular case that all the technologies of all the countries are the same). λ_i is a parameter between zero and one which represents the speed of technology catch up. Obviously, if $\lambda_i = 0$ there will be no spillovers, and, if $\lambda_i = 1$, country i will be able to adopt all the value of $SPILL_{i,t}$ in just one period. Thus, $\lambda_i SPILL_{i,t}$ represents the size of the technology spillover for country i at time t from which the country can benefit at time $t+1$. $\varepsilon_{i,t}$ is a country specific technology shock which is assumed to be independent across countries and time. Additionally, $\varepsilon_{i,t}$ is assumed to follow a normal distribution with mean zero and standard deviation σ_i.

The difference between this model and that of Bernard and Jones, lies on the definition of $SPILL_{i,t}$. In Bernard and Jones the spillover is a function of the productivity gap between the most advanced economy and economy i. Here, however, it is defined with reference to the average level of technology in the countries that are more ad-

[38] The way in which the distribution of the logarithm of technology evolves over time will be made explicit later in this section.

vanced than country i. More formally, the spillover for country i at time t is defined as

$$SPILL_{i,t} = E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right] \tag{2.52}$$

where $E[\bullet]$ is the expectations operator across countries. Therefore, the further the logarithm of technology in a country lies to the lower left of the distribution, the more that particular country can benefit from spillovers from all of the more advanced economies. Clearly, the spillover will depend on the spread of the distribution at each point in time. If the distribution collapses to just one point at time t, that is, if all the technologies are identical across countries, then, there will be no scope for spillover effects.

To develop this model any further, some distributional assumptions are needed. Assuming that technology across countries is distributed according to a logistic distribution at time t, the technology spillover can be written as follows (see Appendix I.i at the end of this thesis for the derivation of this result):

$$E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right] = \frac{\sqrt{3}S\left[\mathbf{a}_t\right]}{\pi}\left(1 + \phi_{i,t}\right)\ln\left(\frac{1 + \phi_{i,t}}{\phi_{i,t}}\right)$$

where

$$\phi_{i,t} = \exp\left[\frac{\pi}{\sqrt{3}}\frac{a_{i,t} - E\left[\mathbf{a}_t\right]}{S\left[\mathbf{a}_t\right]}\right]$$

and $E[\mathbf{a}_t]$ and $S[\mathbf{a}_t]$ are the unconditional mean and standard deviation of the distribution of technology, $\mathbf{a}_t$, across countries at time t. It is clear from this equation that the size of the spillover depends directly on the mean and the standard deviation of the distribution of technology across countries.

To understand the impact of spillovers in this model, the size of the spillovers across all countries needs to be summarized. To keep the analysis simple, it is further assumed that $g_i = g$ and $\lambda_i = \lambda$ for all countries. The first assumption, that is, that the asymptotic rates of technology growth in all the countries are the same, is implicitly assuming that the countries will exhibit convergence of technology. If we allow the g_i's to differ there will be no convergence of technology. Nevertheless, because of the existence of technology spillovers through time which continuously pulls countries' technologies together, they will not diverge forever. Unequal lambdas for the countries will only make a difference in that (assuming that the innovation rates are the same) it would take longer for the countries with low lambdas to converge because they cannot adopt as much technology in each period.

With these two assumptions in place, the expected value and the variance of the spillover across countries at time t can be calculated (see Appendix I.ii for the derivation of the following results). The expected value of the spillover across countries at time t is given by

$$E\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] = \frac{\lambda\pi}{2\sqrt{3}} S\left[\mathbf{a}_t\right] \tag{2.53}$$

and the variance of the spillover across countries at time t is given by

$$S^2\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] = \lambda^2\left(1 - \frac{\pi^2}{12}\right) S^2\left[\mathbf{a}_t\right]$$

Therefore, both of them depend on $S\left[\mathbf{a}_t\right]$; that is, they depend on the spread of the distribution of technology across countries at time t. Thus, the wider the distribution of $\mathbf{a}_t$ across countries, the bigger the expected value of the spillover. The reason being that since countries technologies differ considerably there is more scope for technology transfers.

2.4.2.2 The dynamics of the cross country distribution of the spillover.

Since the variance of the distribution of technologies determines the expected value and the variance of the spillover, it is important to assess the behaviour of this standard deviation across countries as time goes by. The derivation of the following results can also be found in Appendix I.iii at the end of this thesis. The mean of the distribution of $a_{i,t}$ evolves according to the following process over time

$$E\left[\mathbf{a}_{t+1}\right] = g + E\left[\mathbf{a}_t\right] + \frac{\lambda\pi}{2\sqrt{3}} S\left[\mathbf{a}_t\right]$$

Thus, the mean of the distribution of technology grows at the asymptotic rate of growth of technology for all the countries, g, but again, it also depends on the spread of the distribution of $\mathbf{a}_t$ on the last period. This is obvious since the countries at the bottom of the distribution can benefit from big technology transfers, which will increase greatly their level of technology in the following period and therefore push the mean of the distribution of $\mathbf{a}_t$ upwards.

The variance of the distribution of $\mathbf{a}_t$ evolves over time according to

$$\begin{aligned} S^2\left[\mathbf{a}_{t+1}\right] &= \sigma^2 + S^2\left[\mathbf{a}_t\right]\left[1 + \lambda^2\left(1 - \frac{\pi^2}{12}\right) - \frac{6\lambda}{\pi^2}\Omega\right] \\ &= \sigma^2 + S^2\left[\mathbf{a}_t\right]\mathbf{C} \end{aligned} \tag{2.54}$$

where $\Omega = 1.2$ and $\mathbf{C} = 1 + \lambda^2\left(1 - \frac{\pi^2}{12}\right) - \frac{6\lambda}{\pi^2}\Omega$. This is a first order difference equation on $S^2\left[\mathbf{a}_t\right]$. Since $\mathbf{C}$ is less than one for values $0 < \lambda \leq 1$, the variance of the distribution of technology across countries over time approaches a steady state value which is given by the following expression

$$S^2\left[\mathbf{a}_{ss}\right] = \frac{\sigma^2}{\frac{6\lambda}{\pi^2}\Omega - \lambda^2\left(1 - \frac{\pi^2}{12}\right)} \tag{2.55}$$

This variance is always positive as long as $0 < \lambda \leq 1$ and it is influenced by the size of the shocks to technology (σ^2) and the speed of technology transfers, λ. As σ^2 increases, that is, as the variance of the shocks to technology increases, the long run variance obviously increases and as λ increases, that is, countries are able to absorb more of the potential spillover each period, the long run variance decreases. In particular, if $\lambda = 0$ and there are no spillovers, then the coefficient of $S^2\left[\mathbf{a}_t\right]$ in equation (2.54) is equal to 1, as expected, and the variance of the distribution of the logarithm of technologies increases with time. The size of the increase in each period depends on the size of the shocks to the logarithm of technology. For values $0 < \lambda \leq 1$, the coefficient of $S^2\left[\mathbf{a}_t\right]$ is positive and less than one. As lambda increases, the coefficient decreases and in particular when $\lambda = 1$ this coefficient is equal to 0.44677.

For simplicity of exposition, let us assume away for the moment the shocks to technology; that is, let us concentrate on a deterministic model. If this is the case, the variance of the distribution of technology across countries in the steady state depicted in equation (2.55) would be zero and consequently, the expected value of the spillover once the economy reaches the steady state (see equation (2.53)) would also be zero. This is clear because in the absence of shocks to the logarithm of technology, the existence of spillovers will result in convergence of technologies across countries to the same level and therefore no further potential for spillovers.[39] In the long run, all countries will still grow at the same common rate, g, as in the absence of spillovers, but the countries which benefited from spillovers will achieve a higher level of technology at each point in time that they would have achieved otherwise. Intuitively, this will also result in a higher level of output at each point in time as will be shown later in this section.

Figure 2.7 shows the adjustment paths of the logarithm of technology (left hand column) and the value of the variable $SPILL_{i,t}$ (right hand column) for three different cases. In the same way as in section 2.4.1.2, the parameter values are selected assuming a time period unit equal to a quarter. Also, to be consistent with the data used in chapter 6, a group of seven countries is employed to generate the series. The following values of the parameters are common for the three sets of graphs: the asymptotic growth rates for the seven economies, g_i, are generated from a uniform distribution between 0.002 and 0.005, the $\varepsilon_{i,t}$'s are drawn from a normal distribution with mean zero and standard deviation equal to 0.02 and the initial values of the logarithm of technologies are calculated from the actual data used in chapter 6. The differences between the three cases are the values of λ. In the first case, lambda is set to zero, and therefore there are no spillovers. Since countries are growing at different asymptotic growth rates, the distribution of the logarithm of technology widens by the end of the period; that is, countries technologies are diverging. In the second case lambda is equal to 0.05. Even though countries are growing at different rates the spillover effect compresses the distribution by the end of the 100 time periods. Allowing for a higher value of lambda ($\lambda = 0.3$) in

[39] This is obviously only in the case that the g_i's are the same across countries. If this is not the case, even in the absence of shocks there will be still scope for spillovers because technologies in different countries will be growing at different rates, and therefore will be diverging and then pulled together by the spillovers and then diverging again and so on.

the third case, narrows the distribution quicker and even after 100 periods it is easily seen that countries are not diverging as a result of the spillover.

2.4.2.3 The dynamics of the extended OG model with technology spillovers.

The next step, therefore, will be to incorporate this idea of technology spillovers across countries to the OG model of section 2.2.2.3 For simplicity, in the remaining of this section the indices of the countries will be omitted.

As in section 2.2.2.3, individuals live two periods. They work and save in the first period and in the second, they live off the benefits of their first period savings. New individuals are continuously being born while old individuals are continuously dying. Production is carried out following the same Cobb-Douglas production function used in equation (2.17), which again can be written in terms of effective units as in equation (2.18). Population grows at rate n (see equation(2.19)). However, technology does not only grow at g, as it was the case before, but it can temporarily grow quicker due to technology spillovers from other countries. Therefore equation (2.20) is replaced by the following equation[40]

$$A_t = \left(g + \lambda E\left[\mathbf{a}_{t-1} - a_{t-1} \mid \mathbf{a}_{t-1} > a_{t-1}\right]\right) A_{t-1}$$

Domestic markets are again perfectly competitive and capital depreciates at a constant rate δ which is set equal to 1 as before. Thus, in equilibrium, the same two equations are obtained for the marginal products of labour and capital; see equations (2.21) and (2.22) respectively. The dynamics of the economy can be characterized, following

[40] For simplicity and to be able to compare the results of this model with the results of the closed economy model in Section 4.2, the shocks to technology are omitted.

similar steps to those in section 2.2.2.3, by the following equation

$$k_{t+1} = \frac{\beta(1-\alpha)}{(1+\beta)(1+n)(1+g+\lambda SPILL_t)} k_t^{\alpha} \tag{2.56}$$

Note that there is a difference between the equation corresponding to the closed economy model (2.30) and equation (2.56) and that is the term $\lambda SPILL_t$ in the denominator. As time goes by, this term will tend to zero as countries exhaust all the possibilities from imitation. Therefore, in the steady state this term disappears and, thus, countries converge to the same steady state level of capital per effective worker, k^*, which is constant and given by the same expression as before (see equation (2.31))

$$k^* = \left(\frac{\beta(1-\alpha)}{(1+\beta)(1+n)(1+g)}\right)^{\frac{1}{1-\alpha}}$$

This steady state value is globally stable, that is, no matter where the economy starts (apart from zero) it will always converge to the steady state value. However, the difference is that the level of technology for the countries which benefited from the spillovers is higher than it would have been in the absence of spillovers, and this implies that the level of capital is also higher because $K_t = k_t A_t L_t$. Output is higher as well because both the level of capital and the level of technology are higher.

In conclusion, the presence of technology spillovers across countries does not alter the rate of growth of output in the long run if all the countries are assumed to innovate at the same rate. However, because relatively backward countries can benefit from technologies discovered in the most advanced countries, there would be convergence of the levels of technology. Even if there are random shocks to technology, the levels of technology will not diverge due to the effect of the spillovers. In these circumstances, capital and output will be higher for the countries which benefited from the technology spillovers than it would have been otherwise. In the case that the asymptotic rates of technologies are different across countries, again, the levels of technology across countries will be pulled together by the spillover effects and therefore there will not be divergence.

2.5 Concluding comments.

This chapter has given a brief overview of the literature in growth and convergence and many of the conflicting views in the literature have been put forward. The rest of the thesis builds on this literature and extends it in several different ways. First, chapter 3 looks again at the issue of convergence of output growth across large groups of countries. The approach used in this chapter differs from the main strands of the literature in two ways: it employs the two equation system implied by the Ramsey-Cass-Koopmans model as opposed to the typically analysed Solow growth model and a more rigorous continuous time approach is applied. chapters 5 and 6 are based on the ideas of the two theoretical models illustrated in section 2.4 in this chapter. Chapter 5 analyses the presence of a balance of payments constraint on growth of the type described in section 2.4.1. Chapter 6 considers the effect of technology spillovers across countries making use of the ideas in section 2.4.2. Both chapters concentrate on a smaller group of countries, the G7, and apply nonlinear econometric methods with the aim of studying the process of growth in more detail. The econometrics of the type of nonlinear models used in these chapters are described and reviewed in chapter 4.

Table 2.1. Cross section estimation including human capital. 81 countries.

Dependent variable: ln difference of GDP per head 1960-1989			
Regressor	Coefficient	Standard Error	T-Ratio
Intercept	-0.5431	0.3551	-1.5294
$\ln(y60)$	-0.3411	0.0655	-5.2062
$\ln(s)$	0.4251	0.0767	5.5405
$\ln(n+g+\delta)$	-0.7564	0.0771	-9.8051
$\ln(school)$	0.3313	0.0700	4.7306
R^2	0.5776	F-statistic F(3,77)	35.0985
$\bar{R}^2$	0.5612	S.E. of Regression	0.3483
RSS	9.3421	Mean of Dependent vble.	0.5453
S.D. of Dependent vble.	0.5258	Max. Log-likelihood	-27.4573
DW-statistic	1.6930	LM tests of rests χ^2 (1)	1.8562[0.17]
Diagnostic Tests			
Test Statistics	LM Version		F Version
Serial Correlation	$\chi^2(1) = 1.9478$		F(1,76)=1.8726
Functional Form	$\chi^2(1) = 7.6649$		F(1,76)=7.9435
Normality	$\chi^2(2) = 2.5126$		Not applicable
Heteroskedasticity	$\chi^2(1) = 5.6706$		F(1,79)=5.9469

Table 2.2. Cross section estimation including human capital. 49 countries.

Dependent variable: ln difference of GDP per head 1960-1989			
Regressor	Coefficient	Standard Error	T-Ratio
Intercept	-0.7764	0.5618	-1.3820
$\ln(y60)$	-0.4453	0.0946	-4.7071
$\ln(s)$	0.6359	0.1655	3.8421
$\ln(n+g+\delta)$	-0.9378	0.1562	-6.0041
$\ln(school)$	0.3018	0.1617	1.8669
R^2	0.4520	F-statistic F(3,45)	12.3713
$\bar{R}^2$	0.4154	S.E. of Regression	0.3510
RSS	5.5452	Mean of Dependent vble.	0.6836
S.D. of Dependent vble.	0.4591	Max. Log-likelihood	-16.1451
DW-statistic	2.0934	LM tests of rests χ^2 (1)	0.7893[0.37]
Diagnostic Tests			
Test Statistics	LM Version		F Version
Serial Correlation	$\chi^2(1) = 0.1820$		F(1,44)=0.1640
Functional Form	$\chi^2(1) = 4.5941$		F(1,44)=4.5521
Normality	$\chi^2(2) = 4.5618$		Not applicable
Heteroskedasticity	$\chi^2(1) = 2.3330$		F(1,47)=2.3496

Table 2.3. Cross section estimation including human capital. 18 countries.

Dependent variable: ln difference of GDP per head 1960-1989			
Regressor	Coefficient	Standard Error	T-Ratio
Intercept	-0.8700	1.2391	-0.7021
$\ln(y60)$	-0.4621	0.0772	-5.9866
$\ln(s)$	0.6844	0.1945	3.5190
$\ln(n+g+\delta)$	-0.9829	0.2193	-4.4817
$\ln(school)$	0.2985	0.1435	2.0804
R^2	0.7706	F-statistic F(3,14)	15.6770
$\bar{R}^2$	0.7215	S.E. of Regression	0.1412
RSS	0.2792	Mean of Dependent vble.	0.8381
S.D. of Dependent vble.	0.2676	Max. Log-likelihood	11.9555
DW-statistic	1.3956	LM tests of rests χ^2 (1)	0.0106[0.92]
Diagnostic Tests			
Test Statistics	LM Version		F Version
Serial Correlation	χ^2 (1) = 0.0649		F(1,13)=0.0471
Functional Form	χ^2 (1) = 7.0493		F(1,13)=8.3685
Normality	χ^2 (2) = 0.0103		Not applicable
Heteroskedasticity	χ^2 (1) = 0.3098		F(1,16)=0.2802

Table 2.4. Panel data estimation using Least Squares Dummy Variables. 81 countries.

Dependent variable: $\ln y\,(t_2)$			
Regressor	Coefficient	Standard Error	T-Ratio
Intercept	0.7710	0.1840	4.1899
$\ln y(t_1)$	0.8184	0.0268	30.5593
$\ln(s)$	0.1147	0.0193	5.9294
$\ln(n+g+\delta)$	-0.1147	0.0193	-5.9294
R^2	0.9911	F-statistic F(87,398)	507.3275
$\bar{R}^2$	0.9891	S.E. of Regression	0.1078
RSS	4.6273	Mean of Dependent vble.	7.8026
S.D. of Dependent vble.	1.0332	Max. Log-likelihood	441.3746
DW-statistic	2.0835	LM tests of rests χ^2 (1)	2.2552[0.13]
Diagnostic Tests			
Test Statistics	LM Version		F Version
Serial Correlation	χ^2 (1) = 1.1560		F(1,397)=0.9465
Functional Form	χ^2 (1) = 11.0722		F(1,397)=9.2554
Normality	χ^2 (2) = 24.9594		Not applicable
Heteroskedasticity	χ^2 (1) = 5.8225		F(1,484)=5.8688

Table 2.5. Panel data estimation using Least Squares Dummy Variables. 49 countries.

Dependent variable: $\ln y\,(t_2)$			
Regressor	Coefficient	Standard Error	T-Ratio
Intercept	0.6628	0.2324	2.8525
$\ln y(t_1)$	0.8127	0.0364	22.3563
$\ln(s)$	0.1490	0.0269	5.5416
$\ln(n+g+\delta)$	-0.1490	0.0269	-5.5416
R^2	0.9916	F-statistic F(55,238)	510.1318
$\bar{R}^2$	0.9896	S.E. of Regression	0.0916
RSS	1.9965	Mean of Dependent vble.	8.2199
S.D. of Dependent vble.	0.9000	Max. Log-likelihood	316.6851
DW-statistic	2.0712	LM tests of rests χ^2 (1)	1.0397[0.31]
Diagnostic Tests			
Test Statistics	LM Version		F Version
Serial Correlation	$\chi^2\,(1) = 0.4757$		F(1,237)=0.3841
Functional Form	$\chi^2\,(1) = 4.9174$		F(1,237)=4.0314
Normality	$\chi^2\,(2) = 41.6159$		Not applicable
Heteroskedasticity	$\chi^2\,(1) = 18.0784$		F(1,292)=19.1318

Table 2.6. Panel data estimation using Least Squares Dummy Variables. 18 countries.

Dependent variable: $\ln y\,(t_2)$			
Regressor	Coefficient	Standard Error	T-Ratio
Intercept	2.6071	0.5455	4.7794
$\ln y(t_1)$	0.6876	0.0489	14.0567
$\ln(s)$	0.0586	0.0571	1.0263
$\ln(n+g+\delta)$	-0.0586	0.0571	-1.0263
R^2	0.9924	F-statistic F(24,83)	452.4608
$\bar{R}^2$	0.9902	S.E. of Regression	0.0462
RSS	0.1775	Mean of Dependent vble.	9.1091
S.D. of Dependent vble.	0.4676	Max. Log-likelihood	192.9588
DW-statistic	2.1918	LM tests of rests χ^2 (1)	1.7576[0.18]
Diagnostic Tests			
Test Statistics	LM Version		F Version
Serial Correlation	$\chi^2\,(1) = 1.8033$		F(1,82)=1.3924
Functional Form	$\chi^2\,(1) = 4.0434$		F(1,82)=3.1894
Normality	$\chi^2\,(2) = 1.0671$		Not applicable
Heteroskedasticity	$\chi^2\,(1) = 7.4093$		F(1,106)=7.8077

Table 2.7. Descriptive Statistics of the Estimated Coefficients from the Solow Growth model. Unrestricted regressions.

Non-oil countries (n=81)				
	constant	**trend**	$\ln(y_{t-1})$	**Implied** g
Mean	1.47970	0.00278	0.80184	0.00847
Std. Error	1.14366	0.00389	0.14858	0.00489
Std. Error of Mean	0.08992	0.00062	0.01306	0.00337
Median	1.35324	0.00212	0.81910	0.01288
Std. Deviation	0.80427	0.00556	0.11679	0.03016
Kurtosis	-0.15035	5.61788	0.47467	5.04333
Skewness	0.67508	1.77102	-0.83249	-1.32506
Minimum	0.14640	-0.00575	0.43513	-0.12641
Maximum	3.64278	0.02825	0.98458	0.07121
Intermediate group (n=49)				
	constant	**trend**	$\ln(y_{t-1})$	**Implied** g
Mean	1.39077	0.00343	0.82469	0.01329
Std. Error	1.05901	0.00359	0.13172	0.00435
Std. Error of Mean	0.11818	0.00079	0.01585	0.00370
Median	1.16382	0.00256	0.84205	0.01650
Std. Deviation	0.81875	0.00550	0.10982	0.02566
Kurtosis	0.22367	9.68982	0.95941	4.44722
Skewness	0.85662	2.59061	-1.04195	-0.95381
Minimum	0.14640	-0.00575	0.51959	-0.08584
Maximum	3.64278	0.02825	0.98458	0.07121
OECD countries (n=18)				
	constant	**trend**	$\ln(y_{t-1})$	**Implied** g
Mean	1.46340	0.00297	0.83821	0.01808
Std. Error	0.88440	0.00267	0.10103	0.00303
Std. Error of Mean	0.17933	0.00040	0.01963	0.00143
Median	1.47871	0.00311	0.83784	0.01772
Std. Deviation	0.73942	0.00164	0.08093	0.00588
Kurtosis	-1.28204	-1.00776	-1.20776	0.13072
Skewness	0.29345	0.09674	-0.25495	-0.30012
Minimum	0.54335	0.00031	0.70356	0.00511
Maximum	2.67459	0.00580	0.94153	0.02778

Table 2.8. Descriptive Statistics of the Estimated Coefficients from the Solow Growth model. Lambda and g restricted.

Non-oil countries (n=81)				
	constant	**trend**	$\ln(y_{t-1})$	**Implied** g
Mean	0.33273	-0.00022	0.96000	-0.00544
Std. Error	0.05045	0.00020	0.00659	0.00020
Std. Error of Mean	0.00564			
Median	0.33183			
Std. Deviation	0.05044			
Kurtosis	-1.24396			
Skewness	-0.03790			
Minimum	0.24384			
Maximum	0.41348			
Intermediate group (n=49)				
	constant	**trend**	$\ln(y_{t-1})$	**Implied** g
Mean	0.25969	-0.00027	0.97150	-0.00936
Std. Error	0.06145	0.00024	0.00782	0.00024
Std. Error of Mean	0.00463			
Median	0.25239			
Std. Deviation	0.03205			
Kurtosis	-1.37443			
Skewness	0.04183			
Minimum	0.20489			
Maximum	0.31463			
OECD countries (n=18)				
	constant	**trend**	$\ln(y_{t-1})$	**Implied** g
Mean	0.74831	0.00130	0.91822	0.01586
Std. Error	0.09672	0.00033	0.01122	0.00034
Std. Error of Mean	0.00793			
Median	0.75830			
Std. Deviation	0.03269			
Kurtosis	6.15166			
Skewness	-2.36002			
Minimum	0.64512			
Maximum	0.78012			

Figure 2.1. Histogram of per capita income growth rates for the period 1960-1989.

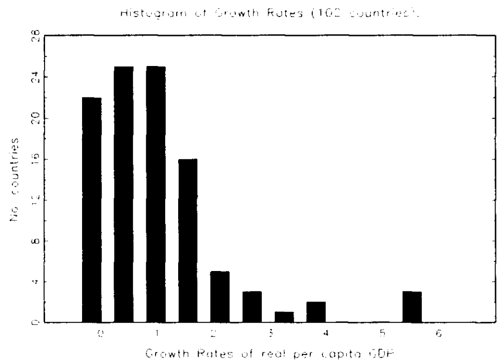


Figure 2.2. Histogram of per capita GDP in 1960.

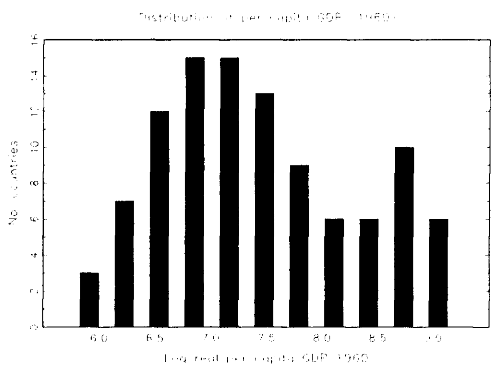


Figure 2.3. Histogram of per capita GDP in 1989.

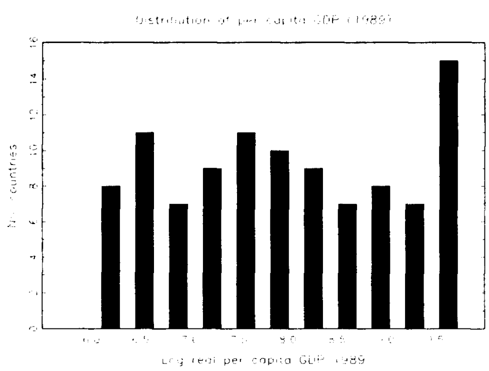


Figure 2.4. Scatter plot of growth of GDP per capita on initial GDP.

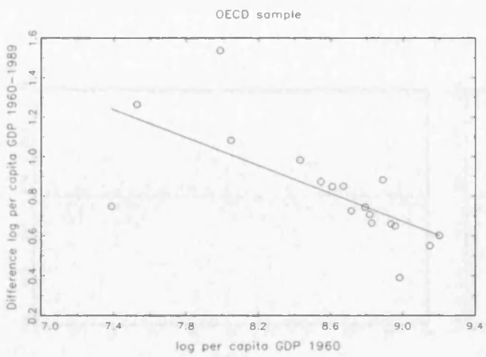


Figure 2.5. Bias of the estimated λ in the group of OECD countries.

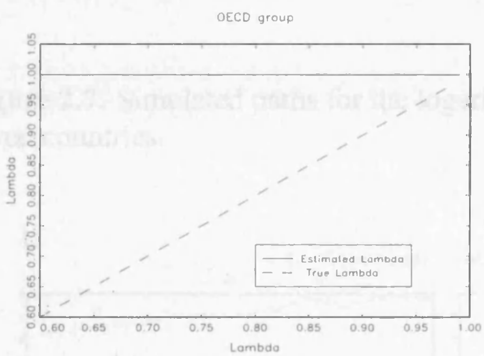


Figure 2.6. Simulated paths of capital per effective units in the deterministic and stochastic open economy OG model with international capital market imperfections.

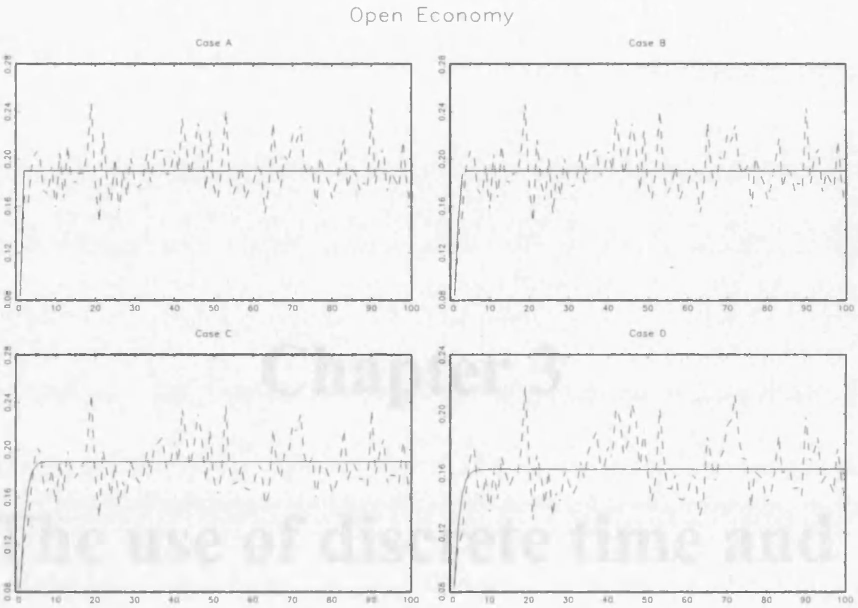
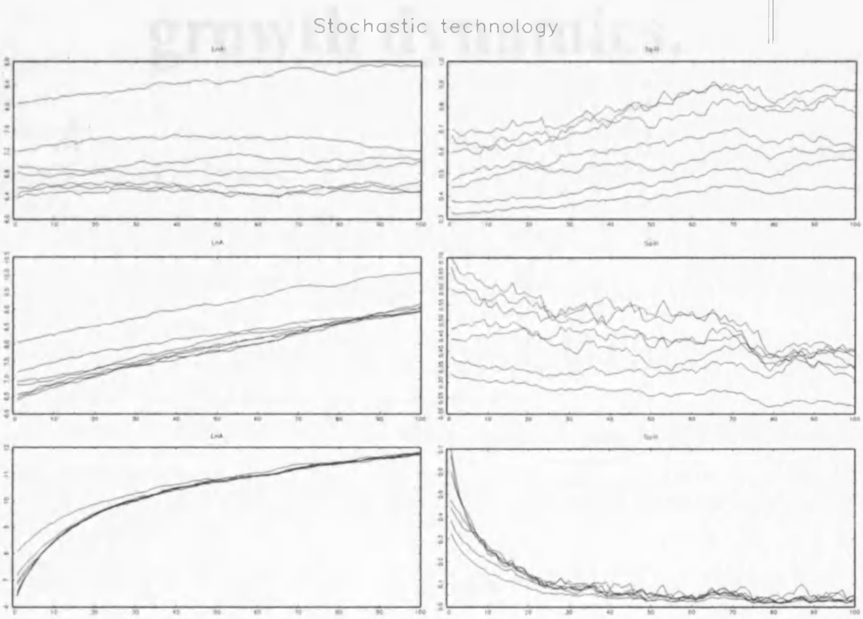


Figure 2.7. Simulated paths for the logarithm of technology and the spillover for seven countries.



Chapter 3

The use of discrete time and continuous time modelling in growth dynamics.

3.1 Introduction.

As outlined in chapter 2, there have been a wealth of empirical studies concerned with cross country growth and convergence in recent years. The starting point for most of these studies is a discrete time version of the Solow growth model. In this chapter, this empirical literature is reassessed by considering two developments of the typical applied exercise. The first considers the use of a continuous time approach rather than the discrete time approach usually found in the literature. The second considers the two equation system developed by Ramsey-Cass-Koopmans (as described in chapter 2) in which both output growth and the savings rate are determined endogenously. This model encompasses the Solow growth model and represents a well known generalisation of Solow which, like the continuous time extension, most authors appear to consider relatively innocuous.

In section 3.2, the differences raised in analysing the discrete time and continuous time versions of the Solow model are elaborated on. The Solow model can be derived either in discrete time (see, for example, Lee, Pesaran and Smith (1995) or Obstfeld and Rogoff (1996)) or in continuous time (for example, Barro and Sala-i-Martin (1992) or Romer (1996)). However, whilst the theoretical predictions of the model are the same regardless of whether it is derived using discrete or continuous time, the exact discretization of the continuous model gives rise to an econometric model which differs from that derived using discrete time. In this section, it will be shown that deriving the Solow growth model in discrete time, as in Lee, *et al* (1995), is equivalent to the

Euler approximation of the continuous time Solow growth model. Since the Euler approximation is known to be an invalid discretization of any continuous time model, the estimates obtained from this model are inconsistent (see Gourieroux and Monfort (1996)).

In section 3.3, the Ramsey-Cass-Koopmans (RCK) model of growth is briefly introduced. This model encompasses the Solow growth model but it produces more sophisticated dynamics. In this section, in contrast to Mankiw *et al* (1992) and Lee *et al* (1995) who wrote of "taking Robert Solow seriously", the Ramsey-Cass-Koopmans model is taken "seriously". It examines whether the more sophisticated transitional dynamics of this model, relative to those of the Solow model, provide additional insights. The exact discrete model corresponding to the continuous time model is, therefore, derived together with the Euler approximation of the model.

In section 3.4, both models will be estimated and compared for three different samples of countries. In this empirical exercise, both extensions of the standard analysis (namely, continuous time econometric analysis in place of discrete time econometric analysis and the inclusion of an endogenous savings rate *a la* RCK) are found to have a significant impact on modelling growth and convergence. Finally, section 3.5 concludes.

3.2 The Solow growth model in discrete and continuous time.

The Solow growth model is the basic reference point for many growth analyses. This model has been derived in continuous time (see for example Romer (1996)) but it has also been derived using discrete time (see for example Obstfeld and Rogoff (1996)).

In chapter 2, the discrete time version of the Solow growth model was presented; here, the continuous time derivation of the model will be outlined. Once both versions of the Solow growth model have been derived, it will be shown that the discrete time derivation of the Solow growth model corresponds to the Euler approximation of the continuous time Solow growth model. This approximation obviously differs from the exact discrete model derived from the continuous time Solow model. From a theoretical point of view both versions are equivalent in the sense that they share the same predictions about growth and convergence. However, this is not the case from the estimation point of view. Gourieroux and Monfort (1996) showed that since the Euler approximation of a continuous time model is not the right discretization of the model, the parameter estimates found from this approximation are inconsistent. It is shown below that this means that inferences made on the speed of convergence of countries to the long run path based on the approximation may be invalid and misleading.

3.2.1 The derivation of the model.

The derivation here follows that of Romer (1996). This derivation is analogous to the discrete time derivation of chapter 2, but, instead of using discrete variables, their continuous counterparts are used. For example, ignoring for the time being country subscripts, output, Y_t, becomes now the continuous variable $Y(t)$. Deriving the Solow model in the same manner as the discrete time, the following nonlinear differential equation is obtained

$$\frac{dk(t)}{dt} = sk(t)^{\alpha} - (n + g + \delta) k(t)$$

where d is the differential operator. Using Taylor's approximation around the steady state, this differential equation can be written in terms of the logarithm of per capita output as follows (see Appendix II.i. for a derivation of this result):

$$\begin{aligned}\frac{d\ln\tilde{y}(t)}{dt} &= g+(1-\alpha)(n+g+\delta)\ln A_0+ \\ &+\alpha(n+g+\delta)(\ln s-\ln(n+g+\delta))+ \\ &+(1-\alpha)(n+g+\delta)gt-(1-\alpha)(n+g+\delta)\ln\tilde{y}(t)\end{aligned}$$

where $\tilde{y}(t)$ stands for per capita output at time t, and the parameters n, g, s, α, δ and A_0 are the same as those in section 2.2.2.1; n and g represent the population and technology growth rates respectively, s is the savings rate, α is the parameter of the Cobb-Douglas production function, δ is the rate of depreciation and A_0 is the initial level of technology. This linear differential equation can be written in terms of the parameters λ and μ defined in chapter 2 (see equations (2.5) and (2.6)) as:

$$d(\ln\tilde{y}(t))=[\mu+(1-\lambda)g+(1-\lambda)gt-(1-\lambda)\ln\tilde{y}(t)]\,dt \tag{3.1}$$

where

$$\begin{aligned}\lambda &= 1-(1-\alpha)(n+g+\delta) \\ \mu &= \lambda g+(1-\lambda)\left[\ln(A_0)-\frac{\alpha}{1-\alpha}\ln(n+g+\delta)+\frac{\alpha}{1-\alpha}\ln(s)\right]\end{aligned}$$

For estimation purposes, similar to the discrete Solow model, this equation can be augmented with the equivalent of a disturbance in continuous time, $dW(t)$, where

$W(t)$ is a standard Brownian motion; hence, the process $dW(t)$ can be written in differential equation form as $dW(t) = \epsilon\sqrt{dt}$ where $\epsilon\tilde{}N(0,1)$. Additionally, the values of $dW(t)$ for two different time intervals are independent; that is, the ϵ's are also independent. Allowing for cross country differences, different tastes and different technologies can also be assumed. With these extensions, equation (3.1) becomes the following linear stochastic differential equation

$$\begin{aligned} d(\ln \tilde{y}_i(t)) &= [\mu_i + (1-\lambda_i) g_i + (1-\lambda_i) g_i t \\ &\quad - (1-\lambda_i) \ln \tilde{y}_i(t)] dt + \sigma_i dW_i(t) \end{aligned} \tag{3.2}$$

where the i subscript again denotes country i. In general, it is not possible to find the explicit exact discrete model corresponding to a differential equation. However, in this case, the exact discretized version of equation (3.2) is as follows (see Appendix II.ii. for the derivation of this result)

$$\begin{aligned} \ln \tilde{y}_{it} &= \{1 - \exp[-(1-\lambda_i)]\} \left(\frac{\mu_i - \lambda_i g_i}{1-\lambda_i}\right) + g_i \exp[-(1-\lambda_i)] + \\ &\quad + g_i \{1 - \exp[-(1-\lambda_i)]\} t + \exp[-(1-\lambda_i)] \ln \tilde{y}_{it-1} + \\ &\quad + \frac{\sigma_i}{[2(1-\lambda_i)]^{\frac{1}{2}}} \{1 - \exp[-2(1-\lambda_i)]\}^{\frac{1}{2}} \epsilon_{it} \end{aligned} \tag{3.3}$$

The Euler approximation of equation (3.2) can also be introduced by simply approximating the derivative of the two continuous variables by their discrete change at time t and by substituting each continuous variable on the right hand side by its discrete counterpart at time $t-1$. The Euler approximation of equation (3.2) is therefore

$$\ln \tilde{y}_{it} = \mu_i + (1-\lambda_i) g_i t + \lambda_i \ln \tilde{y}_{i,t-1} + \sigma_i \epsilon_{it} \tag{3.4}$$

This equation is equivalent to equation (2.7), that is, the equation obtained by deriving the Solow growth model in discrete time where $\varepsilon_{it} = \sigma_i \epsilon_{it}$. In principle, both equations (3.3) and (3.4) could be estimated by using maximum likelihood, for example. In general, it is not possible to determine the analytical form of the likelihood function. In these cases, indirect inference is required to estimate the continuous time model. However, in this case, indirect inference is not required since it is easy to find the analytical form of the likelihood function. However, the estimated parameters obtained from the model which uses the Euler approximation (i.e. equation (3.4)) are inconsistent. This is because the discretization of equation (3.2) is not the right one.

3.2.2 The extent and consequences of the discrete time bias.

The direction of the bias is very easy to see in this particular model since both equations (3.3) and (3.4) have the same structure. The right hand side of both equations consists of a constant, a time trend, the first lag of the dependent variable and an error term. However, these variables are multiplied by different parameters in each model. Here, particular attention will be paid to the estimation of the two parameters which are of most interest: λ_i and g_i. The coefficient λ_i determines the speed of convergence to the steady state for country i, given by $1 - \lambda_i$ and has been the focus of considerable interest in the convergence literature. The parameter g_i is important since this parameter gives the long run growth rate of per capita output for country i which, after all, dominates the convergence issue ultimately.

First, attention is restricted to the parameter λ_i. Using equation (3.4), that is, the approximation to the model, the coefficient of the lagged logarithm of per capita output would be taken as an estimate of λ_i. However, it is evident from the exact model in equation (3.3), that the estimated coefficient is not an estimate of λ_i, but an estimate of $\exp\left[-\left(1-\lambda_i\right)\right]$. The theoretical values of λ_i lie in the interval $(0, 1)$. Thus, the estimate of λ_i obtained from the approximated model is biased upwards, since $\exp\left[-\left(1-\lambda_i\right)\right] > \lambda_i$ when $0 < \lambda_i < 1$. This is shown more effectively in Figure 3.1. The straight line represents the true values of λ_i whereas the curve represents $\exp\left[-\left(1-\lambda_i\right)\right]$, and therefore, gives the values that would be wrongly interpreted as λ_i based on the approximation (3.4) for each value of λ_i. In this figure, the bias is the difference between the two lines and it is clear that the bias depends on the value of λ_i; the lower λ_i, the bigger the bias. Another important point that should be made here, which is also a consequence of this bias, is related to the spread of the distribution of the λ_is across countries. The values taken by the function $\exp\left[-\left(1-\lambda_i\right)\right]$ when $0 < \lambda_i < 1$ lie in the interval $\left(\exp\left(-1\right), 1\right)$. Therefore, these values lie in a smaller interval than the actual values of λ_i. As a consequence, if the approximated model is (wrongly) used to obtain estimates of λ_i for a group of countries, it could be wrongly inferred that the λ_i's of these countries are closer together than they are in fact. This could lead to failure to reject the null hypothesis of equality of λ_i's across countries even if speeds of convergence do differ across countries.

The other parameters that are of particular interest are the g_i's, the long run growth of per capita output across countries. An estimate of g_i is easily found in the approximated model (see equation (3.4)) by dividing the coefficient of the time trend by one minus the coefficient of the lagged logarithm of per capita output,

$$\frac{\left(1-\lambda_i\right) g_i}{1-\lambda_i} = g_i$$

In the exact discrete model of equation (3.3), the estimate of g_i would be found in the same manner, by dividing the coefficient of the time trend by one minus the coefficient of the lagged logarithm of per capita output,

$$\frac{[1-\exp\left[-\left(1-\lambda_i\right)\right]]\,g_i}{[1-\exp\left[-\left(1-\lambda_i\right)\right]]}=g_i$$

The important point made here, is that even though the two coefficients used to find g_i in the approximated model are biased, the biases are such that they cancel out. Therefore, the same estimate of g_i should be expected from both the approximated and the true model.

3.3 The Ramsey-Cass-Koopmans growth model in discrete and continuous time.

The Ramsey-Cass-Koopmans (RCK) model was derived in chapter 2. This section starts from the two nonlinear differential equations (2.13) and (2.14) in chapter 2 and derives both the exact discrete model corresponding to the continuous time model and the Euler discretization of the model. Again dropping the country specific subscripts for the time being for clarity of exposition, equations (2.13) and (2.14) can be written as

$$dy = \alpha y\left[y^{\frac{\alpha-1}{\alpha}}s-\Gamma\right]dt \tag{3.5}$$

$$ds = (1-s)\left[\alpha y^{\frac{\alpha-1}{\alpha}}\left(s-\frac{1}{\theta}\right)-\alpha\Gamma+\frac{\Phi}{\theta}\right]dt \tag{3.6}$$

where

$$\Gamma = n + g + \delta$$
$$\Phi = \rho + \theta g + \delta$$

The steady state values of y and s are:

$$y^* = \left(\frac{\alpha}{\Phi}\right)^{\frac{\alpha}{1-\alpha}} \tag{3.7}$$

$$s^* = \frac{\alpha\Gamma}{\Phi} \tag{3.8}$$

In this model, y is predetermined and s is a jump variable and consequently, the economy exhibits saddle path stability. Once the economy has reached the steady state, its behaviour is identical to that of the Solow model on the balanced growth path. Capital, output and consumption per unit of effective labour are constant and the saving rate is also constant. All per capita variables grow at the rate of growth of technology, g, and variables in levels grow at the sum of technology and population rates of growth, $n+g$.

However, this model implies a higher speed of convergence to the steady state than the Solow model because the savings rate is allowed to change over time. The speed of convergence in Ramsey's model is given by the negative eigenvalue of the system of differential equations (3.5) and (3.6)

$$\Psi = \frac{\Phi - \Gamma - \sqrt{(\Phi - \Gamma)^2 - \frac{4(1-\alpha)(\alpha\Gamma - \Phi)\Phi}{\alpha\theta}}}{2} \tag{3.9}$$

and depends not only on n, g, δ and α as in the Solow model, but also on the household's willingness to shift consumption between periods (θ) and the rate at which households discount future utility (ρ). Therefore, the same value of λ in both models implies a higher speed of convergence to the steady state in the RCK model than in the Solow model.

Output per worker is given by $\tilde{y}(t) = A(t)\, y(t)$. Taking natural logarithms, differentiating and using a first order Taylor's approximation around the steady state of equations (3.5) and (3.6), the following linear system of differential equations in matrix form is obtained (see Appendix II.iii. for the derivation of this result):

$$dX = [PX + Q + Rt]\, dt \tag{3.10}$$

where $X' = \begin{pmatrix} \ln \tilde{y} & \ln s \end{pmatrix}$ and P, Q and R are (2×2), (2×1) and (2×1) matrices of coefficients respectively with individual elements, p_{ij}, q_{ij} and r_{ij} given by

$$\begin{aligned}
p_{11} &= -(1-\alpha)\,\Gamma \\
p_{12} &= \alpha\Gamma \\
p_{21} &= -\frac{(1-\alpha)(\Phi-\alpha\Gamma)(\alpha\theta\Gamma-\Phi)}{\alpha^2\theta\Gamma} \\
p_{22} &= \Phi-\alpha\Gamma \\
q_{11} &= g+\Gamma\left[(1-\alpha)\ln A(0)-\alpha\ln\Gamma\right] \\
q_{21} &= -(\Phi-\alpha\Gamma)\left[\frac{\Phi}{\alpha\theta\Gamma}\ln\left(\frac{\alpha}{\Phi}\right)+\ln\Gamma-\frac{(1-\alpha)}{\alpha}\left(1-\frac{\Phi}{\alpha\theta\Gamma}\right)\ln A(0)\right] \\
r_{11} &= (1-\alpha)\,\Gamma g \\
r_{21} &= \frac{(1-\alpha)(\Phi-\alpha\Gamma)(\alpha\theta\Gamma-\Phi)}{\alpha^2\theta\Gamma}g
\end{aligned}$$

This system can be augmented for estimation purposes with two independent Brownian motions, $W' = \begin{pmatrix} W_1 & W_2 \end{pmatrix}$, to obtain the following linear system of stochastic differential equations

$$dX = [PX + Q + Rt]\, dt + VdW \tag{3.11}$$

where V is the following lower triangular matrix

$$V = \begin{pmatrix} \sigma_1 & 0 \\ \varrho & \sigma_2 \end{pmatrix}$$

This matrix takes into account the expected correlation between the shocks to the logarithm of output and the logarithm of the savings rate. In the derivation of the model, it was stated that the system exhibits saddle path stability since it is assumed that output is a predetermined variable and the savings rate is a jump variable. Therefore, if a shock to output pushes the system off the saddle path, the savings rate will jump so as to return the system to the saddle path.

3.3.1 The Euler Approximation and the exact discretisation of the system.

The direct Euler approximation of the system of stochastic differential equations in (3.11) can be introduced now:

$$X_t = (Q - R) + Rt + (P + I)\, X_{t-1} + V\epsilon_t \tag{3.12}$$

where $X_t' = \begin{pmatrix} \ln \tilde{y}_t & \ln s_t \end{pmatrix}$, I is the identity matrix and $\epsilon_t' = \begin{pmatrix} \epsilon_{1t} & \epsilon_{2t} \end{pmatrix}$. Note that the Solow model is nested within the Ramsey-Cass-Koopmans model. If it is assumed that the savings rate is constant as in the Solow model, the system collapses to the equation in output; that is, the first row of the system in equation (3.12). Additionally, since the savings rate is now constant, the term $p_{12} \ln s_{t-1}$ would be included in the intercept leaving the same output equation as in the Solow model (see equation (3.4)).

The exact discretization of the system in equation (3.11) can also be found which turns out to be the following autoregressive system of equations (see Appendix II.iv.)

$$\begin{aligned} X_t &= \exp(P) X_{t-1} + [\exp(P) - I] P^{-1} \left(Q + P^{-1} R + Rt\right) \\ &\quad - \exp(P) P^{-1} R + \xi_t \end{aligned} \tag{3.13}$$

where

$$\exp(P) = \sum_{j=0}^{\infty} \frac{1}{j!} P^j \tag{3.14}$$

and

$$\xi_t = \sum_{l=1}^{2} \left[\int_{t-1}^{t} \exp(P(t-s)) V^l dW_l(s) \right]$$

where V^l stands for the lth column of matrix V. Thus, the covariance matrix of ξ_t has the following structure (see Appendix II.iv.)

$$E[\xi_t \xi_t'] = \int_0^1 \exp(Ps) VV' \exp(P's) ds \tag{3.15}$$

Note that the invertibility of P derives from the assumption on absence of infinite lifetime utilities, $\rho - n - (1-\theta) g > 0$. The determinant of P is equal to

$$|P| = \frac{(1-\alpha)\left[\alpha(n+g+\delta) - (\rho + \theta g + \delta)\right](\rho + \theta g + \delta)}{\alpha\theta}$$

It is known that $0 < \alpha < 1$, $\theta > 0$, $(n + g + \delta) > 0$ and $(\rho + \theta g + \delta) > 0$. Therefore, $|P| = 0$ if and only if $\alpha\,(n + g + \delta) = \rho + \theta g + \delta$. In addition, the assumption of no infinite lifetime utilities can be written as $\rho + \theta g + \delta > n + g + \delta$. Since $0 < \alpha < 1$, it follows that $n + g + \delta > \alpha\,(n + g + \delta)$. Consequently, $\rho + \theta g + \delta > \alpha\,(n + g + \delta)$ and the determinant of P is different from zero. Therefore, the matrix P is invertible.

3.3.2 The extent of the discrete time bias.

The aim of this section is to find an analytical expression for the bias of the two parameters that are of most interest in this type of growth studies, namely, λ and g. It is important to reiterate here that $1 - \lambda$ in this case is not the speed of convergence to the equilibrium unlike in the Solow growth model. This was discussed earlier in section 3.3. The actual speed of convergence in the RCK model is given by the negative eigenvalue of the system, given in equation (3.9) and it is higher than $1 - \lambda$. However, since the parameter λ in both models refers to the same underlying parameters of both models, it will be interesting to assess how the estimates of this parameter differs in the two models. Even in the case that the actual estimates in both models are the same, the speed of convergence to the steady state implied by the RCK model will be higher. If the estimates of λ are different, this will imply additional differences in the speed of convergence to the equilibrium.

Finding the direction of the bias in the estimates of λ and g in the RCK model when the approximation is wrongly used, is not as straightforward as in the Solow model. First, attention is restricted to g. To find an estimate of g when the approximated model in equation (3.12) is wrongly used, the coefficient of the time trend in the first equation is divided by the coefficient of the lagged logarithm of per capita output also in the first equation. By looking at the exact discrete model in equation (3.13), it is immediately obvious that in fact the first element of the column vector $[\exp(P) - I]\, P^{-1}R$ is being divided by 1 minus the element in row 1 and column 1 of the matrix $\exp(P)$, namely, $[\exp(P)]_{11}$. The elements of the matrix R can be easily written in terms of g and the elements of the matrix P as $r_{11} = -p_{11}g$ and $r_{21} = -p_{21}g$. In addition, the inverse of the matrix P, matrix P^{-1}, is the following

$$P^{-1} = \frac{1}{|P|}\begin{pmatrix} p_{22} & -p_{12} \\ -p_{21} & p_{11} \end{pmatrix}$$

where $|P| = p_{11}p_{22} - p_{12}p_{21}$. Therefore, the matrix $P^{-1}R$ is quite simply obtained as follows

$$P^{-1}R = \frac{1}{|P|}\begin{pmatrix} p_{22} & -p_{12} \\ -p_{21} & p_{11} \end{pmatrix}\begin{pmatrix} -p_{11}g \\ -p_{21}g \end{pmatrix} = \begin{pmatrix} -g \\ 0 \end{pmatrix} \tag{3.16}$$

Therefore, the element of interest comes from the multiplication of the first row of the matrix $[\exp(P) - I]$ times the column vector in equation (3.16). The coefficient of the time trend in the first equation can, then, be written as $[1 - [\exp(P)]_{11}]\, g$. Now, dividing this coefficient by 1 minus $[\exp(P)]_{11}$, g is obtained. The conclusion for the estimate of this parameter is the same as it was in the Solow model. Even though both estimated coefficients are biased when the approximated model is used, the biases cancel out when g is estimated.

In the case of λ, however, the estimates obtained from the approximated model are biased. Nevertheless, it is not easy to establish the direction of the bias because it depends not only on the value of λ (as it was the case in the Solow model) but also on the value of several other parameters of the model. From the approximated model, the estimate of λ would be taken as the coefficient of the lagged logarithm of per capita output in the first equation. However, this coefficient does not provide an estimate of λ, but an estimate of the element $[1,1]$ of the matrix $\exp(P)$. This matrix can be written as in equation (3.14). As a result, element $[1,1]$ of this matrix can be written as follows

$$\begin{aligned}
[\exp(P)]_{11} &= 1-(1-\alpha)\Gamma+\frac{1}{2!}\left[P^2\right]_{11}+\frac{1}{3!}\left[P^3\right]_{11}+\frac{1}{4!}\left[P^4\right]_{11}+\ldots \\
&= \lambda+\frac{1}{2!}\left[P^2\right]_{11}+\frac{1}{3!}\left[P^3\right]_{11}+\frac{1}{4!}\left[P^4\right]_{11}+\ldots
\end{aligned}$$

Thus, in fact, this element is not just an estimate of λ but an estimate of λ plus an infinite sequence of additional terms. Consequently, the estimate of λ that is obtained from the approximated RCK model will be higher or lower than λ depending on the sign of the remainder. The remainder, however, is a complicated function of the underlying parameters of the RCK model and, as a result, it is difficult to assess the direction of the bias. However, the biggest contribution to the bias comes from the first term, namely, $\frac{1}{2!}\left[P^2\right]_{11}$. The structure of this element in terms of the underlying parameters of the model can be easily derived

$$\frac{1}{2!}\left[P^2\right]_{11}=(1-\alpha)^2\Gamma^2-\frac{(1-\alpha)(\Phi-\alpha\Gamma)(\alpha\theta\Gamma-\Phi)}{\alpha\theta}$$

The first term is always positive, but the second term can be positive or negative depending on the value of θ. From the model assumptions, it is known that $(1-\alpha)>0$, $(\Phi-\alpha\Gamma)>0$ and $\alpha\theta>0$. However, the sign of $(\alpha\theta\Gamma-\Phi)$ is unclear. Two different cases can be distinguished:

Case 1 : $\theta \leq 1$. In this case, $\alpha\theta\Gamma \leq \alpha\Gamma$, and it is known that $\Phi > \alpha\Gamma$. As a result $\Phi > \alpha\theta\Gamma$ and $\frac{1}{2!}\left[P^2\right]_{11} > 0$. Thus, in this case, this first element of the remainder will result in an upward bias.

Case 2 : $\theta > 1$. In this case, $\alpha\theta\Gamma > \alpha\Gamma$. On the other hand, it is known that $\Phi > \alpha\Gamma$. However, this does not clarify the relation between $\alpha\theta\Gamma$ and Φ and therefore, nothing definite can be said about it. However, if θ is small, it is very likely that $\Phi > \alpha\theta\Gamma$, resulting in an upwards bias. If θ is very big, it is very likely that $\Phi < \alpha\theta\Gamma$. However, since θ also appears in the denominator, the second term, although negative is likely to be small relative to the first term. Therefore, it is likely that $(1-\alpha)^2\,\Gamma^2$ will dominate, resulting in an upward bias.

Obviously, this is just an indication on how the bias is likely to manifest itself, since in the derivation of this result attention was restricted to the first term of the remainder. However, since the terms are divided by factorials of increasing numbers, this term is likely to have the biggest contribution towards the bias.

The following section investigates whether these two model developments are as innocuous as is implicitly assumed in the empirical work on growth and convergence. To do this, the next section presents results for the discrete version of the Solow model and its continuous time counterpart using the widely employed Penn World Tables dataset. It then considers also the RCK development, estimating both systems of equations, the Euler approximation to the continuous time model and the exact discrete model.

3.4 Comparison of the estimated discrete time and continuous time models.

This section concentrates on the empirical analysis of both the Solow growth model and the Ramsey-Cass-Koopmans model. All the models in this section are estimated by means of full information maximum likelihood. section 3.4.1 focuses on the Solow growth model. In this section, both the Euler approximation of the Solow model in equation (3.4) and the exact Solow growth model as depicted in equation (3.3) are estimated. The Euler approximation of the Solow growth model has already been estimated in the literature either using different methods of estimation to the one used here or using slightly different groups of countries.[41] However, this estimation exercise is carried out so that the parameter estimates of this model can be compared with those obtained from the exact Solow growth model. Additionally, the extent of the bias in the parameter estimates will be evaluated. It will be shown how the estimates arising from the approximated version of the Solow model are systematically different to those obtained from the exact discrete version. The discussion will deliberately focus on the parameters that determine the speed of convergence to the equilibrium and the long run growth rate.

In section 3.4.2 the corresponding Euler approximation and exact discrete RCK models will be estimated, that is, equations (3.12) and (3.13) respectively. The Solow growth model can be considered as a special case of the RCK model. This model has more sophisticated dynamics than the Solow growth model, and, therefore, it allows a more detailed investigation of the empirics of growth. The parameter estimates of both

[41]This model was also estimated in chapter 2 using OLS to illustrate some of the more important points made in the literature about growth and convergence. However, it is worth reestimating this model to ensure that differences in the estimated parameters are not due to differences in the estimation methods.

versions of the model will be compared paying special attention to the biases arising from the approximation. In addition, the results will also be compared to those obtained from the Solow growth model.

The data used in this empirical exercise are from the Penn World Tables (PWT) version 5.6 (see chapter 2, section 2.3 for a detailed description). The samples are those of chapter 2. A sample of 81 countries and two different subsamples of it are selected.

1. a set of 81 non-oil producing countries
2. an intermediate group of 49 countries which excludes those countries whose population in 1960 was less than one million and countries which are thought to have poor data estimates.
3. a group of 18 OECD countries with populations over 1 million.

3.4.1 Comparison of alternative parameterisations of the Solow growth model.

In this section, the two versions of the Solow model derived in section 3.2 are estimated and compared. First, the parameters of the Solow growth model obtained by using the Euler approximation (see equation (3.4)) are estimated, which is exactly the same model obtained by deriving the Solow growth model in discrete time. For clarity of exposition, this equation is reproduced below

$$\ln \tilde{y}_t = \alpha_1 + \alpha_2 t + \alpha_3 \ln \tilde{y}_{t-1} + \sigma \epsilon_t$$

where $\alpha_1 = \mu$, $\alpha_2 = (1 - \lambda) g$ and $\alpha_3 = \lambda$. This equation is estimated by full information maximum likelihood, separately for each of the countries in the study. Tables 3.1, 3.2 and 3.3 present some descriptive statistics of the distribution of the parameter estimates for the three samples highlighted earlier. The estimates of the growth rate, $\hat{g}$, are obtained from

$$\hat{g} = \frac{\hat{\alpha}_2}{1 - \hat{\alpha}_3}$$

Their corresponding standard errors are computed with the usual first order Taylor's approximation of $\hat{g}$ around the true parameter vector and using the sample estimates in place of the unknown parameters. Therefore, the standard error of $\hat{g}$ is computed according to the following formulae

$$s.e.\left(\hat{g}\right) = \sqrt{D\left(\hat{g}\right)' V\left(\hat{\alpha}_2, \hat{\alpha}_3\right) D\left(\hat{g}\right)}$$

where $V\left(\hat{\alpha}_2, \hat{\alpha}_3\right)$ is the $[2 \times 2]$ variance covariance matrix of the two parameters $\hat{\alpha}_2$ and $\hat{\alpha}_3$, and $D\left(\hat{g}\right)$ is the $[2 \times 1]$ column vector of the derivatives of the function $\hat{g}$ with respect to the two parameters, that is,

$$D\left(\hat{g}\right) = \begin{pmatrix} \frac{1}{1-\hat{\alpha}_3} \\ \\ \frac{\hat{\alpha}_2}{(1-\hat{\alpha}_3)^2} \end{pmatrix}$$

The estimation was carried out using all the countries in the sample. However, the summary statistics of these tables, like those in the remainder of the section, are calculated with the exclusion of countries whose estimated g differs more than 3.5 standard deviations of the mean value. This is done on the grounds that the figures are dominated otherwise by a small number of countries. The rest of the parameters, apart from the estimated growth rate do not differ significantly. The mean of the estimated λ's across countries is (standard errors in brackets) 0.80177 (0.10608), 0.82460 (0.09943) and 0.83803 (0.09642) in the non oil, intermediate and OECD group of countries respectively. These estimated parameter values imply that the average speeds of convergence are of the order 20%, 18% and 17% per year for the non oil, intermediate and OECD samples. The standard deviations of the distributions of the estimated lambdas across countries are 0.11678, 0.10981 and 0.08097 for the non oil, intermediate and OECD samples respectively. These standard deviations outline the heterogeneity of the values of lambda across countries. This point can also be highlighted by looking at the minimum and maximum values of the estimates of lambda in each group. As an example, for the non oil group of countries the minimum value of lambda across this group is equal to 0.43513 which implies a speed of convergence of around 56% per year. The maximum value of the estimated lambdas across this group is, however, 0.98453 which implies a very slow speed of convergence, i.e. 2% per year.

The other parameter of importance in studies of growth and convergence across countries is the long run growth, g_i. The average values of the estimated growth rates, $\hat{g}$, (standard errors in brackets) are 0.00848 (0.02553), 0.01330 (0.02016) and 0.01810 (0.00957).[42] Similar to the distribution of the lambdas across countries, it is found that the estimated long run growth rates also differ across countries, although they seem to be closer together in the OECD group. The standard deviations of the distributions across the non oil, intermediate and OECD groups are respectively 0.03014, 0.02563 and 0.00587. However, for the non-oil and intermediate samples, these average growth rates are not significantly different from zero.

[42]The exclusion of the outlier, makes substantial difference to the estimated growth rate. To give an idea of its size compared to the estimates from the rest of the countries the non oil group is used as an example. The value of the estimated growth rate for the outlier is equal to -0.45, compared to an average value without the outlier of 0.00848. The measures of the mean, standard deviation, kurtosis and skewness when the outlier is included are 0.00276, 0.05957, 44.32844 and -5.50799 respectively.

The main focus of this section is, nevertheless, the estimation of the Solow growth model using the exact discrete model corresponding to the continuous time model which will give consistent estimates of the parameters. In section 3.2, it was shown that the estimated λ_i's from the approximated model are biased upwards. It was also shown that as a consequence of this bias, the distribution of the λ_i's across groups of countries would appear narrower than it actually is. However, the estimated long run growth rate, g_i, obtained from the approximated model is not biased simply because the biases of the parameters used to compute the long run growth rate cancel out.

Tables 3.4, 3.5 and 3.6 present the summary statistics of the parameter estimates obtained using the exact Solow growth model. First, attention is focused on the estimates of λ_i across the three groups of countries. The estimated average values of λ are (standard errors in brackets) 0.76726 (0.13493), 0.79731 (0.11898) and 0.81876 (0.09919) for the non oil, intermediate and OECD groups respectively. In all these cases, the mean of the estimated values of λ is lower than those obtained using the approximated Solow growth model. This is not only true for the mean value, but for every single value of λ obtained. These differences show the upward bias in the estimates of this parameter when the approximated model is used. The "true" average of the speeds of convergence in the non oil, intermediate and OECD samples are, therefore, 23%, 20% and 19% per year. Given that the order of magnitude of the differences of the average speeds of convergence is small when comparing the estimates obtained from the approximated and the exact model, it would seem harmless to use the approximated model. However, this only looks at the average across countries and therefore, two points should be made here. Firstly, it has already been mentioned that there is a lot of heterogeneity on the estimates of the speed of convergence across countries.

Consequently, the average value in itself is not a very good representation of the speed of convergence. Secondly, it was mentioned in section 3.2 that the extent of the bias in the estimates of λ_i depends on the value of λ_i itself; the higher λ_i, the lower the bias and vice versa. Hence, the lowest estimated values of λ_i will carry the biggest biases. This can be better appreciated by looking at the minimum estimated values of λ_i across the three groups of countries. For the non oil group of countries, the minimum value of λ_i when the discrete model is used is equal to 0.43513, implying a speed of convergence to the equilibrium of 56% per annum. However, the minimum value of the λ_i's when the exact model is used, is equal to 0.16795, implying a speed of convergence to the equilibrium of 83% per year, which is a lot higher than that obtained from the approximated model. Most importantly, the biases in the estimated λ_i's have implications on the dispersion of the estimates across countries. The standard deviations of the estimated lambdas across countries when the exact model is used are equal to 0.15978, 0.14646 and 0.0918 for the non oil, intermediate and OECD samples respectively. This represents an increase in the dispersion of the estimates compared to those obtained from the estimation of the approximated Solow growth model. This fact suggests that the speed of convergence across the countries in this study varies even more than what the approximated model would lead one to believe.

Now the estimates of the long run growth rates g_i are considered. These estimates are the same as those obtained by using the approximated Solow growth model as it was anticipated in section 3.2. It should be recalled at this point that the estimate of the long run growth rate was obtained as the ratio $\hat{\alpha}_2/\left(1-\hat{\alpha}_3\right)$. It has been already shown that the estimates of λ (i.e. $\hat{\alpha}_3$) obtained from the approximated model are biased upwards and therefore the denominator $1-\hat{\alpha}_3$ will be biased downwards. When the estimates of $\hat{\alpha}_2$ obtained from the two versions of the Solow growth model are compared, it is easy to see that the approximated version always underestimates the true value of the parameter. However, the estimated long run growth rates, g, obtained from the approximated model are not biased; these estimates are the same as those obtained by using the exact model. This is simply because the bias in the parameter $\hat{\alpha}_2$ is exactly equal to the bias of the denominator, $1 - \hat{\alpha}_3$ and therefore the biases cancel each other out in the expression used to calculate the long run growth rate.

In this section it has been shown how the parameter estimates obtained from the Euler approximation of the Solow growth model differ systematically from those found from the exact discrete model obtained from solving the stochastic differential equation obtained from the continuous time model. Additionally, the distribution of the estimated lambdas have a higher dispersion than the distribution found from the Euler approximation to the continuous time model. Therefore, the estimated speeds of convergence to the equilibrium are even more different than the approximated model shows. These issues outline the importance of obtaining consistent estimates by using the discrete exact model corresponding to the theoretical continuous time model instead of using the approximated model.

3.4.2 Comparison of alternative parameterisations of the Ramsey-Cass-Koopmans model.

The focus of this section is the estimation of the RCK growth model. First, the parameter estimates from the Euler approximation of the continuous time RCK model will be presented. The estimated λ's and g's will then be compared to the estimates obtained earlier from the Euler approximation of the Solow growth model. Then, the exact discrete model corresponding to the continuous time model which will give consistent estimates of the parameters will be estimated. It will be clear at that point the extent of the bias in the speed of convergence to the equilibrium and the differences of both the speed of convergence to the equilibrium and the long run growth rate obtained from the Solow model and the RCK model.

For clarity of exposition, it is useful to write the system in equation (3.12) found by using the Euler approximation as follows

$$\ln \tilde{y}_t = \alpha_{11} + \alpha_{12} t + \alpha_{13} \ln \tilde{y}_{t-1} + \alpha_{14} \ln s_{t-1} + \sigma_1 \epsilon_{1t} \tag{3.17}$$

where

$$\begin{aligned}
\alpha_{11} &= g + (1-\alpha)\Gamma \left[\ln A(0) - g - \frac{\alpha}{1-\alpha} \ln \Gamma \right] \\
\alpha_{12} &= (1-\alpha)\Gamma g \\
\alpha_{13} &= 1 - (1-\alpha)\Gamma = \lambda \\
\alpha_{14} &= \alpha \Gamma
\end{aligned}$$

and

$$\ln s_t = \alpha_{21} + \alpha_{22} t + \alpha_{23} \ln \tilde{y}_{t-1} + \alpha_{24} \ln s_{t-1} + \varrho \epsilon_{1t} + \sigma_2 \epsilon_{2t} \tag{3.18}$$

where

$$\begin{aligned}
\alpha_{21} &= -(\Phi - \alpha\Gamma) \left[\frac{\Phi}{\alpha\theta\Gamma} \ln\left(\frac{\alpha}{\Phi}\right) + \ln \Gamma + \frac{(1-\alpha)}{\alpha} \left(1 - \frac{\Phi}{\alpha\theta\Gamma}\right) (g - \ln A(0)) \right] \\
\alpha_{22} &= \frac{(1-\alpha)(\Phi - \alpha\Gamma)(\alpha\theta\Gamma - \Phi)}{\alpha^2\theta\Gamma} g
\end{aligned}$$

$$\alpha_{23} = -\frac{(1-\alpha)(\Phi - \alpha\Gamma)(\alpha\theta\Gamma - \Phi)}{\alpha^2\theta\Gamma}$$

$$\alpha_{24} = 1 + \Phi - \alpha\Gamma$$

Now it is easy to see that there is one restriction between the two equations in the system, namely,

$$\frac{\alpha_{12}}{1-\alpha_{13}} = \frac{-\alpha_{22}}{\alpha_{23}} \tag{3.19}$$

The estimates of this model are obtained by using maximum likelihood with the restriction in equation (3.19) imposed. Tables 3.7, 3.8 and 3.9 show some descriptive statistics of the estimated coefficients of this model for the non-oil, intermediate and OECD groups. The estimated mean values of λ for the three samples (standard errors in brackets) are 0.75943 (0.10585), 0.78032 (0.08812), 0.83063 (0.04684) for the non-oil, intermediate and OECD group respectively. These average values of λ are lower than those obtained from the approximated Solow model and very similar to those obtained from the exact Solow model. However, in the RCK model, the speed of convergence is not $1-\lambda$ (as is the case in the Solow growth model) but it is given by the negative eigenvalue of the system (see equation (3.9)). Therefore, the calculated speeds of convergence to the steady state are 37%, 35% and 28% per year for the non oil, intermediate and OECD groups of countries respectively. For the approximated Solow model, the speeds of convergence to the equilibrium were 20%, 18% and 17% for the same three groups of countries respectively. Therefore, the speeds of convergence found using the RCK model are almost double those found by using the Solow model. This is a consequence of allowing the savings rate to be a variable instead of using a constant savings rate (usually the average for the sample period) as in the Solow model.

The second parameter of interest is the long run growth, g_i. The mean growth rates are (standard errors in brackets) 0.01510 (0.01301), 0.01790 (0.00850) to 0.01953 (0.00884) for the non oil, intermediate and OECD samples respectively. It is immediately obvious that these mean growth rates are substantially higher than those found when using the Solow growth model. Therefore, even though the in both models the same parameter is being estimated, the RCK model gives an estimate of the long run growth rate which is higher than that obtained from the Solow growth model. The average of the estimated long run growth rate for the non oil group of countries does not seem to be significant, which was also the case when using the Solow model. However, for the intermediate group of countries, the average of the estimated long run growth rate is well determined in contrast to what was found in the Solow model. Also of interest is the dispersion of the estimated values of g which is smaller in both the non-oil and the intermediate group, although the decrease in dispersion is more noticeable in the non-oil group. However, this is not the case for the OECD group. Therefore, it seems to be the case that when the more general RCK model is used, the distribution of the long run growth rates in the non oil and intermediate samples is slightly narrower than would appear to be the case when the Solow growth model is used. Consequently, the long run growth rates appear more similar across these groups. Nevertheless, this is not the case for the OECD subsample. In this instance, the long run growth rates appear to differ across countries more than they did before.

Having found the parameter estimates using the approximated RCK model, it will be interesting to investigate how the results change when the exact discrete model corresponding to the continuous time RCK model is estimated. When estimating the exact discrete system (equation (3.13)) a numerical calculation of the covariance matrix of the error terms is needed. For this purpose, the integral in equation (3.15) is approximated by using expression (3.14) in the integral, integrating and then taking the first 10 elements of the expansion.[43] Tables 3.10, 3.11 and 3.12 present the summary statistics of the parameter estimates of the exact RCK model.

First, attention is concentrated on the parameter defining the long run growth rate. In section 3.3.2, it was shown analytically how the estimated long run growth rates obtained when using the approximated model are not biased, even though all the parameters used to calculate this growth rate carried a bias. For this particular parameter, it was simply the case that the biases cancelled each other out, resulting on an unbiased estimate of the long run growth rate. Consequently, in this case, differences in the estimated growth rates are not expected. When the mean values of the estimated g's from the exact RCK model are compared to those obtained from the approximated RCK model, it is found that they are almost the same for the intermediate and OECD groups and very similar for the non-oil group of countries. Looking at the median, however, it is found that the only difference between these estimates is on the fifth decimal point. The median of the estimates obtained from the approximated model are 0.01779, 0.01923 and 0.01871 for the non-oil, intermediate and OECD samples, compared to 0.01775, 0.01926 and 0.01869 for the same samples in the exact discrete model. Therefore, it can be concluded, as expected, that there is no difference in the estimated long run growth rates when the exact model is used as opposed to the approximation.

Having established that the estimated long run growth rates are, in fact, unbiased even when using the approximated model, attention is turned to the other parameter

[43]Estimation was repeated for a higher number of elements in the expansion, but this did not affect the results.

of interest, namely the speed of convergence. However, it is worth first discussing the differences in the estimated λ's. The first important issue to notice is that the mean of the estimated values of λ are lower than those obtained from the approximation of the RCK model (standard errors in brackets); 0.70028 (0.12753), 0.72256 (0.09562) and 0.80109 (0.05044). Additionally, these are even lower if compared to the approximation of the Solow model. The fact that the estimated λ's are lower when using the exact model instead of the approximation is consistent with the analytical derivation of the bias in section 3.3.2. Another important issue to highlight here with respect to the estimated λ's across countries refers to the difference in the dispersion of the distribution of this parameter across countries when using the exact RCK model as opposed to the approximated model. Once more, it is found that the dispersion of this parameter around the mean increases. Even though the estimates of λ are lower in the case of the RCK model as opposed to the Solow model, it cannot yet be concluded that the average speeds of convergence are higher than those of the Solow model. This is because the speed of convergence cannot be calculated in the RCK model as $1-\lambda$ as was the case in the Solow growth model. The differences in the average speed of convergence below will be discussed later, but first, it is very important to highlight the fact that a different estimate of λ implies differences in the parameters of the underlying model. The average speed of convergence to the steady state in the exact RCK model are 52%, 49% and 36% per year for the non oil, intermediate and OECD groups of countries. These convergence speeds are more than double the speeds of convergence found by using the approximated Solow growth model (20%, 18% and 17% per year for the same three groups of countries respectively) and, therefore, they imply transitional periods which are a lot shorter than would have been concluded from using the Solow growth model.

This section has successfully demonstrated the differences arising in the parameters of interest with respect to other studies of growth when the two developments proposed here are employed. The first development considers the use of the exact discrete model that can be found by integrating the continuous theoretical model as opposed to its approximation which is normally used in the literature. It has been shown how the parameter estimates obtained from the simple Euler approximation differ systematically from those found from the exact discrete model. As a direct consequence of this bias in the parameter estimates, the distribution of the estimated speeds of convergence to the equilibrium are found to have higher dispersion around the mean than the distribution found from the Euler approximation to the continuous time model. These differences are of greatest importance when trying to assess the growth processes across countries since the estimated speeds of convergence across countries are even more different than the approximated model would lead one to believe. Consequently, the idea of groups of countries having similar growth processes is overwhelmingly rejected when the exact model is used. The second development considers the use of the system of equations model by Ramsey-Cass-Koopmans as opposed to the Solow growth model. The addition of an endogenously time varying savings rate produces different estimates of both parameters of interest in any growth studies, the long run growth rate and the speed of convergence to the steady state. One one hand, the estimated long run growth rates are found to be higher than those obtained from the Solow model. On the other hand, and perhaps the most dramatic difference arises in the speeds with which countries converge to their own steady state. These speeds of convergence are very high, more than double those found by using the Solow model. Therefore, the transition periods are found to be considerably shorter.

These issues outline the importance of obtaining consistent estimates of the parameters by using the discrete exact model corresponding to the theoretical continuous time model instead of using the approximated model. They also highlight the significance of using more sophisticated models of the growth process.

3.5 Conclusions.

The Solow growth model is the reference point of most growth analysis. This model can be derived in both continuous time and discrete time. From the theoretical point of view, both derivations are equivalent in the sense that they give rise to the same predictions about growth and convergence. Furthermore, it might appear easier for estimation purposes to derive the model in discrete time since data in economics is recorded at discrete time intervals. However, from the estimation point of view, the models obtained for estimation from the discrete time and continuous time derivations are not equivalent. In this chapter, it has been shown that deriving the Solow model using discrete time actually delivers the Euler approximation of the continuous time model. The discrete model, which has the same properties as the continuous time model, can be obtained by solving either a stochastic differential equation in the case of the Solow model or a system of stochastic differential equations in the case of the RCK model. Since Euler approximation is not the right discretization of the continuous time model, the estimates obtained based on the discrete version are inconsistent. Therefore, any analysis of growth based on discrete versions of models is unreliable and could potentially lead to the wrong conclusions.

In this chapter, the Ramsey-Cass-Koopmans (RCK) model of growth has also been introduced as an alternative to the Solow growth model. The RCK model is an optimization neoclassical model which introduces endogeneity of the savings rate in contrast to a constant, exogenously given savings rate assumed in the Solow growth model. As a first step, both the approximated and the exact Solow growth model were estimated. It was clear from this estimation exercise that the estimates arising from the approximated version of the Solow growth model were systematically different to those obtained from the exact discrete version. As a direct consequence of these systematic differences, the estimated speeds of convergence to the steady state were found to be even more heterogenous across countries than what the approximated Solow growth model would lead one to believe. Finally, the RCK model was estimated in both its approximated and exact forms. Once more, the systematic differences in the parameter estimates from the exact and approximated models were evident. Additionally, it was found that both parameters of interest, the long run growth rate and the speed of convergence to the equilibrium were quantitatively higher than those found by using the Solow growth model. In particular, speeds of convergence to the steady states across countries were found to be more than double those obtained using the approximated form of the Solow growth model implying much shorter transitional periods.

To conclude, this chapter highlights the importance of both of the two developments of the standard discrete time version of the Solow growth model used so widely in the literature to consider cross country growth and convergence. It has been found that estimating the exact discrete model corresponding to the theoretical continuous time model can generate substantially different conclusions on growth dynamics than those attained from simple discrete time versions of the model. In addition, it has also found that introducing the savings rate an endogenous variable in a growth model, as suggested by RCK, provides more sophisticated dynamics which again generate different empirical conclusions to those based on Solow. Certainly, these results suggest care should be taken in interpreting many of the conclusions drawn on growth and convergence in the literature.

Table 3.1. Descriptive statistics of the estimated coefficients from the Solow growth model. Euler approximation.

Non-oil countries (n=81)					
Dependent variable ln $(\tilde{y}_t)$					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\sigma}$	$\hat{g}$
Mean	1.48025	0.00279	0.80177	0.04838	0.00848
Mean Std. Error	0.78672	0.00279	0.10608	0.00640	0.02553
Std. Error	0.08992	0.00062	0.01306	0.00282	0.00337
Median	1.35389	0.00212	0.81901	0.04027	0.01288
Std. Deviation	0.80431	0.00556	0.11678	0.02527	0.03014
Kurtosis	-0.15189	5.61622	0.47366	1.27899	5.02950
Skewness	0.67460	1.77050	-0.83196	1.06666	-1.32153
Minimum	0.14676	-0.00575	0.43513	0.01376	-0.12628
Maximum	3.64313	0.02826	0.98453	0.14116	0.07121

Table 3.2. Descriptive statistics of the estimated coefficients from the Solow growth model. Euler approximation.

Intermediate group (n=49)					
Dependent variable ln $(\tilde{y}_t)$					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\sigma}$	$\hat{g}$
Mean	1.39155	0.00343	0.82460	0.03803	0.01330
Mean Std. Error	0.77499	0.00269	0.09943	0.00507	0.02016
Std. Error	0.11819	0.00079	0.01585	0.00278	0.00370
Median	1.16468	0.00256	0.84181	0.03549	0.01653
Std. Deviation	0.81888	0.00550	0.10981	0.01929	0.02563
Kurtosis	0.22019	9.68775	0.95645	2.58847	4.41593
Skewness	0.85571	2.59005	-1.04095	1.40635	-0.94650
Minimum	0.14676	-0.00575	0.51959	0.01376	-0.08554
Maximum	3.64313	0.02826	0.98453	0.10679	0.07121

Table 3.3. Descriptive statistics of the estimated coefficients from the Solow growth model. Euler approximation.

OECD group (n=18)					
Dependent variable ln $(\tilde{y}_t)$					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\sigma}$	$\hat{g}$
Mean	1.46501	0.00297	0.83803	0.02118	0.01810
Mean Std. Error	0.84107	0.00258	0.09642	0.00298	0.00957
Std. Error	0.17944	0.00040	0.01964	0.00111	0.00142
Median	1.48122	0.00311	0.83756	0.02096	0.01774
Std. Deviation	0.73983	0.00164	0.08097	0.00458	0.00587
Kurtosis	-1.28522	-1.01028	-1.21102	0.54835	0.12554
Skewness	0.29263	0.09757	-0.25423	0.42922	-0.29803
Minimum	0.54549	0.00031	0.70340	0.01376	0.00515
Maximum	2.67621	0.00580	0.94128	0.03174	0.02779

Table 3.4. Descriptive statistics of the estimated coefficients from the Solow growth model. Exact model.

Non-oil countries (n=81)					
Dependent variable ln $(\tilde{y}_t)$					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\sigma}$	$\hat{g}$
Mean	1.72686	0.00336	0.76726	0.05464	0.00848
Mean Std. Error	0.98803	0.00338	0.13493	0.00809	0.02474
Std. Error	0.12145	0.00077	0.01786	0.00336	0.00337
Median	1.49753	0.00232	0.80035	0.04605	0.01288
Std. Deviation	1.08625	0.00688	0.15978	0.03001	0.03014
Kurtosis	0.74984	8.03005	2.02797	2.50838	5.02832
Skewness	1.04187	2.19790	-1.30596	1.25525	-1.32153
Minimum	0.14856	-0.00620	0.16795	0.01515	-0.12628
Maximum	4.87680	0.03792	0.98441	0.17987	0.07121

Table 3.5. Descriptive statistics of the estimated coefficients from the Solow growth model. Exact model.

Intermediate group (n=49)					
Dependent variable ln $(\tilde{y}_t)$					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\sigma}$	$\hat{g}$
Mean	1.59938	0.00410	0.79731	0.04229	0.01331
Mean Std. Error	0.91794	0.00317	0.11898	0.00609	0.01911
Std. Error	0.15575	0.00101	0.02114	0.00324	0.00370
Median	1.26906	0.00281	0.82780	0.03903	0.01653
Std. Deviation	1.07910	0.00697	0.14646	0.0225	0.02564
Kurtosis	1.37148	12.40575	2.30541	2.28723	4.41549
Skewness	1.22703	3.03350	-1.45787	1.40476	-0.94696
Minimum	0.14856	-0.00620	0.34534	0.01515	-0.08555
Maximum	4.86815	0.03792	0.98441	0.11916	0.07121

Table 3.6. Descriptive statistics of the estimated coefficients from the Solow growth model. Exact model.

OECD group (n=18)					
Dependent variable ln $(\tilde{y}_t)$					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\sigma}$	$\hat{g}$
Mean	1.63787	0.00332	0.81876	0.02320	0.01811
Mean Std. Error	0.86819	0.00249	0.09919	0.00313	0.00751
Std. Error	0.21735	0.00047	0.02381	0.00131	0.00143
Median	1.61376	0.00357	0.82274	0.02369	0.01774
Std. Deviation	0.89616	0.00192	0.09818	0.00538	0.00588
Kurtosis	-1.14059	-0.94004	-1.06665	-0.41562	0.10915
Skewness	0.40194	0.17632	-0.37462	0.18953	-0.30010
Minimum	0.56157	0.00032	0.64822	0.01515	0.00515
Maximum	3.16877	0.00688	0.93950	0.03358	0.02779

Table 3.7. Descriptive statistics of the estimated coefficients from the Ramsey-Cass-Koopmans growth model. Euler approximation.

Non-oil countries (n=81)						
	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	$\hat{\alpha}_{13}$	$\hat{\alpha}_{14}$	$\hat{\sigma}_1$	$\hat{g}$
Mean	1.67127	0.00391	0.75943	0.04599	0.04516	0.01510
Mean Std. Error	0.71234	0.00339	0.10585	0.05446	0.00885	0.01301
Std. Error	0.10089	0.00073	0.01543	0.01086	0.00273	0.00277
Median	1.41533	0.00291	0.79022	0.05148	0.03870	0.01779
Std. Deviation	0.90237	0.00651	0.13802	0.09715	0.02443	0.02479
Kurtosis	-0.30289	3.42230	-0.04923	1.22768	3.50656	0.60447
Skewness	0.55595	1.28138	-0.74996	-0.32520	1.54948	-0.27375
Minimum	-0.23605	-0.00695	0.40455	-0.25436	0.01260	-0.05003
Maximum	3.73409	0.03076	0.97951	0.32395	0.14987	0.07924
	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	$\hat{\alpha}_{23}$	$\hat{\alpha}_{24}$	$\hat{\sigma}_2$	$\hat{\varrho}$
Mean	0.51776	0.00173	0.04394	0.64789	0.13107	0.03453
Mean Std. Error	2.26821	0.00852	0.33955	0.15470	0.01918	0.02908
Std. Error	0.44267	0.00161	0.06475	0.02522	0.01005	0.00611
Median	0.70892	0.00022	0.00160	0.65992	0.12758	0.04100
Std. Deviation	3.95938	0.01441	0.57910	0.22556	0.08994	0.05468
Kurtosis	6.46667	4.30620	5.63451	0.16954	3.39253	0.15407
Skewness	-1.26559	1.49501	1.07175	-0.65596	1.50059	-0.47076
Minimum	-18.76587	-0.02645	-1.47157	0.05485	0.01703	-0.12343
Maximum	10.57146	0.05638	2.73752	1.04911	0.50319	0.14361

Table 3.8. Descriptive statistics of the estimated coefficients from the Ramsey-Cass-Koopmans growth model. Euler approximation.

Intermediate group (n=49)						
	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	$\hat{\alpha}_{13}$	$\hat{\alpha}_{14}$	$\hat{\sigma}_1$	$\hat{g}$
Mean	1.59975	0.00431	0.78032	0.04558	0.03515	0.01790
Mean Std. Error	0.61076	0.00253	0.08812	0.05537	0.00638	0.00850
Std. Error	0.13068	0.00083	0.01974	0.01588	0.00255	0.00292
Median	1.34112	0.00316	0.81969	0.05297	0.03390	0.01923
Std. Deviation	0.90540	0.00576	0.13677	0.11000	0.01766	0.02021
Kurtosis	-0.09595	8.80810	0.48005	1.05368	5.16909	2.56218
Skewness	0.85509	2.18896	-1.01147	-0.44182	1.83831	-0.52791
Minimum	0.24914	-0.00695	0.40455	-0.25436	0.01260	-0.05003
Maximum	3.73409	0.03076	0.97951	0.32395	0.10852	0.06176
	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	$\hat{\alpha}_{23}$	$\hat{\alpha}_{24}$	$\hat{\sigma}_2$	$\hat{\varrho}$
Mean	0.53171	0.00026	0.06317	0.67033	0.09040	0.04551
Mean Std. Error	1.84286	0.00629	0.27378	0.15101	0.01214	0.02101
Std. Error	0.35232	0.00145	0.04837	0.03114	0.00832	0.00631
Median	0.75627	0.00016	0.00160	0.66166	0.08698	0.04606
Std. Deviation	2.44094	0.01003	0.33512	0.21575	0.05768	0.04375
Kurtosis	-0.24502	10.80109	-0.19728	-0.28450	-0.00436	0.71923
Skewness	0.21087	1.57922	0.04606	-0.47610	0.82308	-0.08304
Minimum	-3.98964	-0.02645	-0.76916	0.14127	0.01703	-0.06168
Maximum	6.28126	0.04738	0.73443	1.04911	0.24742	0.14361

Table 3.9. Descriptive statistics of the estimated coefficients from the Ramsey-Cass-Koopmans growth model. Euler approximation.

OECD group (n=18)						
	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	$\hat{\alpha}_{13}$	$\hat{\alpha}_{14}$	$\hat{\sigma}_1$	$\hat{g}$
Mean	1.51873	0.00324	0.83063	-0.00563	0.02012	0.01953
Mean Std. Error	0.32544	0.00177	0.04684	0.05099	0.00751	0.00884
Std. Error	0.18356	0.00061	0.02518	0.02719	0.00140	0.00307
Median	1.27976	0.00274	0.87574	0.02592	0.02047	0.01871
Std. Deviation	0.75685	0.00252	0.10381	0.11209	0.00588	0.01266
Kurtosis	0.61098	-0.58608	-0.26121	0.53525	4.25502	3.59067
Skewness	1.17753	0.72456	-0.90655	-0.50001	1.88012	1.51255
Minimum	0.78795	0.00032	0.62331	-0.25436	0.01260	0.00352
Maximum	3.29060	0.00820	0.97332	0.19822	0.03793	0.05639
	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	$\hat{\alpha}_{23}$	$\hat{\alpha}_{24}$	$\hat{\sigma}_2$	$\hat{\varrho}$
Mean	2.24359	0.00321	-0.15504	0.73298	0.03605	0.04385
Mean Std. Error	0.57592	0.00269	0.09233	0.11526	0.00506	0.01402
Std. Error	0.37414	0.00097	0.04443	0.04179	0.00420	0.00291
Median	1.93079	0.00255	-0.12324	0.75655	0.03385	0.04358
Std. Deviation	1.54260	0.00399	0.18318	0.17229	0.01743	0.01224
Kurtosis	-0.50863	-0.04855	0.58331	-1.13804	3.07532	4.18350
Skewness	0.04019	0.15474	0.38391	-0.34812	1.86542	2.26620
Minimum	-0.85858	-0.00499	-0.45903	0.44364	0.01703	0.03165
Maximum	4.76320	0.01016	0.27982	0.98104	0.08541	0.07607

Table 3.10. Descriptive statistics of the estimated coefficients from the Ramsey-Cass-Koopmans growth model. Exact model.

Non-oil countries (n=81)						
	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	$\hat{\alpha}_{13}$	$\hat{\alpha}_{14}$	$\hat{\sigma}_1$	$\hat{g}$
Mean	2.05277	0.00484	0.70028	0.06390	0.03861	0.01603
Mean Std. Error	0.79342	0.00876	0.12753	0.07596	0.02188	0.01882
Std. Error	0.13956	0.00091	0.02197	0.01601	0.00274	0.00311
Median	1.64496	0.00344	0.75212	0.06783	0.03547	0.01775
Std. Deviation	1.24828	0.00818	0.19654	0.14319	0.02467	0.02779
Kurtosis	0.04311	3.38451	1.22322	1.31647	3.83549	5.97611
Skewness	0.78727	1.29979	-1.14321	-0.39300	1.57127	1.08200
Minimum	-0.36945	-0.01101	0.03252	-0.35964	0.00000	-0.04577
Maximum	5.37428	0.03877	0.99146	0.44976	0.14519	0.14963
	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	$\hat{\alpha}_{23}$	$\hat{\alpha}_{24}$	$\hat{\sigma}_2$	$\hat{\varrho}$
Mean	0.57239	0.00271	0.07644	0.50750	0.07228	0.05151
Mean Std. Error	2.70244	0.02708	0.42164	0.23874	0.06852	0.04766
Std. Error	0.72869	0.00259	0.10744	0.04460	0.00708	0.00743
Median	0.77825	0.00024	0.01593	0.58564	0.06284	0.04872
Std. Deviation	6.51763	0.02317	0.96102	0.39894	0.06368	0.06688
Kurtosis	5.80418	5.42576	4.93310	1.20071	2.01142	0.31735
Skewness	-1.12572	1.70193	1.05806	-1.19658	1.20373	-0.38196
Minimum	-29.67441	-0.04505	-2.12696	-0.67134	0.00000	-0.12081
Maximum	15.58769	0.10209	4.29998	1.06290	0.31442	0.20247

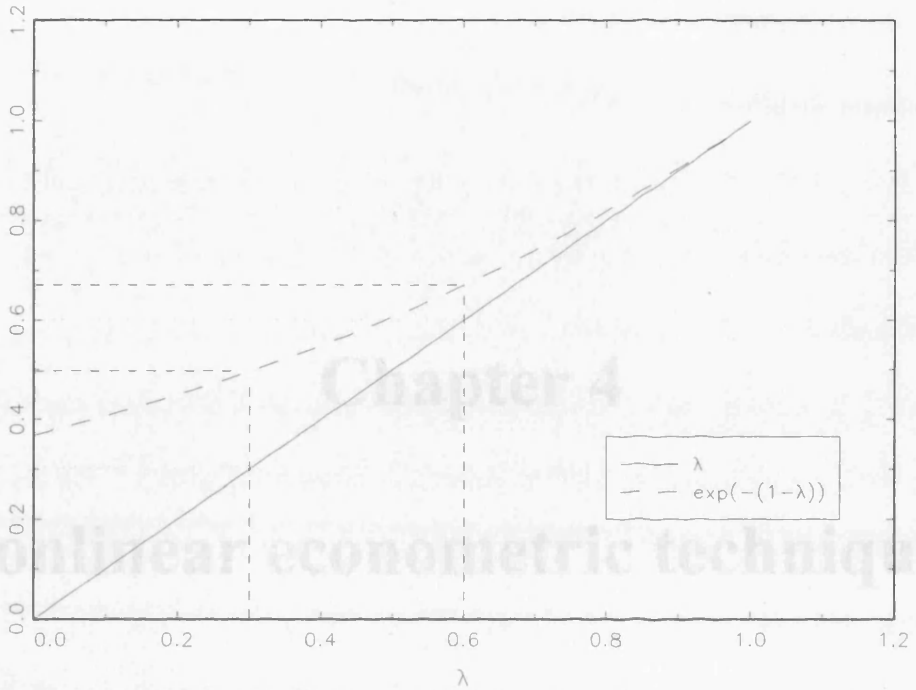
Table 3.11. Descriptive statistics of the estimated coefficients from the Ramsey-Cass-Koopmans growth model. Exact model.

Intermediate group (n=49)						
	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	$\hat{\alpha}_{13}$	$\hat{\alpha}_{14}$	$\hat{\sigma}_1$	$\hat{g}$
Mean	1.98856	0.00543	0.72256	0.06155	0.03021	0.01799
Mean Std. Error	0.59124	0.00783	0.09562	0.07074	0.02103	0.03688
Std. Error	0.18936	0.00108	0.02997	0.02353	0.00252	0.00283
Median	1.56320	0.00344	0.78668	0.07343	0.02471	0.01926
Std. Deviation	1.31192	0.00745	0.20765	0.16304	0.01764	0.01957
Kurtosis	0.28244	7.85109	2.33309	1.09046	5.92567	1.87276
Skewness	1.06960	2.21904	-1.53388	-0.42218	1.86729	-0.29754
Minimum	0.17333	-0.00735	0.03252	-0.35964	0.00246	-0.04142
Maximum	5.37428	0.03877	0.99146	0.44976	0.10494	0.06176
	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	$\hat{\alpha}_{23}$	$\hat{\alpha}_{24}$	$\hat{\sigma}_2$	$\hat{\varrho}$
Mean	0.48333	-0.00036	0.12118	0.54667	0.04926	0.06205
Mean Std. Error	1.73234	0.00958	0.26742	0.20117	0.05663	0.03871
Std. Error	0.54441	0.00224	0.07721	0.05193	0.00561	0.00768
Median	0.77825	0.00012	0.01593	0.57541	0.03548	0.05229
Std. Deviation	3.77176	0.01555	0.53493	0.35980	0.03928	0.05374
Kurtosis	0.30611	8.78524	0.58410	0.59977	-0.81607	1.42479
Skewness	0.28309	0.96192	0.40623	-0.92714	0.53042	-0.01295
Minimum	-6.52870	-0.04505	-1.10054	-0.53196	0.00000	-0.09825
Maximum	10.14906	0.06779	1.47336	1.06290	0.14148	0.20247

Table 3.12. Descriptive statistics of the estimated coefficients from the Ramsey-Cass-Koopmans growth model. Exact model.

OECD group (n=18)						
	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	$\hat{\alpha}_{13}$	$\hat{\alpha}_{14}$	$\hat{\sigma}_1$	$\hat{g}$
Mean Std. Error	1.80144	0.00380	0.80109	-0.01423	0.01872	0.01952
Std. Error	0.38087	0.00148	0.05044	0.04866	0.00372	0.00562
Std. Error	0.24751	0.00074	0.03250	0.03929	0.00157	0.00306
Median	1.57785	0.00296	0.86295	0.03160	0.01866	0.01869
Std. Deviation	1.02051	0.00304	0.13401	0.16198	0.00664	0.01264
Kurtosis	-0.18162	-0.25495	-0.12819	0.99077	3.18620	3.56201
Skewness	0.96341	0.89561	-1.06909	-0.48637	0.95538	1.50695
Minimum	0.68830	0.00033	0.54136	-0.35964	0.00654	0.00353
Maximum	3.99659	0.00998	0.94731	0.31097	0.03752	0.05627
	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	$\hat{\alpha}_{23}$	$\hat{\alpha}_{24}$	$\hat{\sigma}_2$	$\hat{\varrho}$
Mean	2.86466	0.00393	-0.19508	0.65495	0.03073	0.03995
Mean Std. Error	0.59606	0.00230	0.09924	0.13568	0.01056	0.01150
Std. Error	0.57049	0.00128	0.06275	0.06574	0.00483	0.00389
Median	2.39667	0.00291	-0.15552	0.72069	0.02985	0.04154
Std. Deviation	2.35220	0.00526	0.25873	0.27106	0.02051	0.01649
Kurtosis	0.39679	0.37831	1.12662	-0.56476	0.96223	1.05320
Skewness	0.46955	-0.20162	0.54725	-0.68389	0.77757	-0.40734
Minimum	-1.60375	-0.00807	-0.59757	0.08270	0.00000	0.00148
Maximum	8.01945	0.01300	0.45099	1.00211	0.08109	0.07184

Figure 3.1. Bias of the parameter λ_i when the approximated Solow growth model is used instead of the exact discrete model.



Chapter 4

Nonlinear econometric techniques: the switching regressions model.

4.1 Introduction.

In chapter 2, sections 2.4.1 and 2.4.2, two possible extensions to the theoretical Overlapping Generations model of output growth were described. Both of these extensions showed how the growth of output in an economy might be generated by different processes in two different regimes. In the model of section 2.4.1, the regime depended on whether a country is in balance of payments difficulties or not whereas in the model of section 2.4.2, the regime depended on whether there are technology spillovers across countries at each point in time. Empirically, this type of behaviour may be best characterized by a "switching regime" model. The remainder of the thesis concentrates on the behavioural models considered in chapter 2 and their empirical counterparts. This chapter focuses on the econometric issues raised in switching regression models, and then, applies these modelling techniques in two empirical exercises in chapters 5 and 6. The empirical model in chapter 5 investigates the role of a balance of payments constraint on growth, while chapter 6 investigates the role of technology spillovers across countries as a determinant of growth.

The remainder of this chapter is devoted to a discussion of econometric issues concerned with switching regime models. The purpose of the chapter is both to review the modelling techniques employed, evaluating some of the alternative approaches suggested in the literature, and also to set out and define measures, procedures and statistics that will be used in the subsequent empirical analyses of chapters 5 and 6.

Section 4.2 provides a literature review of a subset of nonlinear models which are relevant to the empirical chapters 5 and 6 and can be grouped under the heading of switching regime models. Section 4.3 describes the particular the switching regime model which will be used in the econometric estimation of chapters 5 and 6. Section 4.4 presents the likelihood function of the model. Additionally, in this section it is also shown how problematic the maximization of the likelihood function of the switching regime model considered here can be. The discussion considers some of the alternative methods of estimation proposed in the literature and proposes the use of a global optimisation algorithm, namely, the simulated annealing algorithm, instead of using traditional optimisation algorithms in the context of maximum likelihood estimation of switching regime models.[44]

Section 4.5 focuses on the different testing procedures which can be used in switching regime models. Here, discussion is concerned with diagnostic tests based on the score of the likelihood function, likelihood ratio tests for testing nested hypothesis. It also examines a potentially interesting question specific to the applied work in the next two chapters which is whether both regimes share the same long run growth rate even if the short run dynamics appear to be different.

In section 4.6, the measurement of the persistence of shocks to output in switching regime models is considered. If shocks to output are persistent, it is important to quantify the impact and duration of their effect on output. In linear models, the persistence of a shock is measured by means of an impulse response. However, for multivariate nonlinear models, the impulse response functions are history, shock and composition

[44]In subsequent chapters, it will be examined how useful this algorithm is compared to traditional methods and the Expectation-Maximization algorithm widely used in the econometric literature to estimate switching regime models.

dependent. Thus, in nonlinear models, the impulse response function is actually a random variable. This section details how of the Generalized Impulse Response functions of Koop, Pesaran and Potter (1996) can be applied to the switching regime model studied here. Finally, section 4.7 concludes.

4.2 Switching regime models and their use in the growth literature.

The majority of the literature on output growth has concentrated on linear models (see chapter 2), although recently there has been a switch towards nonlinear models in the study of output growth. The main reason for using linear models is that they are easy to estimate with standard software packages and their dynamic properties can be precisely described by means of their calculated impulse responses. However, nonlinear models allow a more general specification than linear models and linear models can be seen as a special case of any class of nonlinear models (although testing linearity against a nonlinear model is not straightforward due to the presence of nuisance parameters under the null hypothesis).[45] The literature in nonlinear models is rapidly expanding and there are many different classes of models that have been proposed. In this discussion, attention is restricted to a set of models which can be grouped under the heading of *switching regime models*. In switching regime models, there are at least two regimes or states under which the dependent variable is generated by different processes and there are regime switches over time according to some imposed structure. They can be classified into models for which the change in regime is determined

[45]There is one exception, however, namely the chaos deterministic models. This class of models does not include linearity as a special case.

by an observable variable, models for which the change in regime is unobserved and a mixture of the two.

Observable change in regime.

The most important models under this heading are *Threshold models*. These models are considered in detail in Tong (1990). The dependent variable in these models can be generated at each point in time by one of a set of different processes (typically two or three). The state which generates the dependent variable at each point is determined by an indicator function whose value depends on whether the d-lag (d is called the delay parameter) of a certain variable is above or below a threshold μ. The states are usually modelled as autoregressive processes and, in this case, the model is known as a Threshold Autoregressive (TAR) process. If the variable in the indicator function is the same as the dependent variable the model is known as a Self-Exciting Threshold Autoregression (SETAR). Applications include, among others, the SETAR models of US GNP growth of Tiao and Tsay (1994) and Potter (1995) and Peel and Speight's (1998) SETAR models for output growth for five of the G7 countries (namely, Canada, Germany, Japan, the UK and the US).

Although TAR and SETAR models appear close to linear ones in the sense that it seems only a minor modification, they provide an enormous increase in flexibility in fitting data and capturing dynamics. Correspondingly, the dimensions of modelling work rises quickly (for example in terms of the number of estimated parameters and choices for undertaking specification searches). It is therefore worth describing in a little more detail the modelling work carried out in the applications described above.

Tiao and Tsay (1994) use a SETAR model for output growth with two regimes, each one being an AR(2) process, a delay parameter $d = 2$ and a threshold $\mu = 0$. They later expand this model to a four regime model by splitting each of the two former regimes into two, depending on the relative size of the first lag of the dependent variable with respect to the second lag. Therefore, the four regimes are as follows: a contraction period in which the economy goes further into recession (if the first and second lags of growth are negative or equal to zero and the first lag is equal or even more negative than the second), a contraction period in which the economy starts improving (if the second lag is zero or negative, but the first lag is greater than the second one), an expansion period in which the economy is still improving (both lags are positive with the first lag being greater than the second) and a period of declining growth (the second lag of growth is positive but the first lag is equal or less than the second lag). They found evidence of asymmetric behaviour of US real GNP in these four regimes, although the dynamics of last two regimes were quite similar. Potter (1995) also uses a SETAR model but with only two regimes, each one being an AR(5) process, a delay parameter $d = 2$ and a threshold $\mu = 0$. The third and fourth lags of the AR processes in both regimes are later restricted to zero although the fifth lag is retained.[46] He further shows the asymmetric behaviour of US GNP by making use of nonlinear impulse response functions and finds that the post-1945 US output is more stable. He also finds that negative shocks tend to be alleviated and even reversed in periods of recession.

Peel and Speight (1998) estimate SETAR models for both the trend stationary and difference stationary GNP/GDP for five of the G7 countries.[47] They found that the estimated SETAR models imply a reduction in the residual variance of the models com-

[46]Potter justifies the presence of the fifth lag and the omission of the third and fourth lags by arguing that the fifth lag in both regimes are significant and improve the fit of the model. However, he finds this rather peculiar.

[47]They argued along similar lines as Pippinger and Goering (1993) that threshold nonlinearities could lead to an erroneously identification of a unit root and therefore, they estimate SETAR models for both the trend stationary case and the difference stationary case.

pared to a linear models and that, apart from Canada, these nonlinear models imply asymmetric behaviour of output for all the countries.

Beaudry and Koop's (1993) model of US output growth consists of two regimes also with endogenous switching similar to the models presented earlier. The difference is that instead of using a fixed lag (or lags) of output as the variable governing the switch between regimes they defined a new variable, called the current depth of the recession (CDR_t) which is not fixed over time. CDR_t is defined as the difference between the former maximum level of output and the present level if this quantity is positive and zero otherwise. This variable has the effect of dampening the negative shocks to output (relative to a positive shock) when the economy is in a depression if added to a linear AR model. To assess whether there are asymmetries in the responses of output to shocks they also used impulse responses. The main conclusion of the paper is that the responses of output to shocks are very different as including their persistence. Positive shocks are found to be very persistent while negative shocks are temporary. It has been shown by Pesaran and Potter (1997) that Beaudry and Koop's model of US output growth can be written as a two regime TAR model with endogenous switching depending on whether output is below its previous maximum. Pesaran and Potter took Beaudry and Koop's model of US output growth one step further by defining three rather than two regimes referred to as floor, ceiling and corridor. This is accomplished by redefining the CDR_t variable of Beaudry and Koop together with a new variable, the overheating variable, OH_t. The redefined CDR_t variable allows for a non-zero threshold, so that, small falls in output do not set off the dampening of fluctuations and, therefore, this variable characterizes the floor regime. The OH_t variable is nonzero if for the last two periods the growth rate is above a certain level. Therefore, this OH_t

variable allows for a dampening of positive shocks when an economy has been in a period of expansion. This variable then, characterizes the ceiling regime.[48] The third regime, the corridor regime, takes place when the economy is neither in the floor nor the ceiling regimes. Using Generalised Impulse Response functions (see section 4.6) they found the same type of asymmetry as Beaudry and Koop in the floor regime and also the reverse type of asymmetry in the other two regimes, although this was not as pronounced.

The transitions from one regime to another in these models are very sharp and it is argued that maybe a smooth transition period would be more appropriate to model the economy. To obtain smooth transitions in TAR models it is sufficient to replace the indicator variable by a continuous function of the difference between the variable used for the transition and the threshold. These models are known as Smooth Transition Autoregression (STAR) models.[49] There are a number of different models that arise depending on the form of the continuous function used. For example, by using a logistic or exponential distribution, the model becomes a LSTAR or a ESTAR model respectively. Terasvirta and Anderson (1992) apply these models to an index of industrial production for the US, Japan, Belgium, Canada, Federal Republic of Germany, Italy and an aggregate index of industrial output for European countries members of the OECD. They found that the dynamics of recessions and expansions are different and that in many cases this family of nonlinear models seems to be appropriate for describing the type of nonlinearities found in the series.

[48] If the floor regime is in effect, this specification does not allow a recovery after a recession to set off the ceiling regime.

[49] A generalization of this STAR model is one which includes other explanatory variables. These models are known as Smooth Transition Regression (STR).

Unobservable change in regime.

It is possible, however, that the change in regime is not observable and a literature has also evolved relating to models of this form. Quandt (1972) proposed a switching regression model in which the dependent variable is generated by one out of two or more regimes, each one characterized by a linear regression with differing parameters. There is an unknown probability λ that each observation belongs to the first regime and therefore a probability $1 - \lambda$ that each observation belongs to the second regime. This is also called a mixture model since one can think of the parameters as random variables coming from a mixture distribution. In this model, the probability that an observation belongs to one of the regimes is independent of past realizations of the regime. Goldfeld and Quandt (1973b) relaxed this assumption by introducing a matrix of transition probabilities which, in the case of two regimes, is a 2×2 matrix. This makes the switching process a Markov chain. Based on the specification of Goldfeld and Quandt but applying it to a dynamic context, Hamilton (1989) proposed a model for US real GNP growth. The model is a fourth order autoregression around two different means corresponding to each one of the two regimes. The switch in regime is generated exogenously by an unobserved Markov Chain with constant transition probabilities which depend only on the regime the system was in the previous period. This is known as a Markov switching model. He found evidence of these shifts in the mean of the process in US real GNP data.

Hamilton (1990) also proposes an alternative Markov switching model in which the parameters of the autoregressive process are the ones switching between regimes instead of the mean of the process.[50] This alternative model has the advantage of being able to generate different responses to shocks in both regimes and, is, therefore, capable of dealing with the increasing evidence of asymmetries and nonlinearities in output.[51] If the probability of being in either regime at each point in time is independent of the regime that was active in the last period, this model reduces to a dynamic version of the switching regression model of Quandt (1972) since there is no persistence in the regimes. Models like this have been used extensively in the literature. In the case of output growth it has been used for example by Hansen (1992). Hansen (1992) estimates a Markov switching model of US GNP as proposed in Hamilton (1990) using a fourth order autoregression in each regime.[52] However, the restriction that there is no persistence in the switching process cannot be rejected and therefore, the Markov switching model is rejected in favour of a switching regression model with regimes independent over time.

Possible generalizations of these models have also been suggested in the above literature. For example Goldfeld and Quandt (1973) introduced the idea that rather than the regimes switching with a constant probability (either independent over time or through an unobserved Markov chain), exogenous variables can be introduced so that these probabilities are functions of the exogenous variables. The following section is concerned with a dynamic specification of the switching regressions model of Goldfeld and Quandt (1973a) with the extension of an exogenous variable governing the switch

[50] Hamilton also indicates the possibility of making the innovation variance-covariance matrix dependent on the regime.

[51] Some authors, however, do not find any evidence of asymmetries: for example, Mills (1995a,b), who focuses on UK output amongst other macroeconomic time series and Koop and Potter (2000,2001), who look at real US GDP growth amongst other variables.

[52] Hansen initially estimates the switching regime model with only a switch in mean proposed in Hamilton (1989). However the Markov switching model with shifts in the parameters fits the data better.

between regimes. This type of model will be later used in the empirical chapters 5 and 6.

4.3 An introduction to the "switching regressions" model.

In chapter 2, sections 2.4.1 and 2.4.2, two possible extensions to the theoretical Overlapping Generations model were considered. Both models showed how the growth of output in an economy might be generated by one of two different processes at each point in time. In the model of section 2.4.1, the switching between these two processes was governed by the position of the current account relative to output, that is, whether there was a surplus or a deficit. In the model of section 2.4.2, which process was in effect depended on the presence or absence of technology spillovers across countries. This type of behaviour may be best characterized by a "switching regressions" model of the type described by Goldfeld and Quandt (1973a). Specifically, let $y_{i,t}$ denote the logarithm of output in country i at time t and $\Delta y_{i,t}$ denote output growth. Let $z_{i,t}$ denote the switching variable for country i; that is, any extra information that is available to classify the dependent variable into the two regimes. Then, output growth in economy i can be modelled as

$$\begin{aligned} \Delta y_{i,t} = & \begin{cases} \beta_{i,1,0} + \sum_{j=1}^{k_{i,1}} \beta_{i,1,j} \Delta y_{i,t-j} + \varepsilon_{i,1,t} & \text{if } y^*_{i,t} < 0 \\ \beta_{i,2,0} + \sum_{j=1}^{k_{i,2}} \beta_{i,2,j} \Delta y_{i,t-j} + \varepsilon_{i,2,t} & \text{if } y^*_{i,t} \geq 0 \end{cases} \\ & \text{and} \\ y^*_{i,t} = & \; \beta_{i,3,0} + \beta_{31} z_{i,t} + \varepsilon_{i,3,t} \end{aligned} \tag{4.1}$$

where $y^*_{i,t}$ is a latent variable which classifies the dependent variable into the two regimes and is independent of the past realizations of the process. A more general model, as discussed in the last section, will make the probability of being in a particular regime dependent on the regime the economy was in the last period. That is, a Markov switching model. One of the diagnostic tests which will be introduced in section 4.5 tests for the presence of omitted Markov effects. If the hypothesis of no omitted Markov effects is rejected, then it would be necessary to estimate a Markov model instead of the switching regression model. The errors are assumed to have the following joint distribution

$$\begin{pmatrix} \varepsilon_{i,1,t} \\ \varepsilon_{i,2,t} \\ \varepsilon_{i,3,t} \end{pmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \sigma_{i,1}^2 & \sigma_{i,1,2} & 0 \\ \sigma_{i,1,2} & \sigma_{i,2}^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \tag{4.2}$$

Note that the variance covariance matrix includes the standardization of the variance of $\varepsilon_{i,3,t}$ that is needed for identification. Each regime of this model can be augmented with other additional variables. For example, in chapter 5, each regime is augmented with lags of the mean growth outside the country under study, namely,

$$\bar{y}_{i,t-j} = \frac{1}{P-1} \sum_{p=1, p\neq i}^{P} y_{p,t-j}$$

where P is the number of countries in the sample. That is,

$$\Delta y_{i,t} = \begin{cases} \beta_{i,1,0} + \sum_{j=1}^{k_{i,1}} \beta_{i,1,j} \Delta y_{i,t-j} + \sum_{j=1}^{\tilde{k}_{i,1}} \gamma_{i,1,j} \Delta \bar{y}_{i,t-j} + \varepsilon_{i,1,t} & \text{if } y_{i,t}^* < 0 \\ \beta_{i,2,0} + \sum_{j=1}^{k_{i,2}} \beta_{i,2,j} \Delta y_{i,t-j} + \sum_{j=1}^{\tilde{k}_{i,2}} \gamma_{i,2,j} \Delta \bar{y}_{i,t-j} + \varepsilon_{i,2,t} & \text{if } y_{i,t}^* \geq 0 \end{cases}$$

and

$$y_{i,t}^* = \beta_{i,3,0} + \beta_{i,3,1} z_{i,t} + \varepsilon_{i,3,t} \tag{4.3}$$

with the same joint distribution of the errors as before. Since this only affects the number of parameters to be estimated but it does not affect the econometric discussion, the model in equation (4.1) will be used for simplicity of exposition. Also, the superscripts i's designating the countries will be omitted as long as there is no scope for misinterpretation.

It is worth noting that the nonlinear effects of this model arise not only because of the changes in regime over time, but also because the variances of the errors are allowed to differ between regimes. An important property of this model is that, even though the regime is not actually observed, one can formulate probability statements for each observation in the sample period regarding the regimes. For example, conditioning on z_t, the probability of being in the first regime is $\Phi\left(-\beta_{3,0} - \beta_{3,1} z_t\right)$, while conditioning also on Δy_t, the probability is

$$\frac{\Phi\left(-\beta_{3,0} - \beta_{3,1} z_t\right) f_{1,t}}{\Phi\left(-\beta_{3,0} - \beta_{3,1} z_t\right) f_{1,t} + \Phi\left(\beta_{3,0} + \beta_{3,1} z_t\right) f_{2,t}}$$

where

$$f_{r,t} = \phi\left(\frac{\Delta y_t - \beta_{r,0} - \sum_{j=1}^{k_r} \beta_{r,j} \Delta y_{t-j}}{\sigma_r}\right) \frac{1}{\sigma_r} \qquad r = 1, 2 \tag{4.4}$$

Now that the switching regression model has been introduced, the following section is devoted to the different estimation methods which have been proposed in the literature, their advantages and their shortcomings. The next section also introduces the methodology which will be used to find the estimates of the parameters of the switching regressions model in chapters 5 and 6.

4.4 Estimation methods for the switching regression model.

The present section focuses on the methodology proposed in the literature for the estimation of the switching regressions model in equation (4.1) and their shortcomings. Section 4.4.1 presents the likelihood function of the switching regime model (4.1), given the assumption of normality of the errors. In practice, an obvious estimate of the parameters could then be found by maximizing this function using conventional algorithms for optimisation. This is, however, not straightforward. Two potential problems should be taken into account in the estimation of these models. First, due to the highly nonlinear structure of the model, the likelihood function might present several local maxima. Analytically, this does not represent a problem, the global maximum among all the possible local maxima of the function need to be selected. However, in practice, the optimisation of the function is usually carried out using conventional algorithms. These algorithms are "local optimisation" algorithms. Consequently, if the likelihood function is multimodal, conventional algorithms will stop at the first optimum found. Thus, several trials with different initial values need to be run. The second problem arises from the likelihood function itself. This function might become unbounded for some parameter values. If the likelihood function is unbounded, conventional algorithms will produce higher and higher values of the likelihood function at each step without ever converging, as will be shown in section 4.4.1.[53]

Some authors have, therefore, proposed different methods of estimation which they claim are better suited for this kind of problem. These methods are discussed in section 4.4.2. In section 4.4.3, the use of maximum likelihood is supported. How-

[53]In practice, the algoritms will converge eventually to a point clearly in the unbounded region which depends on the precision of the computer used.

ever, instead of using conventional algorithms for the maximization of the likelihood function, the use of a global optimisation algorithm, namely the simulated annealing algorithm is proposed. This algorithm helps with the two difficulties of the likelihood function. Firstly, since it is a *global* optimisation algorithm, the potential multimodality of the likelihood function is not a problem. Secondly, the output of this algorithm enables the user to learn about the likelihood function and, therefore, it is very easy to spot the unbounded region of the likelihood function. Once this is known, it is straightforward to constrain the algorithm from going into that region.

4.4.1 Maximum likelihood estimation.

The likelihood function for the observed data of the switching regime model (4.1), given the assumption in equation (4.2) of normality of the errors, can be calculated as follows (see, for example, Goldfeld *et al* (1971)):

$$\begin{aligned} L &= \prod_{t=1}^{T}\left[\Phi\left(-\beta_{3,0}-\beta_{3,1}z_t\right)\phi\left(\frac{\Delta y_t-\beta_{1,0}-\sum_{j=1}^{k_1}\beta_{1,j}\Delta y_{t-j}}{\sigma_1}\right)\frac{1}{\sigma_1}+\right. \\ &\qquad \left.+\Phi\left(\beta_{3,0}+\beta_{3,1}z_t\right)\phi\left(\frac{\Delta y_t-\beta_{2,0}-\sum_{j=1}^{k_2}\beta_{2,j}\Delta y_{t-j}}{\sigma_2}\right)\frac{1}{\sigma_2}\right] \qquad (4.5) \\ &= \prod_{t=1}^{T}\left[\Phi\left(-\beta_{3,0}-\beta_{3,1}z_t\right)f_{1,t}+\Phi\left(\beta_{3,0}+\beta_{3,1}z_t\right)f_{2,t}\right] \qquad (4.6) \end{aligned}$$

It was stated in section 4.2 that this is a "mixture model" because at any point in time both regimes have non zero contributions to the model and the density of Δy_t is a mixture of the densities of both regimes. This switching regime model can be considered as one with varying parameters in the sense that, although the parameters of each regime are constant, the contributions of each regime (and thus the individual parameter contributions) to the structural model vary over time. It is important to note at this point that the covariance between the two regimes, σ_{12}, is not identified (see, for example, Maddala and Nelson (1975)), since it does not enter the likelihood function.

Various problems have been noted with switching regressions models. One problem is that the likelihood function for these models becomes unbounded for certain parameter values (see, for example, Maddala and Nelson (1975) and Maddala (1983)). To illustrate the unboundness of the likelihood function in equation (4.6), it is useful to recall that this likelihood function is a mixture of the density functions of the two regimes, $f_{1,t}$ and $f_{2,t}$, weighted by their corresponding probabilities of being in each regime. Thus, each element of the likelihood is made up of two components. It is immediately obvious that a problem could arise if the standard deviation of one of the regimes tends to zero since the density function of that regime will tend to infinity. Let the standard deviation of one of the regimes be different from zero, for example that of the first regime, $\sigma_1 \neq 0$. If at some value of t, say t^*, $\Delta y_{t^*} = \hat{\beta}_{20} + \sum_{j=1}^{k_2} \hat{\beta}_{2j} \Delta y_{t^*-j}$, that is, for that observation and the current parameter estimates there is a perfect fit in regime 2, then, the density function of the second regime for that observation will tend to infinity as $\sigma_2 \to 0$, $f_{2,t^*} \to \infty$ whereas this density function for the rest of the observations will tend to zero, $f_{2,t} \to 0\ \forall t \neq t^*$since all the mass of this density function is concentrated at $t = t^*$. Since $f_{1,t}$ will be finite for any t, the likelihood function L tends to infinity. Following the same reasoning, the likelihood function also tends to infinity as $\sigma_1 \to 0$ if $\sigma_2 \neq 0$ and $\Delta y_t = \hat{\beta}_{10} + \sum_{j=1}^{k_1} \hat{\beta}_{1j} \Delta y_{t-j}$. Thus, there are some parameter values for which the likelihood function is unbounded. In practice, this means that any maximization algorithm might produce higher and higher values of the likelihood function as one of the standard deviations tends to zero, without converging or eventually converging to a value close to zero for one of the standard deviations. Hence, attempting to locate a global maximum in these cases leads to inconsistent estimates of the parameters because even if the algorithm finally converges (in the region of the unbounded solution), the estimates will not be consistent since this in not a suitable maximum of the likelihood function.

Kiefer (1978) proved that when the likelihood function is unbounded, a local maximum of the function is consistent, asymptotically normal and efficient if the likelihood is unimodal away from the bounds. Hence, in the case of unbounded likelihoods it is enough to find the local maximum of the function situated in the region where the likelihood is not unbounded. However, if the likelihood function is unbounded and, is multimodal away from the bounds, there is no information as to which of all the possible maxima gives consistent estimates of the parameters. Hartley (1978) mentioned that, in these circumstances, "*presumably* the root which maximizes L is the consistent one". In practice, this means that a search through different parameter values is needed to assess whether there is in fact only one local maximum in the case that the likelihood function is unbounded and, if this is not the case, to try to locate the maximum among all the local maxima.

4.4.2 Alternative methods of estimation.

Some alternative methods of estimating switching regime models have been suggested. Quandt and Ramsey (1978), aware of the problems of maximum likelihood, suggested estimation using a moment generating function (MGF) estimator for the switching regression model with constant probabilities; that is, for the special case in which there is no z_t in the model. The parameter estimates in this case are obtained by minimizing the sum of the squared differences between the theoretical and the sample moment generating functions. However, these functions depend on some extra parameters θ_j for $j = 1, \ldots, J$ since the moment generating function of a variable, say y, is equal to the expectation of $\exp(\theta y)$. The actual values of these parameters, as well as the number of them to be used, J, have to be selected before minimization. The number of θ's needs to be at least equal to the number of parameters that are to be estimated. However the choice of the values of θ_j is more problematic. Quandt and Ramsey stated that they should be chosen in such a way that the system of equations of the first order conditions for minimization is nonsingular and presented some Monte Carlo evidence supporting the good performance of this estimator. In discussing this paper, Kiefer (1978) pointed out that although the MGF estimator gives consistent estimates, these are not efficient and the choice of θ's is subjective. Therefore, he suggested using the MGF estimator to find consistent estimates and then, supply these estimates as initial values for maximum likelihood.

A different approach was followed by Hartley (1977, 1978). He proposed the use of an algorithm equivalent to the Expectation - Maximization (EM) algorithm to solve the system of equations of the first order derivatives of the likelihood function of the switching regressions model instead of using conventional algorithms to maximise directly the likelihood function. He presented some limited Monte Carlo evidence of the robustness of this approach in the sense that the algorithm seems to converge always to a local maximum of the likelihood function. His evidence suggested that the EM algorithm could be quicker to start with, in the sense that it is less sensitive to the choice of initial parameter values, but might take longer to converge than maximum likelihood optimised with traditional algorithms near the final parameter estimates. This indicates that a useful procedure would be to start the estimation with the EM algorithm and then switch to maximum likelihood. However, Kiefer (1978) found that the EM algorithm tends to drift away from the solution in models that have unbounded likelihood functions. Moreover, he encountered this problem even when the initial values given to the algorithm were the true parameter values (Maddala (1983)). In chapter 5, the performance of the EM algorithm in conjunction with maximum likelihood and traditional optimisation algorithms will be compared with the approach proposed in section 4.4.3, namely, the use of the Simulated Annealing algorithm to maximize the likelihood function.

4.4.3 A preferred method of estimation using the Simulated Annealing algorithm.

In section 4.4.1, two problems in the maximization of the likelihood function for the switching regression model were pointed out: unboundness and multimodality. In this section, a different algorithm for the maximization of the likelihood function is proposed, namely the "Simulated Annealing" (SA) algorithm (see for example Corana *et al* (1987) and Goffe *et al* (1994)). This algorithm helps with the two difficulties in the maximization of the likelihood function. Firstly, since it is a *global* optimisation algorithm, even if the likelihood function is multimodal, the algorithm can get away from local optima once it has reached them and find the global optimum. Secondly, the output of this algorithm is such that allows learning about the likelihood function. Therefore it is very easy to spot the unbounded region of the likelihood function. Once this is known, it is straightforward to constrain the algorithm from going into that region. Although global optimisation algorithms are very good at finding the maximum when the function is multimodal, they might take too long to converge towards the exact point depending on the complexity of the function. However, the running time of the algorithm can be dramatically reduced if only a point in the close vicinity of the maximum is needed. Once such a point is found, it is more efficient to revert to local optimisation algorithms, which are faster, using the final point obtained from the SA algorithm as the starting point for the new search.

Simulated annealing algorithms are based on an analogy with thermodynamics and the way liquids solidify when the temperature falls. If the drop in temperature is slow, the molecules in the material become settled in a highly ordered way, forming a pure crystal which is the global minimum energy state for the material. However, if the drop in temperature is too fast, the solidified material will contain defects; that is, a local minimum with higher energy than the crystal. Maximizing a function with conventional algorithms, which only accept points which yield a higher value of the function to be maximized (uphill moves), is like cooling a liquid rapidly. Hence, these algorithms are very likely to reach only a local maximum instead of the global one. However, the simulated annealing algorithm allows not only uphill moves but also downhill moves which are controlled with the temperature. A step by step description of the algorithm is presented in Appendix III.ii, but without going into the technicalities of the algorithm, it is worth providing a general idea of how the algorithm operates. At high temperatures, the algorithm forms a rough view of the function and as the temperature falls the algorithm concentrates on the parts of the function that look more favourable. The simulated annealing algorithm generates a sequence of points from a given set of initial parameter values. These new points are generated around the current parameters using random moves along the direction of each element in the parameter vector in turn. The new coordinate values are uniformly distributed in intervals centered around the corresponding element of the current vector of parameters. Half the size of these intervals along each coordinate is recorded in the step vector. If the point falls outside the definition domain, a new point is randomly generated in the definition domain.[54] If the value of the function evaluated at the new point is higher than it was previously, the point is accepted. If the function at this new point is lower, then, the point is accepted or rejected according to a given probabilistic criterion (namely, the Metropolis criterion (Metropolis *et al* (1953)). These steps are then repeated a number of times until a convergence condition is satisfied (see Appendix III.ii for further details of this algorithm).

This section has discussed the different methods used in the literature to estimate switching regression models highlighting the shortcomings of each approach. To overcome the shortcomings of the approaches used in the literature, a combination of local or traditional optimisation algorithms and Simulated Annealing was proposed to maximize the likelihood function of these models. The next section deals with the different testing procedures that will be used in the subsequent empirical chapters 5 and 6 after estimation of the switching regressions models.

[54]Before running the SA algorithm, some parameters need to be supplied (see *Step 1* of the SA algorithm in Appendix III.ii for a more detailed explanation of these). One of these is the definition domain, that is, the lower and upper bounds for each parameter to be estimated. These lower and upper bounds can be set to a very big negative and positive number respectively, if no restrictions on the parameters are required. Alternatively, they can be set so that the parameters are restricted to a certain area of interest.

4.5 Evaluation of the switching regressions model.

This section describes the testing procedures that are applicable to switching regression models and that will be used later in the empirical chapters 5 and 6. These procedures require the score (or gradient) and the hessian matrix of the parameters. The gradient is calculated by using the first order analytical derivatives of the likelihood function with respect to the parameters of the model. The hessian matrix is calculated by finding numerically the derivatives of the analytical gradient.

The testing procedures discussed here include diagnostic tests based on the score of the likelihood function and likelihood ratio tests for testing nested hypothesis. Once the model has been estimated, it is very important to check that the model is not misspecified in any way, and tests for the presence of autoregressive errors or conditional heteroskedasticity in each of the regimes are crucial in this regard. In the switching regime model that is the focus of this and subsequent chapters, the regimes are assumed to be independent of past realizations of the regimes. Hence, a test for the presence of omitted Markov effects is essential in this context. These tests are the focus of section 4.5.1.

Section 4.5.2 deals with testing a broader range of important issues. Examples include testing whether the actual dynamics of the regimes are the same or whether the probability of the switch is independent of the exogenous variables. There is an additional procedure derived in section 4.5.3 for testing an important question specific to the applied work in the two next chapters; namely, whether both regimes share the same long run growth rate even if the short run dynamics appear to be different. If this is the case, the changes in regime affect only the short run dynamics but not long run growth.

4.5.1 Diagnostic tests for misspecification.

White (1987) extended Newey's (1985) conditional moment test of first order and formulated a general specification test based on the serial correlation of the score of the likelihood function. This type of test has also been applied by Hamilton (1996) to the markov switching regime model. It can be adapted easily to the case of switching regime models of the type considered here (see Norden and Vigfusson (1996)). If Θ is a $(n \times 1)$ vector of the parameters of interest in the likelihood function, the score, $\mathrm{s}_t(\Theta)$, is just the gradient of the likelihood function with respect to these parameters. Assuming ergodicity and no corner solutions, the gradient at the true parameter estimates, $\mathrm{s}_t(\Theta_{true})$, should be impossible to forecast using any information that was available in the previous period. That information obviously includes as well the lagged score, $\mathrm{s}_{t-1}(\Theta_{true})$. Since Θ is a $(n \times 1)$ vector, this yields $n \times n$ restrictions under the hypothesized lack of first order serial correlation in $\mathbf{s}_t(\Theta_{true})$ which can be tested. In practice, attention is constrained to restrictions that can be easily interpreted. A $(u \times 1)$ vector $\mathbf{q}_t(\Theta)$ is constructed with the u elements of the $\mathrm{s}_t(\Theta)\,\mathrm{s}_t(\Theta)'$ matrix that are of interest in each particular problem and then it is tested whether their expected value, given the information available, is equal to zero. Let $\hat{\Theta}$ be the maximum likelihood estimate of Θ, T the sample size and $\hat{\mathbf{A}}$ the row 2 and column 2 sub-block of the inverse of the following partition matrix

$$T^{-1}\left[\begin{array}{cc}\sum_{t=1}^{T}\mathbf{s}_t\left(\hat{\Theta}\right)\mathbf{s}_t\left(\hat{\Theta}\right)' & \sum_{t=1}^{T}\mathbf{s}_t\left(\hat{\Theta}\right)\mathbf{q}_t\left(\hat{\Theta}\right)' \\ \sum_{t=1}^{T}\mathbf{q}_t\left(\hat{\Theta}\right)\mathbf{s}_t\left(\hat{\Theta}\right)' & \sum_{t=1}^{T}\mathbf{q}_t\left(\hat{\Theta}\right)\mathbf{q}_t\left(\hat{\Theta}\right)'\end{array}\right]$$

Under the null hypothesis of no misspecification in the model, the product

$$T^{-1}\left[\sum_{t=1}^{T}\mathbf{q}_t\left(\hat{\Theta}\right)\right]'\hat{\mathbf{A}}\left[\sum_{t=1}^{T}\mathbf{q}_t\left(\hat{\Theta}\right)\right]$$

will have asymptotically a $\chi^2(u)$ distribution.

Any combination of parameters can be chosen to be used in this test. For a general test of misspecification in the switching regime considered here, the selected parameters are $\beta_{1,0}$, $\beta_{2,0}$, $\beta_{3,0}$, σ_1 and σ_2 (see equation (4.1)). Testing the gradients of $\beta_{1,0}$ and $\beta_{2,0}$ (if each equation has an intercept term) is analogous to testing for omitted serial correlation in the errors $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$. Detection of serial correlation in the gradients of $\beta_{1,0}$ and $\beta_{2,0}$ implies that there are sets of consecutive observations for which the constant should be different and this persistence effect translates into serial correlation in the residuals. Testing the gradient of σ_1 or the gradient of σ_2 is equivalent to testing for ARCH(1) regime specific effects because persistence will imply a changing variance over time. A test for the presence of Markov switching effects involves testing the gradient of $\beta_{3,0}$. Finding serial correlation in this case suggests that the regimes are more persistent than estimated and, if this is the case, the dependence of a regime on past regimes should be modelled as a markov switching regime model[55].

[55]Again there is a procedure available from the Bank of Canada which calculates these diagnostic tests for a switching regime model.

4.5.2 Nested testing procedures.

The switching regime model considered here encompasses many other models that might be of interest. One hypothesis that might need attention is whether the probability of the switch is constant or whether it varies with z_t. This is equivalent to testing whether $\beta_{3,1} = 0$ (constant switching probability) in equation (4.1). Another hypothesis that might be worth testing is whether the difference between the two regimes arises from some sort of heteroskedasticity, although the actual dynamics are the same. This is equivalent to test the joint null hypothesis of $\beta_{1,j} = \beta_{2,j}$ (and $\gamma_{1,j} = \gamma_{2,j}$ if the model to be tested is the one shown in equation (4.3)) for all j. If the null hypothesis cannot be rejected, then the model can be considered to be linear but with a pattern of heteroskedastic errors. Many other different hypotheses can be tested using different mixtures of parameter restrictions which will be imposed depending on the nature of the problem encountered.

Since all these models are nested within the switching regime model, a Likelihood Ratio test can be used. The Likelihood Ratio (LR) statistic is defined as

$$LR = 2\left[\log L\left(\Theta_{unrest}\right) - \log L\left(\Theta_{rest}\right)\right]$$

where $\log L\left(\Theta_{unrest}\right)$ is the natural logarithm of the likelihood function in equation (4.5) evaluated at the unrestricted parameter estimates Θ_{unrest}, while $\log L\left(\Theta_{rest}\right)$ is the natural logarithm of the likelihood function at the restricted parameter estimates Θ_{rest}. Under the null hypothesis of $\Theta_{unrest} = \Theta_{rest}$, the LR statistic has a χ^2distribution with degrees of freedom equal to the number of restrictions imposed.

4.5.3 Testing the equality of long run growth rates across regimes.

A potentially interesting question to address is whether the long run growth rates are the same in both regimes even if the short run dynamics are different. In this section, the long run growth rates will be calculated for the switching regression model in equation (4.3). It is straightforward to restrict this to the case of equation (4.1) where each of the regimes is simply an autoregression by setting all $\gamma_{r,j}$ to be equal to zero for $r = 1, 2$ and $j = 1, \ldots, \tilde{k}_r$.

The long run growth rates for country i in regime r for the switching regression model in equation (4.3) are calculated as

$$growth_{i,r} = \frac{\beta_{i,r,0} + \sum_{j=1}^{\tilde{k}_{i,r}} \left[\gamma_{i,r,j} \left(\frac{\sum_{t=1}^{T} \Delta \bar{y}_{i,t-j}}{T} \right) \right]}{1 - \sum_{j=1}^{k_{i,r}} \beta_{i,r,j}}$$

Thus, testing whether the long run growth rates for both regimes are equal is the same as testing whether

$$\frac{\beta_{i,1,0} + \sum_{j=1}^{\tilde{k}_{i,1}} \left[\gamma_{i,1,j} \left(\frac{\sum_{t=1}^{T} \Delta \bar{y}_{i,t-j}}{T} \right) \right]}{1 - \sum_{j=1}^{k_{i,1}} \beta_{i,1,j}} - \frac{\beta_{i,2,0} + \sum_{j=1}^{\tilde{k}_{i,2}} \left[\gamma_{i,2,j} \left(\frac{\sum_{t=1}^{T} \Delta \bar{y}_{i,t-j}}{T} \right) \right]}{1 - \sum_{j=1}^{k_{i,2}} \beta_{i,2,j}} = 0 \tag{4.7}$$

This restriction can also be written as

$$\begin{aligned} \varpi(\psi_i) = \; & \left(1 - \sum_{j=1}^{k_{i,2}} \beta_{i,2,j}\right) \left(\beta_{i,1,0} + \sum_{j=1}^{\tilde{k}_{i,1}} \left[\gamma_{i,1,j} \left(\frac{\sum_{t=1}^{T} \Delta \bar{y}_{i,t-j}}{T} \right) \right] \right) - \\ & \left(1 - \sum_{j=1}^{k_{i,1}} \beta_{i,1,j}\right) \left(\beta_{i,2,0} + \sum_{j=1}^{\tilde{k}_{i,2}} \left[\gamma_{i,2,j} \left(\frac{\sum_{t=1}^{T} \Delta \bar{y}_{i,t-j}}{T} \right) \right] \right) = 0 \end{aligned} \qquad r = 1, 2 \tag{4.8}$$

where ψ_i is the $\left[\left(2 + \sum_{j=1}^{2} \sum_{s=1}^{2} k_{i,s,j} \right) \times 1 \right]$ vector of the parameters of the model needed to compute the growth rate. The advantages of using the multiplicative formulation of the null hypothesis in equation (4.8) as opposed to the formulation in equation

(4.7) will be discussed later. For the moment, the problem is to test the hypothesis, $\varpi(\psi_i) = 0$, which involves a nonlinear function of the parameters. The test statistic used is

$$\frac{\varpi\left(\hat{\psi}_i\right)}{\sqrt{\text{estimated } Var\left[\varpi\left(\hat{\psi}_i\right)\right]}}$$

which is normally distributed. The estimated standard error can be found by using a first order Taylor's approximation to $\varpi\left(\hat{\psi}_i\right)$ around the true parameter vector ψ_i,

$$\varpi\left(\hat{\psi}_i\right) \approx \varpi(\psi_i) + \left(\frac{\partial\varpi(\psi_i)}{\partial\psi_i}\right)'\left(\hat{\psi}_i - \psi_i\right)$$

so that

$$Var\left[\varpi\left(\hat{\psi}_i\right)\right] \approx \left(\frac{\partial\varpi(\psi_i)}{\partial\psi_i}\right)' Var\left[\hat{\psi}_i\right]\left(\frac{\partial\varpi(\psi_i)}{\partial\psi_i}\right) \tag{4.9}$$

Since the derivatives in equation (4.9) are functions of the unknown population parameters, their sample estimates are used to compute them.

In the present case, the Wald test is very convenient since estimation of the restricted model is unnecessary. Asymptotically in a Wald test, the way a hypothesis is formulated under the null hypothesis is not important. However, in finite samples, the test statistic of the Wald test is not invariant to the algebraic representation of the hypothesis. Since the Wald test is derived from a Taylor series expansion, different but analytically equivalent forms of a nonlinear expression lead to actual differences in their respective Taylor series. This implies that in practice, the test statistics corresponding to the null hypotheses in equation (4.7) and (4.8) have different numerical values, even though they are testing the same restriction. Gregory and Veall (1985, 1987) presented some Monte Carlo evidence suggesting that in problems like the present case, the multiplicative form in equation (4.8) is more reliable than that in equation (4.7). Phillips and Park (1988) also studied this problem of the Wald statistic by using asymptotic Edgeworth expansions of the distribution of the Wald statistic. They showed analytically that if the denominator in equation (4.7) is small (as might be the present case), the multiplicative form in equation (4.8) would give a much better approximation to the asymptotic value of the Wald statistic. Hence, to test the equality of the long run growth rates in both regimes, the null hypothesis of equation (4.8) will be employed since it is more reliable than that of equation (4.7).

4.6 Impulse response functions in the analysis of output growth.

In any study of output growth, it is important to assess the effects of shocks to the economy. If a sudden random shock to output dies out eventually, so that the shock is not "persistent", then the economy would return back to its normal position. However, if any shocks to output do not die out, that is, they are persistent, the economy will not go back to this position. Additionally, even if a random shock is not persistent the speed and path of adjustment are also very important. If a shock is not persistent but it dies out very slowly, that is has a low speed of adjustment, it would take the economy many years to go back to its original position. Also, a shock may not be persistent but generate large responses or even large swings in the responses while the adjustment process is taking place. Such large responses might be undesirable. An impulse response function measures the effect of a shock to a certain data series over time. For that reason, impulse responses are extremely important in decision-making contexts where the whole time profile of the effect of a shock on a variable can be important, influencing the timing and magnitude of subsequent responses for example.

Impulse responses are derived as the difference between two different realizations of the variable of interest over a certain specified number of periods or horizon. It is assumed that for one of the realizations, the system is only hit by one shock, V_t, at time t. For the other realization it is assumed that the system is not hit by any shocks at all. Thus, the impulse response function measures the effect of a shock, V_t, which hits the system at time t, compared to when the system is not subjected to any shocks over the subsequent periods. Early studies on output fluctuations concentrated on linear models which impose symmetry of responses (see, for example, Cochrane (1988), Demery and Duck (1992) and Cogley (1990)). For example, in linear models, the persistence of positive and negative shocks are restricted to be the same and their response proportional to the size of the initial shock. Furthermore the effects of shocks are not allowed to vary over the business cycle, that is their responses are restricted to be the same no matter where in the business cycle (either a peak or a trough) the economy is. If this symmetry restrictions are not accepted by the data, but they are imposed anyway, they could bias the measures of persistence. Koop, Pesaran and Potter (1996) showed that in a multivariate nonlinear model, the impulse response functions are history- , shock- and composition-dependent. They are *history-dependent* because, in general, the impulse response functions calculated for two different initial values are different. For example, in the switching regime model considered here, one can choose as initial values two different points in time; one in each regime. For a nonlinear model in general, and in particular for the switching regime model considered here, the impulse response functions will be different, since the two regimes have different dynamics. The impulse response functions are also *shock-dependent* in nonlinear models. When calculating the impulse response function for a linear model, the size of the shock does not affect the transmission mechanism but merely scales the measure. This implies that the impulse response functions of linear models for positive and negative shocks are symmetric and, also, proportional to the size of the shock. However, this is not true in general in nonlinear models. It is worth mentioning here that in the switching regressions model, if one conditions on the regime, since each regime is linear, then the impulse response functions, after conditioning on one of the regimes, will be shock independent. The *composition effect* is relevant in multivariate models, both linear and nonlinear. It highlights the fact that the shocks to different equations of the system are correlated and, therefore, it is not appropriate to shock one equation while keeping the shocks to the rest of the equations set to zero. Although impulse responses of this type can be calculated, this is not what it is typically observed so it provides little information about the consequences of shocks. A solution proposed for this problem is to orthogonalise the covariance matrix of the shocks, but, in general, there is no unique transformation and, as a consequence, this method has been criticised as being arbitrary.

The three points raised above mean that the impulse response function is in fact a random variable and is best investigated using what Koop, Pesaran and Potter (1996) termed "Generalized Impulse Response Function (GIR)". The GIR is defined as

$$GI_Y\left(d, V_t, \Omega_{t-1}\right) = E\left[Y_{t+d} | V_t, \Omega_{t-1}\right] - E\left[Y_{t+d} | \Omega_{t-1}\right]$$

where Y_t is a random vector, d is the time horizon, V_t is a vector of random disturbances, $\Omega_{t-1}$is the set of information used to forecast Y_t and $E\left[\bullet\right]$ is the expectation operator. The dispersion of the unconditional GIR measures the long-run persistence of a series. For example, if the series is stationary, then the GIR random vector will converge towards a vector of zeros as time goes to infinity. If the series are random walks, the dispersion stays constant as the time horizon increases. However, to determine whether positive and negative shocks have different persistence, conditional versions of the GIR have to be studied; for example, one might condition on a particular shock and/or on a particular history or subset of the history. For example, in their work on US output growth, Beaudry and Koop (1993) found that the effects of positive shocks to output are highly persistent while negative shocks tend to be temporary. In their study of US output growth, Pesaran and Potter (1997) found that in recessions, negative shocks tend to be inhibited and sometimes reversed but they also found the reverse asymmetry in periods of normal growth and in expansions.

The impulse responses which will be calculated in chapters 5 and 6 are based on Koop, Pesaran and Potter (1996) and adapted for the switching regressions model shown in equation (4.1). The problem here is that the regime is not actually observed. Instead, for each observation, the *a priori* probability of being in a particular regime is known. This means that, until shocks are generated for the switching equation, it is not possible to know in which regime that particular observation actually lies. Additionally, it is also necessary to estimate a model to forecast z_t, the exogenous variable in the switching equation, so that, at each point in time, the actual regime in which an observation lies can be established.

The next section provides details of the computation of the impulse responses for the switching regime model which will be used in the subsequent empirical chapters.

4.6.1 Monte Carlo computation of the impulse responses in a switching regressions model.

There are several techniques available to compute the conditional expectation required for the computation of the impulse responses. One of them is Monte Carlo integration. This involves drawing a large number of innovations from the distribution of the residuals for each history. Using these innovations, the selected history and the estimated model, a realization of the time series can then be computed at each point in time for a certain number of periods. The conditional expectation can then be calculated by averaging across realizations. The discussion below ignores the country subscripts, i, for simplicity of exposition.

In the case of the switching regressions model considered here, a model to forecast the exogenous variable in the switching equation (4.1), that is z_t, is needed to be able to compute the impulse responses. Details of the two models used for forecasting purposes are found in chapters 5 and 6. In this discussion, let us assume that forecasting of z_t is carried out by the following regression

$$z_t = g(\bullet) + \sigma_4 \varepsilon_{4t} \tag{4.10}$$

where $g(\bullet)$ is a linear function of some variables and ε_{4t} has the usual normal distribution with zero mean and variance equal to 1.

The step-by-step description of the algorithm is found below.

Step 1. Pick a history (ω_{t-1}) from the observed values of the time series. Either randomly draw a shock (υ_t) from the joint density of the errors or if the interest is on the estimation of the impulse responses conditional on an specific shock, set υ_t equal to this shock.

Step 2. For a given horizon T, randomly sample $(D+1)\times R$ values of the 4 innovations ε_{1t}, ε_{2t}, ε_{3t} and ε_{4t}.

Step 3. Compute $\triangle y^0_{t+d}(\upsilon_t, \omega_{t-1})$ for $d = 1, 2, \ldots, D$ using the first D random shocks generated in step 2, iterating on the model in equation (4.1) from the initial conditions υ_t and ω_{t-1}.

The random shocks ε_{3t+d} together with the values of z_{t+d} and the estimated parameters for the switching equation $\beta_{3,0}$ and $\beta_{3,1}$ determine the values of $\left(y^*_{t+d}\right)^0$. This variable classifies the dependent variable $\triangle y^0_{t+d}$ either into regime 1 if $\left(y^*_{t+d}\right)^0 < 0$ or into regime 2 otherwise. Once the regime has been determined $\triangle y^0_{t+d}$ can be computed by using the parameters and shocks corresponding to the appropriate regime. The random shocks ε_{4t+d} together with the forecasting equation (4.10) can be used in turn to forecast the value of the exogenous variable in subsequent periods.

Step 4. Compute $\triangle y^0_{t+d}(\omega_{t-1})$ for $d = 0, 1, \ldots, D$ using the D random shocks used in step3 plus one additional shock, iterating on the model from the initial condition ω_{t-1},using the same procedure as in *Step 3.*

Step 5. Repeat steps 3 and 4 R times and form the averages for each individual component

$$\begin{aligned}
\overline{\triangle y}_{R,t+d}(\upsilon_t, \omega_{t-1}) &= \tfrac{1}{R}\sum_{j=0}^{R-1} \triangle y^j_{t+d}(\upsilon_t, \omega_{t-1}) \quad && d = 1, 2, \ldots, D \\
\overline{\triangle y}_{R,t+d}(\omega_{t-1}) &= \tfrac{1}{R}\sum_{j=0}^{R-1} \triangle y^j_{t+d}(\omega_{t-1}) && d = 1, 2, \ldots, D
\end{aligned}$$

By the Law of Large Numbers, the averages of this Monte Carlo replications will converge to the conditional expectations $E\left[Y_{t+d}|v_t, \omega_{t-1}\right]$ and $E\left[Y_{t+d}|\omega_{t-1}\right]$ as $R \to \infty$.

Step 6. Take the difference between the two averages to form a Monte Carlo estimate of the GI, where $\overline{\triangle y_t}\left(v_t, \omega_{t-1}\right)$ is just one realization.

Step 7. Repeat these steps a sufficient number of times to be able to estimate precisely the features of interest of the GI random vector.

In this section, it has been discussed how important both the persistence of shocks to output, as well as the adjustment path of output after a shock are. Since the switching regressions model is a nonlinear model, impulse responses should no be computed in the standard way which is applicable to linear models. It has been shown how the Generalised Impulse Response functions of Koop, Pesaran and Potter (1996) can be adapted to the switching regressions model under study here. These will be applied in chapters 5 and 6 to the models estimated there.

4.7 Conclusions.

In this chapter, some of the econometric issues involved in estimating switching regime models have been outlined. Additionally, several concepts which will be used in subsequent work have been defined. The discussion noted that many researchers have found problems when trying to estimate switching regressions models using maximum likelihood methods. One of these problems is the unboundness of the likelihood function that occurs for certain parameter values when the standard deviation of one of the regressions tends to zero (Maddala and Nelson (1975) and Maddala(1983)). In practice, this means that the algorithm used for maximization might generate incessantly higher values of the likelihood function as the standard deviation of one of the two regimes tends to zero. If this is the case, trying to find the global maximum leads to parameter estimates that are inconsistent. A local maximum (away from the unbounded region) will give consistent estimates of the parameter values in cases like this. In practice, however, the likelihood function is usually multimodal when attention is restricted to the bounded region of the likelihood. A problem then arises which is not inherent in the likelihood function but results from the way conventional algorithms find the roots of a function. Initial conditions play a very important role in the parameter estimation when conventional algorithms are used. Specifically, these algorithms typically stop at the first maximum found. This will be the closest to the initial parameter values supplied to the algorithm but there is obviously no guarantee that this is in fact the global optimum point.

To try to overcome these problems of the likelihood function, it has been argued that a global optimisation algorithm, the Simulated Annealing algorithm, should be used to maximize the likelihood function. Being a global optimisation algorithm it avoids the problem of multimodality. It is also very useful for overcoming the problem of unboundness of the likelihood function. Since it is very easy to spot the unbounded region of the likelihood function from the intermediate output of the algorithm, it is, therefore, easy to restrict the parameter values to avoid this region.

In chapters 5 and 6, two methods will be used to estimate different switching regressions models. First, a combination of the Expectation-Maximization method and maximum likelihood will be employed and second the simulated annealing algorithm will be used to maximize the likelihood function. The first method obviously involves searching through different parameter values to try to locate the value of the parameters that maximizes the likelihood function. These two methods will be then compared and the advantages and disadvantages of each one assessed.

In this chapter, several testing procedures that will be used in the next two empirical chapters have also been described in detail. A set of these testing procedures relate to issues common to all switching regime models, and include tests for the presence of autocorrelation and heteroskedasticity in each of the regimes, as well as tests for omitted Markov effects. All these tests are very important in assessing the adequacy of the estimated model. Testing procedures relating to questions specific to the applied work in chapters 5 and 6 have also been described in detail.

Finally, in the last section of the chapter, the focus has been on the importance of analysing the responses of the system to shocks once the model has been estimated. Since the switching regressions model is a nonlinear model, the standard computation of impulse responses is not appropriate. For this reason, the Generalized Impulse Response functions for nonlinear models of Koop, Pesaran and Potter (1996) have been adapted to the switching regressions model studied here.

In the next two chapters, these modelling techniques will be applied to two important economic questions. Chapter 5 studies the implications of current account deficits on both, the short run dynamics of growth and long run growth. Chapter 6 concentrates on the role of technology spillovers across countries in the process of growth.

Chapter 5

Growth dynamics and the balance of payments constraint; a switching regressions analysis of output growth in the G7.

5.1 Introduction.

The recent literature on output growth has concentrated mainly on the long run characteristics of growth but not enough attention has been paid to understanding the process by which growth is achieved. In chapter 3, it was shown that countries do not seem to be converging towards a common steady state. However, lack of convergence does not necessarily translate into divergence. The aim of the present chapter and that of chapter 6 is to try to shed light into the process of growth by incorporating interdependencies across countries. These links across countries will keep them together so that, even though there is no convergence, in the long run there is no divergence either. To be able to analyse these interdependencies in detail, attention is restricted in both chapters to a smaller group of countries, namely, the G7 countries.

For the purpose of studying long run growth the recent empirical literature has generally used linear models (see chapter 2 for a review of this empirical literature). However, linear models are problematic because implicitly they place many symmetry restrictions that if untrue could bias, for example, the measures of the persistence of shocks to output. Lately there has been a switch towards nonlinear models in the study of output growth which allows for potentially more realistic dynamics than linear models, but the focus has been on studying output on its own. These studies have been reviewed in detail in chapter 3.

In chapter 2 the advantages of analysing growth within a more realistic theoretical framework were stressed, that is, the importance of including other macroeconomic variables that are relevant for the process of growth. The present chapter deals with the econometric estimation of one of the theoretical models illustrated in chapter 2, namely that of section 2.4.1 using the nonlinear techniques highlighted in chapter 4. The model is that of an open economy with imperfect capital markets. Countries are allowed to borrow from abroad, however the maximum amount is restricted to some proportion of the country's output. In an open economy world, it would be rational for countries to borrow capital from abroad when they lack enough capital to achieve the optimum growth rate. Borrowing capital translates into a deficit in the current account of the balance of payments. However, the capital needed to attain this growth rate might exceed the maximum amount that other countries are prepared to lend and this can be regarded as a balance of payments constraint on growth. In an intertemporal framework, growing deficit of the current account of the balance of payments in relation to output in an economy increases the likelihood of a period of constrained growth. In this case the growth of output might be generated by two different processes which achieve the same long run growth rate of output but with very different dynamics depending on the accumulation of surpluses or deficits relative to output. This behaviour can be portrayed by using a "switching regressions" model of the type described in chapter 4 with an accumulation of current account surpluses/deficits over output as the switching variable. Chapter 2, section 2.4.2 also illustrates the role of technology spillovers across countries in the growth process. A more detailed investigation of this issue is left for chapter 6. However, the present chapter tries to include this complementary effect in addition to the balance of payments constraint on growth. This is accomplished by the inclusion of the mean growth rate outside the country under study. This variable is taken as a proxy for the level of technology in the rest of the countries. A high mean growth rate outside the economy could signal a higher level of technology which the economy could achieve if there are spillovers of technology across countries.

The remainder of the chapter is distributed as follows. In section 5.2 an overview of the data including the examination of the order of integration of the series and some preliminary linear estimation is provided. This linear estimation is presented as a benchmark of comparison with the switching regressions model which will be estimated subsequently, although no formal test will be employed. Additionally, in this section, the impulse responses obtained from these linear models will also be presented and analysed.

Section 5.3 concentrates on the nonlinear econometric analysis. To begin with, this section compares traditional optimisation methods with the Simulated Annealing algorithm for the maximization of the likelihood function in the switching regressions model. In chapter 4, it was discussed in detail some of the problems that may arise in the empirical estimation of these models due to the structure of the likelihood function. It is, therefore, of most importance to assess how frequently these problems arise to be able to judge the practical advantages of tools like the Simulated Annealing algorithm. Once the usefulness of this global optimisation algorithm is shown, the rest of the section gives details of the country by country nonlinear econometric analysis of the balance of payments constraint on growth using a combination of maximum likelihood and the Simulated Annealing algorithm. In this section it will be shown that for all the countries under study here apart from France a linear model is too restrictive when trying to model the process of output growth.

Section 5.4 concentrates on the responses of output in the G7 countries to different shocks. In this section, the impulse responses obtained for each country from the switching regressions models of section 5.3 are analysed and compared to those obtained in section 5.2 for the linear models. To obtain the responses of output to shocks using the switching regressions model, the adapted Generalized Impulse Response functions of chapter 4 are used since the impulse responses are history, shock and composition dependent in nonlinear models (see chapter 4). Finally section 5.5 concludes.

5.2 Output growth and the balance of payments in the G7 countries, 1970q1-1994q4.

In the empirical work of this chapter, attention is restricted to a smaller group of countries than that of chapter 3, the G7 countries. This allows for a much more detailed investigation of the process of growth. Additionally, the required length of time series data for a nonlinear study of this type can be found for these countries.

Quarterly data from the **International Financial Statistics** published by the International Monetary Fund is used in this analysis. Output data for the USA, UK, France, Italy and Canada refer to GDP in constant prices while output data for West Germany and Japan refers to GNP in constant prices, all of them seasonally adjusted and in each country's currency. Data on the current account is however not seasonally adjusted and is in US dollars for all the countries. To convert the output data for all the countries into the same currency (US dollars) the exchange rate (market rate) is used, also obtained from the **International Financial Statistics**. The data run from the first quarter of 1970 to the last quarter of 1994. A longer sample period could not be obtained because quarterly data on the current account is not available before 1970 and a consistent series could not be found for West Germany after 1994 due to German reunification.

Using these two series of data, two variables will be constructed which will in turn be used for estimation purposes. Let $y_{i,t}$ be the logarithm of country i's output at time t. It follows then, that the series $\Delta y_{i,t}$ denotes output growth at time t for country i. An additional series needs to be constructed, that is, the variable $z_{i,t}$ in equation (4.3). Following the discussion in the introduction, in this chapter $z_{i,t}$ is constructed as the sum of the current account/output ratio from time $t-8$ to time $t-1$ and it measures, therefore, the extent of the accumulated deficit or surplus in relation to output.[56] This results in a sample of 92 observations from the first quarter of 1972 to the last quarter of 1994.

Figures 5.1 and 5.2 plot both series, $\Delta y_{i,t}$ and $z_{i,t}$, respectively for each of the G7 countries. Additionally, Table 5.1 provides some descriptive statistics of the series for each of the countries. The mean growth rate of output is positive for all the G7 countries. The smallest average growth rate is that of the United Kingdom with 0.49% growth rate per quarter. The highest growth rate is 0.92% per quarter and corresponds to Japan. If individual periods for each country are looked at, the United Kingdom achieves the highest growth rate at 5.95% in just a quarter, although this country is also the one with the highest dispersion of growth around the mean. In the period considered here, the distribution of the growth rates across time in most of the G7 countries is skewed to the left and, therefore, growth above the mean value is more common than growth below. This is the case for the USA, France, Japan and to a lesser extent West Germany. The distribution of the United Kingdom is, however, the only one which is very skewed to the right, signalling the presence of growth rates consistently under

[56]The decision of using 8 lags to construct $z_{i,t}$ is completely ad hoc. However, in the econometric work of subsequent sections we tried to assess the robustness of the results by using 4 and 12 lags. This, however, did not change the conclusions of the analysis.

the average. The distributions of Italy and Canada are quite symmetric. If turning the attention to the accumulation over two years of the current account/output ratio, it is easily seen that the first significant feature of this variable is that the average is only positive for two countries, West Germany and Japan. The lowest average is that of Canada and the highest corresponds to Japan. Looking at specific periods for individual countries, Canada features the worst accumulated deficit in any one period whereas West Germany, followed closely by Japan, depicts the highest accumulated surplus. The dispersion of this variable across time is quite similar for the different countries apart from France which shows the lowest dispersion of all. In all, west Germany and Japan run consistently high surpluses, whereas Canada portrays large and consistent deficits.

The next two subsections carry out some preliminary tests and estimation to get a broad characterization of the data. Section 5.2.1 tries to establish the order of integration of the series. For this purpose, three different testing procedures will be employed: Augmented Dickey-Fuller, Kwiatkowski, Phillips, Schmidt and Shin's test and the "t-bar" test proposed by Im, Pesaran and Shin. Section 5.2.2 estimate linear models for each of the countries in the sample trying to capture in a linear setup the presence or absence of a balance of payments constraint on output growth.

5.2.1 Testing for Unit Roots and Cointegration.

This section tries to establish the order of integration of the two series used for each of the countries. First, a series of Augmented Dickey-Fuller (ADF) tests are carried out which test the null hypothesis of whether a series is integrated of order one. To complement this test, Kwiatkowski, Phillips, Schmidt and Shin's test (1992) will also be used. In this case the null hypothesis is stationarity, contrary to the null hypothesis of the ADF test. Both tests have been criticized for their lack of power and a different test, the "t-bar" test, has been proposed by Im, Pesaran and Shin (1995). This test is based on the values of the ADF statistic across countries and makes use of the panel structure of the data. This test will also be applied here since it is regarded as the most reassuring.

First, a series of ADF tests including an intercept and a linear trend, are applied to the logarithm of output data, $y_{i,t}$ (see Table 5.2). The number of augmentations in the test are chosen using the Akaike Information Criterion (AIC) out of a potential maximum number of lags of 4. The null hypothesis that the logarithm of output was integrated of order one, $I(1)$, cannot be rejected for any of the countries except for West Germany. All the test statistics (excluding West Germany) are in the interval $[-2.91, -1.81]$ with the 95% critical value being -3.4586. For West Germany the test statistic is only marginally significant, -3.4927, compared to a critical value of -3.4586. However, for this country, the optimal number of augmentations selected by the Schwarz Bayesian criterion is zero and completely different to the number selected by the AIC which is four. For any number of augmentations, except for four, the null hypothesis cannot be rejected. In particular, using 3 lags which is enough to eliminate any autocorrelation in the residuals the null hypothesis cannot be rejected with a test statistic of -3.1138 and the same critical value as before. The null hypothesis that the series is $I(2)$ is rejected for all the countries, the test statistics lie in the interval $[-10.14, -3.72]$, thus concluding that the output series for all the countries are $I(1)$.

To complement the ADF test a second test was applied following Kwiatkowski, Phillips, Schmidt and Shin. (1992). In this case, the null hypothesis is stationarity around a deterministic trend for the output series, y_{it}. It is a one sided (upper tail) LM test which uses the following statistic

$$LM = \frac{T^{-2}\sum\limits_{t=1}^{T}\left(\sum\limits_{j=1}^{t} e_j\right)^2}{T^{-1}\sum\limits_{t=1}^{T} e_t^2 + 2T^{-1}\sum\limits_{j=1}^{\varrho}\left[\left(1-\frac{j}{\varrho+1}\right)\sum\limits_{t=j+1}^{T} e_t e_{t-j}\right]}$$

where e_t are the residuals from the regression of the series on an intercept and a time trend, given that the null hypothesis is trend stationarity, and ϱ is the truncation parameter. Note that if the truncation parameter (ϱ) is chosen to be equal to zero, then the denominator of this statistic is reduced to the sum of the squared residuals divided by T. The problem of this test is that it is sensitive to the choice of the truncation parameter and the statistic decreases as the truncation parameter increases. The values of the test statistics for the logarithm of output, $y_{i,t}$, for values of the truncation parameter, ϱ, from 0 to 5 are reproduced in Table 5.3. For the USA and Japan, the null hypothesis of stationarity around a deterministic trend can only be rejected for low values of the truncation parameter ($\varrho = 0$ and 1). For West Germany and the UK, the maximum truncation parameters to be able to reject the null hypothesis are $\varrho = 3$ and $\varrho = 4$ respectively. For the rest of the countries, the null hypothesis is still rejected with the truncation parameter set to 5. In fact, for both Italy and Canada, the null hypothesis is rejected even for $\varrho = 8$, with the value of the test statistic being 0.1817 and 0.1581 respectively. Therefore, taking into account both tests, there seems to be enough evidence to regard the logarithm of output as being integrated of order one.

The same set of tests were applied to $z_{i,t}$. In this case, the regressions included an intercept but no trend. The results of the tests are not as clear as in the case of output.

Using the ADF procedure, the null hypothesis of $I(1)$ is rejected for $z_{i,t}$ in the UK, France and Italy (see Table 5.2) with test statistics -2.9532, -3.6812 and -4.9746 respectively and a 95% critical value equal to -2.8947. Using the LM test (see Table 5.4), the null hypothesis of stationarity is rejected for the USA, Canada and Japan even for a truncation parameter equal to 5. In fact, even with $\varrho = 8$, the null hypothesis cannot be rejected for these 3 countries (the test statistics are 0.6145, 0.5538 and 0.5492 respectively). This seems to indicate that this variable is not stationary for the USA, Canada and Japan. Economic reasoning would lead one to believe that the series should be stationary since countries cannot accumulate debt forever, but it could be that this effect would only show up in a longer run than the sample covers. Coackley, Kulasi and Smith (1996) showed that the solvency constraint implies that the balance of payments as a share of GDP should be stationary.

A potential problem is that the two tests applied here have low power and could fail to reject the unit root hypothesis even when it is false and the series is stationary. The panel structure of the data can be used to apply the "t-bar test" proposed by Im, Pesaran and Shin (1995) to test for unit roots in panels. This test is based on the average value of the augmented Dickey-Fuller statistic obtained across countries. The t-bar statistic is calculated as

$$\bar{t}_{PT} = \frac{\frac{1}{P}\sum_{i=1}^{P} t_{iT}\left(\varkappa_i, \hat{\rho}_i\right) - \frac{1}{P}\sum_{i=1}^{P} E\left[t_T\left(\varkappa_i, 0\right)\right]}{\sqrt{\frac{1}{P^2}\sum_{i=1}^{P} V\left[t_T\left(\varkappa_i, 0\right)\right]}}$$

where P is the number of countries in the sample, t_{pT} is the t-value from the ADF with the number of augmentations equal to $\varkappa$ (this is chosen as before using AIC). The critical values of this statistic are tabulated in Im, Pesaran and Shin (1995). The null hypothesis that the logarithm of output is integrated of order 1 cannot be rejected since the t-bar calculated statistic is -1.4825 compared to a 95% critical value of -2.66. However, the null hypothesis that the order of integration is equal to 2 is definitely rejected with a t-bar statistic equal to -14.3160 (the critical value remains the same). The conclusion, therefore, is that the logarithm of output is integrated of order 1. The same test is applied to $z_{i,t}$, the sum of the current account/output ratio. In this case, the null hypothesis is rejected in favour of the alterative of stationarity around a level with a t-bar statistic of -4.1604 and a 5% critical value of -2.05. Thus, stationarity will be assumed for the variable $z_{i,t}$.

5.2.2 Output growth regression models; a linear analysis.

This section concentrates on estimating linear models for each country which capture the two ideas mentioned in the introduction, namely, the presence of a balance of payments constraint on growth and the presence of technology spillovers as measured by the average growth rate outside the country under study. These linear models will serve as a point of reference when the nonlinear econometric analysis is carried out in the next section.

Since the output series are all integrated of order one, the next logical step is to determine whether a cointegrating relationship can be found among the output series of the G7 countries. First, an unrestricted vector autoregression (VAR) of order 4 was estimated and using AIC, 2 lags were selected for the order of the VAR. Cointegration tests using Johansen's maximum likelihood procedure (see Table 5.5) signalled the presence of one cointegrating relationship between the logarithm of output among the

G7 countries. The presence of a cointegrating vector among the output series of the G7 countries is very interesting because it supports the idea of no divergence. Even if countries do not seem to be converging to the same steady state as it was shown in chapter 3, the presence of a cointegrating vector across the output series suggests that there is something keeping the countries' output from diverging in the long run. The estimated cointegrating vector subject to a just identifying restriction is shown on Table 5.6. The coefficient for Canada is not significant so this over-identifying restriction is imposed and accepted with a likelihood ratio statistic of 0.55013 (95% critical value: 3.841). However, it is very difficult to give an economic interpretation to this cointegrating vector although something along the lines of Lee (1997) could be a plausible explanation, that is, a shock to output in one country will affect output in other countries if the shock causes balance of payments disequilibria. However, the coefficients of the cointegrating vector are found to be very sensitive to small changes in the sample size. Since a cointegrating vector is a long run relationship, significant changes to the parameters of this vector are not expected with minor alterations to the sample size. Since this is not the case, this issue in its present form is not pursued any further but it is important to bear in mind its presence.

Following this line of thought, that is, no divergence of the output series for the G7 countries in the long run, the two ideas discussed in the introduction will be incorporated in a VAR for the output growth series for each country. On one hand, the linear model will be augmented with the accumulated current account/output ratio variable, $z_{i,t}$. This variable will identify whether accumulated deficits constrain growth. It is important to stress again that $z_{i,t}$ is an accumulation from time $t-8$ to time $t-1$ and it is, therefore not determined contemporaneously with output growth at time t. This variable will work in a similar fashion to an error correction mechanism. On the other hand, the model also tries to incorporate the existence of technology spillovers across countries by augmenting the linear models with the mean growth outside the country under study, $\Delta\bar{y}_{i,t-j}$ following Lee (1997). This variable captures the feedback effects between the growth of output in different countries. Output outside country i was already defined in chapter 4 as

$$\bar{y}_{i,t-j} = \frac{1}{P-1} \sum_{p=1,p\neq i}^{P} y_{p,t-j}$$

where P is the number of countries in the sample. Using AIC to select the number of lags, the following linear regressions are estimated

$$\Delta y_{i,t} = \beta_{i,0} + \sum_{j=1}^{k_{i,1}} \beta_{i,,j}\Delta y_{i,t-j} + \sum_{j=1}^{k_{i,2}} \gamma_{i,,j}\Delta\bar{y}_{i,t-j} + \delta_i z_{i,t} + \varepsilon_{i,t} \tag{5.1}$$

for each country i, and then the joint significance of the $\Delta\bar{y}_{i,t-j}$'s is tested (see Table 5.7). One problem that seems to be common to most of these regressions is the failure of the normality test. However, inspection of the residuals reveal that the failure of the normality test is due to one very large error (two in the case of the United Kingdom). It is interesting to note that for Japan and for two of the three European countries for which the regressions fail the normality test (France and Italy), these large negative errors occur in 1974, just after the first oil price shock. The united Kingdom is the only one with two large errors, both positive, one in 1973 (when oil was discovered in the North Sea) and the other one in 1979 (during the second oil price shock). The regression for the USA has one negative large error in 1980. The regressions for these countries were run again but using simple dummy variables in the observation corresponding to the large error. it was found that for all the countries, these dummies were highly significant, but their presence did not affect the other coefficients of the regressions. Furthermore, once these dummies are included, the null hypothesis that the errors have a normal distribution cannot be rejected at normal significance levels for any of the countries. The test statistics (significance in square brackets) are 3.2343[0.198], 3.0496[0.218], 4.1722[0.124], 4.3411[0.114] and 0.10509[0.949] for the USA, United Kingdom, France, Italy and Japan respectively. Additionally, the inclusion of the dummy variable for Italy also solves the problem of serial correlation for this country with a test statistic (significance level in square brackets) of 6.2062[0.184].

Following the discussion of the theoretical model in chapter 2, the coefficient of $z_{i,t}$ is expected to be positive if accumulated balance of payments difficulties constrain the growth of output. The coefficient on the balance of payments variable takes this sign for all the countries except for Japan although the coefficient for Japan is not significantly different from zero. Even though the coefficient of $z_{i,t}$ is positive for the USA, it is not significant either but this may be as expected. Since USA is quite a self sufficient country, the balance of payments variable defined here is perhaps less likely to be significant a priori. Also, Japan ran large current account surpluses and high growth for a lengthy period in the sample so it may be difficult to single out this sort of effect for Japan since, in this situation, it is hard to think of output growth in Japan as being constrained. For the remaining five countries, the accumulation of the current account/output ratio does seem to exert some influence on growth, although this variable is only significant at conventional significance levels for West Germany. Note that, in the model for the UK, neither output growth in the rest of the countries, or its own past growth seem to be significant. Strong evidence is found in five of the seven countries (USA, France, West Germany, Italy and Canada) for the inclusion of the lags of output growth in the rest of the countries strengthening the postulated importance of feedbacks across countries.

Figure 5.3 plots the impulse responses of the logarithm of output for 5 of the G7 countries. The countries excluded are the United kingdom and West Germany, since for these countries the response of output is just equal to the size of the shock. In section 5.4, these impulse responses will be compared to those of the switching regressions model of section 5.3. The horizon for the impulse responses is set to 20 periods, that is, 5 years. The impulse responses level off before reaching the horizon considered here for all the countries except for Japan, for which the response is still increasing after 20 periods (5 years) although at a decreasing rate. In fact, the response of the logarithm of output for Japan levels off after 25 periods, marginally over 6 years. The responses for Italy and Canada level off after just 4 periods which corresponds to one year. The response of the shock for the USA levels off somewhat later, after 5 periods. The full impact of the shock to the logarithm of output for France is reached after approximately 2 years, that is 8 periods. In the long run, the responses to the shocks are magnified for all the countries in Figure 5.3. For the USA, France, Italy and Canada the responses are magnified by factors of 1.3328, 1.2314, 1.1700 and 1.2424 respectively. However, Japan is different in that the response to a shock more than doubles in the long run, specifically, the response is magnified by a factor of 2.1330 after 25 periods. Therefore, Japan seems to be the country for which the effects of a shock adjust slower and have a bigger impact in the long run.

5.3 Output growth models; a nonlinear analysis incorporating balance of payments constraints.

The models estimated in the last section include interdependencies across countries and have already quite sophisticated dynamics due to the inclusion of the accumulation of the current account of the balance of payments and output growth elsewhere. However, these models are linear and, therefore, implicitly impose certain restrictions on the model (see chapter 4). It is therefore important to use in this context nonlinear models of the type described in chapter 4, that is, switching regressions models. This type of models allow for potentially more sophisticated dynamics than linear models and, consequently, are less restrictive. The present section concentrates on the econometric estimation of switching regressions models for each of the G7 countries. However, allowing for more sophisticated dynamics comes with a cost; that is, the estimation of the switching regressions model is far from straightforward in practice. In chapter 4, the practical problems of estimation of this type of models with traditional optimization algorithms were discussed in detail. To overcome these problems, a global optimization algorithm was proposed. The problems discussed in chapter 4 are sample specific and it is, therefore, very important to assess the frequency with which they appear in different samples since the computation time when using a global optimization algorithm is higher. The use of data for seven different countries in this chapter is perfectly suited for that purpose. If these kind of problems are found when estimating the models for all of the countries, then, this gives a good indication that they are more common than it seems to be the consensus in the literature.

Section 5.3.1 introduces the switching regressions model which will be estimated in this chapter. Section 5.3.2 concentrates on the practical problems of estimation of the switching regressions model using traditional optimization algorithms. The performance of these algorithms are compared to the global optimization algorithm proposed in chapter 4, the Simulated Annealing algorithm. In this section, it will be shown that there are actually practical problems in the estimation of all countries even though for some countries there is no indication of problems when traditional optimization algorithms are used. Therefore, this gives a good indication of how important the use of the Simulated Annealing algorithm is when dealing with switching regressions models. Section 5.3.3 details the switching regressions models estimated for each of the G7 countries.

5.3.1 The form of the switching regressions model.

In this section, the empirical switching regressions model which will be estimated subsequently is presented. In view of the significance of output growth in the rest of the countries in the preliminary linear regressions, equation (4.3) is the final estimated nonlinear model. For clarity, this model is reproduced below.

$$\Delta y_{i,t} = \begin{cases} \beta_{i,1,0} + \sum_{j=1}^{k_{i,1}} \beta_{i,1,j} \Delta y_{i,t-j} + \sum_{j=1}^{\tilde{k}_{i,1}} \gamma_{i,1,j} \Delta \bar{y}_{i,t-j} + \varepsilon_{i,1,t} & \text{if } y_{i,t}^* < 0 \\ \beta_{i,2,0} + \sum_{j=1}^{k_{i,2}} \beta_{i,2,j} \Delta y_{i,t-j} + \sum_{j=1}^{\tilde{k}_{i,2}} \gamma_{i,2,j} \Delta \bar{y}_{i,t-j} + \varepsilon_{i,2,t} & \text{if } y_{i,t}^* \geq 0 \end{cases}$$

and

$$y_{i,t}^* = \beta_{i,3,0} + \beta_{i,3,1} z_{i,t} + \varepsilon_{i,3,t} \tag{5.2}$$

and with the errors following the joint distribution depicted in equation (4.2). The switching regressions model should uncover the balance of payments constraint effect that was discussed in chapter 2, that is, different dynamics but the same long run growth. To this end, the model will be estimated allowing not only for different dynamics but also for different long run growth rates. Once the model is estimated, their equality will be explicitly tested.

Assume for the moment that there might be a balance of payments constraint on output growth. As the accumulation of the current account of the balance of payments surplus (relative to output) increases, the probability of the economy being in the unconstrained growth regime increases. Similarly, as the deficit grows the probability of the economy being in the constrained growth regime rises. Assuming that $\beta_{i,3,1} < 0$ and ignoring for the moment $\beta_{i,3,0}$, an accumulated surplus in the current account/output ratio will make the probability of being in the first regime, $\Phi\left(-\beta_{i,3,1} z_{i,t}\right)$, higher than 0.5, so that the high growth regime should be the first, while the constrained growth regime should be the second.

The inclusion of the average growth rate of output in the rest of the countries tries to capture the presence of spillovers of technology across countries. If the rest of the countries are growing at a higher rate because of higher technology, then the presence of technology spillovers across countries will translate into higher growth in the country under study.

In switching regressions models, it is also interesting to look at the roots of the characteristic polynomials associated with each of the regimes since they will provide useful insights into their dynamic properties which will help to understand better the growth process. By looking at these roots, the random shocks are assumed away. However, at this stage, this gives an indication of the dynamics of the regimes. In the next section, the dynamic properties of the models estimated here for each country will be examined taking into account random shocks.

5.3.2 Comparison of traditional estimation approaches and the preferred approach in modelling G7 output growth.

This section focuses on the performance of traditional algorithms used to maximize the likelihood function of the switching regressions model versus the Simulated Annealing algorithm. In chapter 4 the potential difficulties when maximizing the likelihood function of these models were discussed in detail. The major difficulties outlined there were, on one hand, the presence of multiple maxima and, on the other hand, the unboundness of the likelihood function. The presence of several optima is an inconvenience when traditional algorithms are used to find the maximum of the likelihood function. This is because traditional algorithms are dependent on the initial values of the parameters used to start the algorithm. In practice, this means that they will converge to the local maxima which is closest to the values of the initial conditions supplied to the algorithm and stop there. However, this does not guarantee that the global maximum has been found. Consequently several runs of the algorithm with different initial conditions are needed in this case. Even after several runs, there is no guarantee that the global maximum has been located.

The second problem, the unboudness of the likelihood function, is sample specific. If there is a combination of parameter values which happen to fit one of the observations perfectly in one regime, then the likelihood becomes unbounded in the sense that higher and higher values of the likelihood function can be achieved by simply decreasing the variance in that regime. In this case, the maximum of the likelihood away from the unbounded region will give consistent estimates of the parameter values (see chapter 4). In cases of unbounded likelihood functions, the use of traditional optimization algorithms will produce one of two possible outcomes. If the initial conditions are near the unbounded region of the likelihood function, the algorithm will keep on trying to reduce the variance of one of the regimes. This process in theory could continue forever, however, in practice, due to the precision of the computer, the algorithm will eventually stop at a point in which the variance of one regime is very small. Nevertheless, these parameter estimates are not consistent (see chapter 4). It could also be the case that the initial conditions are close to one of the local maxima away from the unbounded region. As a result, the algorithm will converge to this point giving no indication of the unboundness of the likelihood or the presence of other local maxima away from the unbounded region.

A global optimization algorithm, the Simulated Annealing algorithm, was proposed in chapter 4 to overcome these two practical difficulties when maximizing the likelihood function of switching regressions models. The presence of several optima is not a problem since it is a global optimization algorithm and, therefore, independent of the initial conditions. This, in turn, implies that if the likelihood is unbounded, this algorithm will always converge to the unbounded solution. However, the advantage of this algorithm lies in that it is straightforward to restrict the parameter space so as to exclude the unbounded part of the likelihood function. Once the parameter space is restricted, convergence to the global maximum in this restricted space is guaranteed.

The aim of this section is to illustrate how useful the Simulated Annealing algorithm is in the context of switching regressions models. Since data is available for seven different countries, the frequency of the presence of unbounded likelihoods and several optima can be assessed in practice. It will be shown how common unbounded likelihoods are. In fact, the likelihoods for all seven countries used in this analysis are unbounded. The only reason why this does not appear to be the case when using traditional algorithms is because of their dependency on the initial conditions used. It will also be illustrated how difficult it is to find the global maximum with local optimization algorithms even after several runs with different initial conditions.

Model (5.2) was estimated for each country setting the maximum lag to four, $k_{i,1} = \tilde{k}_{i,1} = k_{i,2} = \tilde{k}_{i,2} = 4$. As a first attempt to find a maximum, the initial conditions were generated by splitting the sample into two groups according to whether the values of $\Delta y_{i,t}$ were positive or negative. Then using Ordinary Least Squares for the two subsamples, estimates of the parameters for the first regime (positive values of the dependent variable) and second regime (negative values) were obtained. The initial estimates of the parameters of the switching equation were obtained by regressing the cumulative distribution function of the standardized dependent variable, $\Delta y_{i,t}$, on a constant and the accumulation of the current account/output ratio, $z_{i,t}$. These initial conditions should work well for countries for which the mean growth rate is close to zero (see Table 5.1 for some summary statistics of the growth rates in the G7 countries). The estimation was carried out by using the EM algorithm to start with. When a certain convergence condition for this algorithm was met, the procedure was switched to maximum likelihood using Newton's algorithm.[57] Several problems were encountered. For France the algorithm would not converge and for Italy and Japan it was clear that the algorithm converged to an unbounded solution, that is the estimate of σ_2 in both cases was practically zero. However, for the rest of the countries, a maximum was found with reasonable parameter values and no indication of unboudness of the likelihood function. It was also noted that the EM algorithm was very quick to approach the solution.

[57]The convergence condition was usually set to 0.0001. In this case, the EM algorithm comes to a halt when, for every single parameter, the absolute value of the difference between the estimated parameters in two successive iterations is smaller than 0.0001 times the absolute value of the new parameter.

However, in cases of unbounded likelihood functions, this is an obvious disadvantage when trying to find the maximum away from this region because it jumps too quickly towards the unbounded solution. In view of such problems and as a trial, 100 random initial values between the bounds expected for each of the coefficients were generated and the estimation was carried out again but this time using only maximum likelihood (the EM algorithm was found to crash too often because it was jumping too quickly to an unbounded solution). Convergence to a solution with random initial values seemed to be very unlikely as was expected (as an example, in the trial for Italy only three had converged to a solution after 5000 iterations of the algorithm: the three of them stopped in 3 different points and one of them clearly being in the unbounded region of the likelihood). It was decided then to estimate for each country all the possible different combinations of lags using as initial conditions the Ordinary Least Squares estimates described earlier as initial conditions.[58] It was then apparent how the models' parameters were converging to different points, in the sense that they were converging to different local maxima. The estimated parameters were then placed into different groups according to their values and then all the different combinations of lags were estimated again, but this time, with the initial conditions corresponding to each one of the groups of the parameters identified under the last estimation.[59] This process continued until there were no more new initial conditions that could lead to a different local maximum. At this point, the global maximum was selected among all the local maxima found. However, it was obvious that for Italy and Japan this method was unable to

[58]Note that different number of lags will lead to slightly different initial conditions.

[59]At first sight it seems that the computations could be reduced by simply selecting only the set of parameters for which the likelihood is highest in each step. It was found that the highest likelihood in each step might be in the direction of an unbounded solution (obviously unbounded solutions as it was explained in Chapter 4 lead to a very high likelihood value) that only becomes apparent after the inclusion of more lags in the model, that is, possibly in the following estimation step.

find a local maximum and that all the possible solution candidates ended up pointing towards unbounded solutions.

To assess the usefulness of the simulated annealing algorithm in this context, the likelihood function is maximized once for all the countries using the algorithm.[60] Not surprisingly, when the algorithm was run, it was found that all seven countries had unbounded likelihoods. From inspection of the output it was clear that the likelihood function for all seven countries was being maximized by trying to make the variance of one of the regimes smaller and smaller at every step of the algorithm and therefore, trying to put all the observations in just one regime. However, in difficult cases like these is when the simulated annealing algorithm shows its potential value. From the intermediate output of the algorithm, it is easy to decide how to restrict the bounds of the parameters so as to exclude the points at which the likelihood became unbounded. When the values of the likelihood function obtained with this algorithm were compared to those obtained by using several runs of traditional algorithms it was found that the actual maximum had only been found for two countries, Canada and West Germany, although the differences in the likelihood and parameter values were small in the other three countries.[61] These estimated switching regressions models for each country will be presented in detail in the next section.

In conclusion, even though a run of the simulated annealing takes longer than one run of conventional algorithms, it is quicker overall because only two or three runs of

[60] The following parameter values are used for the simulated annealing algorithm (see the Appendix to Chapter 4 for a detailed description of this parameters): the initial temperature, T^0, is set to 1000, which for this particular dataset guarantees that the step vector, vm, is big enough to cover all the parameter space, the criterion for termination of the algorithm, ϵ, is set to 10^{-6}, the number of temperature reductions before testing for termination , N_ϵ, is equal to 4, the number of cycles before the step vector is adjusted, N_S, is equal to 20, the varying criterion, c, is set to 2 for all the parameters, the number of loops, N_T, is equal to 100, and the reduction coefficient, r_T, is set to 0.85.

[61] Recall that traditional algorithms were unable to find a suitable model for Italy and Japan.

the algorithm are needed in total.[62] When using conventional algorithms, it is important to run them several times with different initial conditions. In some cases, conventional algorithms will not converge at all giving no indication of the possible reason of this behaviour.

5.3.3 Estimated switching regressions models.

The following sections present the estimated switching regressions models for each of the G7 countries. Each country section presents a more detailed account of the estimated models and corresponding test procedures for each country. However, a summary of the main conclusions reached after this estimation exercise is presented first.

Firstly, there is clear evidence of the type of nonlinearities in output growth studied here, that is of distinct regimes, in all the G7 countries with the only exception of France. This is very important since the linear models of section 5.2.2 are unable to capture these effects. There seems to be little evidence that balance of payments deficits have constrained growth in the long run in the countries studied here. Only Canadian data supports this hypothesis. In the case of West Germany, the preferred model seems to indicate that very high and prolonged surpluses in the current account boost growth. Nevertheless, the growth rate in West Germany can hardly be regarded as having been constrained by balance of payments deficits since in the sample period under study here it had long periods of very high accumulated surpluses in comparison to the size

[62] However, sometimes this is not even the case. Every so often, one run of traditional optimisation methods is unable to converge before a run of the simulated annealing algorithm. Sometimes traditional algorithms can run for hours in situations of unbounded likelihoods trying to make the variance of one of the regiems smaller and smaller each time without converging.

of the accumulated deficits. However, there is evidence that for most of the other countries, the accumulation of current account deficits or surpluses over output exerts some influence on the short run dynamics of growth, France is, nevertheless, an exception. A possible explanation for this finding is indicated by the model of constrained borrowing presented in chapter 2 which points towards different dynamics depending on whether countries are constrained or not, but that the balance of payments does not represent a long run growth constraint. It could also be the case that higher deficits than the ones encountered in the sample are needed to find evidence of a long run constraint on growth or that the accumulation over two years fails to pick up the constraint that some other accumulation would single out.

On the other hand, there is wide evidence supporting the inclusion of lagged output growth in the rest of the countries in the model since they are always significant although not always present in all the regimes. This certainly highlights the importance of modelling these kind of interactions among countries and, therefore, these should be studied in more depth.

A more in depth analysis for each of the countries is given below. First, the unrestricted models with 4 lags are presented and some preliminary conclusions are drawn. These are the models obtained in the last section by maximizing the likelihood function using the Simulated Annealing algorithm. Having reported the unrestricted model, a "preferred model" for each of the countries is selected according to the Akaike Information Criterion (AIC) if there is no misspecification.[63] Several tests which were explained in detail in chapter 4, are then carried out in the preferred model for each

[63]Terasvirta (1994) favours AIC as a selection mechanism for nonlinear models.

country. These tests include diagnostic tests for the presence of AR and ARCH errors in each of the regimes, the presence of omitted Markov effects in the switching regression and a general test of misspecification. An important test in the switching regressions model framework is also included, namely, a test of the switching regressions model against a linear model. The linear model selected for this test is not, however, the linear model estimated in section 5.2.2 but a linear model which is nested within the switching regressions model, that is, a model with equal dynamics in both regimes but allowing for different variances in each of the regimes. Being unable to reject the restrictions implied by this model points towards a linear model but with some kind of heteroskedastic pattern in the error term. Additionally, a test which is of special interest in the present economic context is also applied, that is, the equality of the estimated growth rates in each of the regimes. This test will give an indication of whether there is a link between an accumulation of current account deficits with respect to output and the growth rate. The levels of significance associated with different test statistics are shown in square brackets. For the following sections the country subscript will be dropped as long as misinterpretation is not possible.

5.3.3.1 USA.

Table 5.8 presents the estimates of the unrestricted switching regressions model for the USA; that is, the model with k_1, k_2, $\tilde{k}_1$ and $\tilde{k}_2$ equal to four. This model passes the general misspecification test, although there is some evidence of AR effects of order one in the second regime with a test statistic of 3.902 $[0.0482]$. The estimated coefficient of the switching variable, that is the accumulation of current account deficits over output, is significant. However, the estimated growth in the first regime is lower than the estimated growth in the second regime, although they are not significantly different with a test statistic equal to -1.581.

In moving to the "preferred specification", AIC chooses a model with no intercept in the switching equation but with k_1, k_2, $\tilde{k}_1$ and $\tilde{k}_2$ still equal to four. This model is reported in Table 5.9. The restriction cannot be rejected with a likelihood ratio (LR) test of 0.2432 [0.6219], since the intercept of the switching equation is not significant in the unrestricted model. This model is well specified as it passes the joint test for misspecification as well as all the individual tests. The estimated coefficient of the switching variable, -4.0990, is significant with an associated probability equal to 0.0506.

Since this model is well specified and the coefficient of the accumulation of the current account over output is significant, the next logical step is to test this model against a linear model with equal dynamics in both regimes but different variances. Thus, the following restrictions are tested $\beta_{1,,j} = \beta_{2,j}, \gamma_{1,k} = \gamma_{2,k}$ and $\beta_{3,0} = 0$ for $j = 0, \ldots, 4$ and $k = 1, \ldots, 4$. This hypothesis is rejected with a LR test statistic equal to 21.30 [0.0191]. Thus, the linear model with heteroskedastic errors is discarded in favour of the switching regressions model shown in Table 5.9.

Additionally, the lags of output growth in the rest of the countries are jointly significant with a LR test of 25.73 [0.0023]. This fact supports the idea of the inclusion of output growth in the rest of the countries as a determinant of the growth process.

The estimated growth rate in what should be the high growth regime is equal to 0.30% per quarter and in the second regime is 0.83% per quarter, that is, they do not follow the economic theory, however, they are not significantly different at 5% with a statistic of -1.729. Thus, even though the growth rates attained in both regimes are not different, and hence, in the case of the USA no evidence is found of accumulated current account deficits constraining growth, there is a definite effect of this variable on the dynamics of output growth since the hypothesis of equal dynamics in both regimes is rejected.

It is interesting to note that the first regime has higher dispersion than the second, and, therefore, in the case of the USA periods of small accumulated current account over output surpluses are associated with more volatility in output growth. The probability of being in the first regime will be higher than the probability of being in the second regime if there is an accumulated surplus of the current account of the balance of payments. The probabilities of being in the first regime for each point in the sample are shown in Figure 5.4. However, since the growth rates in both regimes are not statistically significantly different these probabilities cannot be associated with periods of constrained and unconstrained growth. Nevertheless, they are important to identify the periods associated with a higher probability of larger shocks to output growth.

The time path of both regimes is stable. The first regime has a pair of complex roots and two real roots with opposite signs, both of them less than one in absolute value (see Table 5.9) which indicates a tendency of the growth rate of output to grow towards a mean value after a mixture of cycles and fluctuations due to the complex roots and due to the opposite signs of the real roots. The second regime has two pairs of complex roots showing cyclical behaviour.

5.3.3.2 United Kingdom.

The parameter estimates of the unrestricted switching regressions model for the UK are shown in Table 5.10. The coefficient of the switching variable is not significant a 5% but it is significant at 10%. The model, however, seems to be misspecified. The joint misspecification test has a test statistic of 13.290, therefore the null hypothesis of no misspecification is rejected with significance equal to 0.0208. The reason appears to be the presence of autoregressive (AR) effects of order one in the first regime. Additionally, the presence of AR effects of order one in the second regime is only just rejected at 5%.

Following the specification search rules mechanically, the model chosen by AIC as the "preferred model" has $k_1 = 1$ $k_2 = 2$, $\tilde{k}_1 = 0$ and $\tilde{k}_2 = 4$. The LR test of the restrictions is equal to 3.334 (significance equal to 0.9496). This model, however, is misspecified. The joint misspecification test has a test statistic equal to 13.77 and significance equal to 0.0171. Furthermore, the absence of AR effects of order one in the first regime is rejected with a test statistic of 4.121 and significance 0.0076. For that reason, the model is augmented with one more lag of output growth in the first regime. The parameter estimates of this model are presented in Table 5.11. The LR test of the restrictions from the unrestricted model is equal to 2.996[0.9346]. Although the model passes the joint misspecification test, the presence of AR effects of order one in the first regime cannot be ruled out at 5% significance level. The coefficient of the switching variable is significant. Additionally, the dynamics of both regimes are different and in fact the restriction of equal dynamics but different variances in both regimes is rejected.[64] Consequently there is evidence of two regimes governed by the accumulation of current account surpluses of deficits over output. The lags of output growth in the rest of the countries are not significant in the first regime. However, in

[64] Although the selected model for the UK has two lags of output growth in both regimes, the first regime has no lags of the average output growth in the rest of the countries, whereas the second regime has four lags. Therefore, the test of equality of dynamics but different variances is based on the restrictions needed to move from the unrestricted model to a model with equal dynamics in both regimes with two lags of output growth and four lags of the average output growth in the rest of the countries.

the second regime the four lags are significant and the test of the restriction of no lags of output growth in the rest of the countries in the second regime is rejected.

The estimated growth rates are, in the first regime 0.49% per quarter and in the second regime 0.44% per quarter. Therefore, the first regime has a slightly higher growth rate than the second regime, nevertheless, they are very similar in magnitude. In fact the hypothesis of equality of growth rates in both regimes cannot be rejected. Consequently, even though the size of the accumulated deficit influences the dynamics of growth, there is no evidence of a balance of payments constraint in the sense that the growth rates achieved in both regimes are the same.

The probability of being in the first regime is higher than the probability of being in the second regime if there is an accumulated surplus of the current account over output or if the accumulated deficit is less than 0.1859. Figure 5.5 shows the probabilities of being in regime one for each of the points in the sample. Again, these probabilities cannot be associated with periods of constrained or unconstrained growth, however, the probability of being in the first regime determines the probability of being in the regime in which the shocks to output have higher dispersion. It is very clear now, the reason why in the linear model estimated for the UK in section 5.2.2 the lags of output growth in the rest of the countries were insignificant. In the sample period, the UK was more likely to be in the first regime overall (see Figure 5.5). In this regime, the lags of output growth in the rest of the countries are insignificant (see Table 5.11). Thus, the linear model concentrates on the most likely dynamics, those of the first regime and, consequently, regards the lags of output growth in the rest of the countries as insignificant, even though it is clear from the obtained switching regressions model that these lags are very significant if there is a large accumulated deficit of the current account over output.

As is the case for the USA the time paths of both regimes are convergent. Both regimes have two real roots with modulus less than one (see Table 5.11). The first regime has two real roots with opposite signs which indicates an oscillating trajectory towards a mean value. The second regime has two positive real roots indicating a tendency of output growth to move towards a mean value.

5.3.3.3 France.

Table 5.12 depicts the parameter estimates of the unrestricted switching regression model for France. The first important difference between the model for France and the models for the USA and the UK is that the switching variable is not significant. Therefore, it seems that the accumulation of the current account over output does not affect neither the growth rates, nor the dynamics. This unrestricted model is misspecified, due mainly to the presence of ARCH effects of order 1 in the second regime.

The model selected by AIC has $k_1 = \tilde{k}_1 = \tilde{k}_2 = 4$, $k_2 = 2$ and no constant in the switching equation (see Table 5.13). The restrictions needed to get to this model from the unrestricted model cannot be rejected. Again, the accumulation of the current account over output does not significantly affect either the dynamics of output growth or growth rates. The model is, however, misspecified since the joint hypothesis of no misspecification is rejected. Closer inspection of the individual tests reveal the presence of AR and ARCH effects of order 1 in the second regime. Both regimes have four lags of output growth in the rest of the countries. The joint hypothesis of no lags of output growth in the rest of the countries, that is, $\gamma_{jk} = 0$ for $j = 1, 2$ and $k = 1, \ldots, 4$ and

$\beta_{23} = \beta_{24} = 0$ is rejected. Therefore, the importance of taking into account these type of interdependencies across countries is established.

However, in the case of France, the most important test to assess the validity of the switching regressions model is a linearity test since the switching regressions variable is not significant. It could be the case that there are in fact two different regimes but the switch between regimes is not governed by the accumulation of current account deficits and surpluses as it is hypothesized here. Alternatively, it could be the case that the model is in fact linear. The hypothesis of equal dynamics but different volatilities in the two regimes is, therefore, tested. In this case, it involves testing the following restrictions $\beta_{1,,j} = \beta_{2,j}, \gamma_{1,k} = \gamma_{2,k}$ for $j = 1, 2$ and $k = 1, \ldots, 4$.[65] In contrast to the preceding countries, this hypothesis cannot be rejected. The LR test statistic is equal to 7.267[0.6094]. This model with equal dynamics passes the joint misspecification test as well as all the individual tests. Furthermore, the coefficient of the accumulation of the current account over output is not significant. Therefore, the final model chosen for France is shown in Table 5.14. In this model, both regimes have the same dynamics, but the variances of the regimes are different. In addition, since the accumulation of the current account over output is excluded from the switching equation, the probability of the switch is constant. The restrictions from the unrestricted model in Table 5.12 cannot be rejected. Since the dynamics are the same both regimes share the same growth rate, 0.56% per quarter. The probability of being in the first regime is constant and equal to 0.70.

[65]For the moment no restrictions are imposed in any of the parameters of the switching equation.

The time paths of both regimes, since they have the same dynamics, are stable. There are two pairs of complex roots (see Table 5.14) which indicates a mixture of cycles. The dispersion in the first regime is higher than that of the second.

Therefore, it is clear that the dynamics of output growth for France can be adequately described by a linear model with a heteroskedastic pattern in the error terms rather than by either the linear model of section 5.2.2 or the switching regressions model of Table 5.13.

5.3.3.4 West Germany.

The unrestricted switching regressions model estimated for West Germany is presented in Table 5.15. The coefficient of the accumulation of the current account over output is significant and, therefore, it can be said that for West Germany, there are two regimes whose switches are governed by this variable. The joint hypothesis of no misspecification cannot be rejected at 5%. Nevertheless, AR effects of order one seem to be present in the first regime. The significance of output growth in the rest of the countries will be tested. These 8 restrictions are openly rejected with a LR test statistic equal to 24.11[0.0022] and, thus, the average growth rate in the rest of the countries seem to be important in determining the growth rate.

Table 5.16 shows the parameter estimates of the model selected by AIC. The model has one lag of output growth in the first regime, and two in the second. Additionally, both regimes have two lags of output growth in the rest of the countries. It is clear from Table 5.16 that this model is well specified, as the joint hypothesis and all the individual hypotheses of misspecification cannot be rejected at 5% level. The coefficient of the accumulation of the current account over output is very significant. At first sight, the dynamics of both regimes seem to be quite different. However, since the number of lags of output growth are not the same in both regimes (regime one has only one lag of output growth whereas regime two has two lags) test the unrestricted model in Table 5.15 is tested against a restricted model with equal dynamics and two lags of both output growth and output growth in the rest of the countries. The parameter estimates of this restricted model can be found in Table 5.17. Surprisingly, these restrictions cannot be rejected at 5%. However, the value of the likelihood drops by such an amount that the AIC is higher for the model with different dynamics. On this basis, the model with different dynamics is selected as the final model for West Germany (see Table 5.16).

The calculated growth rates of the two regimes appear to be quite different and, in fact, the hypothesis of equality of growth rates is markedly rejected. This is in contrast to the other three countries discussed earlier. Growth in the first regime is calculated to be 1.16% per quarter, while in the second regime it is 0.44% per quarter. That is, the growth rate in the first regime is higher than that of the second regime, which seems to indicate the presence of a long run balance of payments constraint on growth. Closer inspection of the probabilities associated with both regimes reveals that the probability of being in the first regime will be higher than 0.5 (and thus higher than the probability of being in the second regime) when the accumulation over the last 2 years of the current account surplus reaches 29.10% of output. This indicates that high and prolonged surpluses boost growth but regime 2 which includes deficits and surpluses could be taken as a normal growth regime and thus there seems to be little indication of deficits constraining growth in the case of West Germany. Nevertheless, this could be because during the sample period West Germany had long periods of high surpluses compared to the size of the deficits (see Figure 5.2), so in this period West Germany can hardly be regarded as having been constrained by balance of payments deficits. The probabilities of being in the first regime for each point in the data sample can be seen in Figure 5.6. It is also worth mentioning at this point that both regimes are stable. The first regime has one real root with a modulus less than one (see Table 5.16) suggesting that the growth rate will tend to the mean value while the second regime has a pair of complex roots indicating cyclical behaviour. This cyclical behaviour is possibly the result of including deficits and small surpluses together in this regime and that might also be the reason why the dispersion in regime 2 appears to be higher than that of the first regime.

5.3.3.5 Italy.

Table 5.18 presents the estimated parameters of the switching regressions model for Italy. The accumulation of the current account over output is not significant at 5% (significance equal to 0.0698) but it is significant at 10%. However, this model is misspecified and there is evidence of AR effects of order one in the first regime and ARCH effects of order one in both regimes.

The model selected by AIC has $k_1 = k_2 = \tilde{k}_1 = 4$, $\tilde{k}_2 = 3$ and no intercept in the switching equation. The two restrictions imposed from the unrestricted model are accepted with a LR statistic and significance equal to 0.2340[0.8896]. The coefficient of switching variable is significant at 5%. However, this restricted model is still misspecified, with evidence of both AR and ARCH effects of order one in the first regime.[66]

For this reason, the final model selected for Italy is shown in Table 5.19. This model has 3 lags of output growth in both regimes. In addition, regime 1 has one lag of

[66] The test statistic for ARCH effects of order 1 in the second regime is equal to 3.1726[0.0749] and therefore not significant at 5%. The joint test for misspecification and the individual tests for AR and ARCH effects of order one in the first regime are equal to 17.5278[0.0015], 10.0726[0.0015] and 11.1880[0.0008] respectively.

output growth in the rest of the countries whereas regime 2 has 3 lags of this variable and the switching equation has no intercept. The LR test accepts the restrictions at 5% significance level. Additionally, this model passes the joint misspecification test as well as all the individual misspecification tests at 5% significance level. Restricting the attention to the switching regression it is immediately obvious that the coefficient of the accumulation of the current account over output is significant at 5%. Therefore, this variable determines two regimes in the process of output growth.

Once more, a test of linearity is essential in this context. Therefore, this switching regressions model is tested against a linear model with a heteroskedastic pattern in the error terms. The restrictions of equal dynamics but different variances in the regimes, that is, $\beta_{1,,j} = \beta_{2,j}, \gamma_{1,k} = \gamma_{2,k}$ for $j = 1, \ldots, 3$ and $k = 1, \ldots, 3$ and $\beta_{30} = 0$ are clearly rejected. Therefore, a linear model is too restrictive to analyse the growth process in Italy.

In analogy with the other countries, the significance of output growth in the rest of the countries is tested. These restrictions are again widely rejected. The next logical step is, therefore, to concentrate on whether the estimated growth rates in each regime differ. At fist sight, the estimated long run growth rates appear to be different in both regimes. The estimated growth rate in the first regime is equal to 0.50% per quarter and lower than the estimated growth rate in the second regime, 0.60% per quarter. However, their equality cannot be rejected, with a test statistic equal to -0.395[0.6930]. The growth process in Italy has therefore two distinct regimes, in the sense of different dynamics and the changes between these two regimes are governed by the variable measuring the extent of the accumulated deficit/surplus of the current account over output.

However, the growth rates achieved in these two regimes are not significantly different so as to portray a balance of payments constraint on growth.

The probability of being in the first regime is higher than the probability of being in the second regime whenever the accumulated current account over output is positive, that is, when there is an accumulated surplus in the current account. The probabilities of being in regime one for each data point are shown in Figure 5.7. In the case of Italy, the dispersion in both regimes is very similar, although somewhat higher in the second regime. Thus, the probabilities of being in each of the regimes are important for assessing the dynamics of the process of growth, although the shocks to output in both regimes are very similar.

The first regime has a pair of complex roots and one real positive root with modulus less than one (see Table 5.19). Therefore, the time path of this regime is convergent. However, the second regime has a pair of complex roots and one real negative root with modulus slightly higher than 1. Therefore, the time path of this regime is explosive. The trajectory will show cyclical behaviour but the cycles will slowly grow in magnitude.

5.3.3.6 Canada.

Table 5.20 shows the parameter estimates of the unrestricted switching regressions model estimated for Canada. The coefficient of the switching variable is significant and the model is well specified. The null hypotheses of the joint test and every individual misspecification tests cannot be rejected at standard significance levels.

The model chosen by AIC has $k_1 = k_2 = 2$, $\tilde{k}_1 = 4$, $\tilde{k}_2 = 3$ and $\beta_{30} = 0$. The restrictions needed to go from the unrestricted model in Table 5.20 to this model cannot be rejected with a LR statistic of 4.597[0.5964]. Nevertheless, even though the joint hypothesis of misspecification cannot be rejected at 5% significance level (the test statistic is equal to 7.987 and the significance level is equal to 0.09205), there is evidence of ARCH effects of order one in the first regime with a test statistic equal to 6.024[0.0141]. Consequently, the first regime is augmented with one more lag of output growth. This final model is shown in Table 5.21. The restrictions imposed on these model with respect to the unrestricted one cannot be rejected. The model is well specified since none of the null hypotheses of the misspecification tests can be rejected at conventional significance levels.

The coefficient of the accumulation of the current account over output is significant, leading to a two-regimes model. Additionally, the lags of output growth in the rest of the countries are jointly significant.[67] Thus, these type of feedback effects across countries are very important and the corresponding lags of this variable need to be included in the model.

After establishing the significance of the accumulation of the current account over output and the lags of output growth in the rest of the countries, the next step is to identify whether a linear model can be rejected in favour of the switching regressions model presented here. The hypothesis of equal dynamics but different volatility is rejected.[68] Thus, a linear model even allowing for heteroskedasticity in the error term is too restrictive to model output growth in Italy as it was also the case for all the countries described earlier with the exception of France.

Most significantly, the accumulation of the current account surpluses and deficits over output not only influences the dynamics of the regimes but also the growth rates in

[67] This hypothesis involves the restrictions $\gamma_{1,k} = \gamma_{2,k} = 0$ for $k = 1, \ldots, 4$ together with $\beta_{1,4} = \beta_{2,3} = \beta_{2,4} = \beta_{3,0} = 0$.

[68] In this case, the null hypothesis is $\beta_{1j} = \beta_{2j}$ for $j = 1, \ldots, 3$, $\gamma_{1,k} = \gamma_{2,k}$ for $k = 1, \ldots, 4$ and $\beta_{14} = \beta_{24} = \beta_{30} = 0$.

each regime since the hypothesis of equality of long run growth is clearly rejected. The growth rate in regime one is estimated to be 1.14% per quarter. In the second regime, the estimated growth rate is negative and equal to -0.26% per quarter. The probabilities associated with regime one for each data point are depicted in Figure 5.8. In the case of Canada, the probability of being in regime 1 is higher than 0.5 (and, therefore, higher than the probability of being in regime 2) if the accumulation of the current account over output is positive. The two periods for which the likelihood of being in the high growth regime is higher than the likelihood of being in the second regime correspond to the only two periods for which the accumulated current account over output was in surplus. That is, at the beginning of the sample during 1972 and from the end of 1983 to the end of 1985. Therefore, in the case of Canada, it is clear that the large accumulate deficits present during the sample period (see Figure 5.2) have adversely affected growth. The model for Canada is perhaps the one which highlights more strongly the dangers of the restrictions imposed by linear models. In the linear model estimated for Canada in section 5.2.2, no evidence was found of any effect of the accumulated current account variable on output growth. However, once the implicit restrictions imposed by this linear model are relaxed, it is quite clear that this variable influences both the short run dynamics of the process of growth and the growth rate itself.

The characteristic roots of the polynomials in each regime are shown in Table 5.21. Both regimes are stable. The first regime has a pair of complex roots and one negative real root with an absolute value less than one. This indicates a cycling behaviour but as the real root takes over the cycles become smaller and smaller in magnitude, finally tending towards a mean value. The second regime has two real roots with opposite signs, both of them less than one in absolute value. Therefore, this regime after some initial oscillations tends towards a mean value.

5.3.3.7 Japan.

Table 5.22 gives the parameter estimates of the switching regressions model for Japan. The coefficient of the accumulation of the current account of the balance of payments over output is significant at standard significance levels. The model is well specified since none of the null hypotheses of the misspecification tests can be rejected at conventional levels of significance.

The parameter estimates of the switching regressions model selected for Japan by the AIC are shown in Table 5.23. This model has two and three lags of output growth in the first and second regime respectively. In addition, both regimes have four lags of the average of output growth in the rest of the countries. This model is again well specified since there is no evidence of misspecification in the tests reported. The coefficient of the switching variable is significant and, consequently, there is evidence of regime switches governed by the accumulation of the current account over output.

When trying to test the significance of the lags of the average output growth in the rest of the countries and the hypothesis of equal dynamics but different volatilities in both regimes, it was found that for the case of Japan, the imposition of these restrictions causes the likelihood function to become too flat and, therefore, it proved impossible to find a maximum with an acceptable degree of confidence.[69]

[69] Obviously, as it was discussed extensively in Section 5.3.2, the likelihood is unbounded. Consequently, we can always find convergence of the likelihood function in this area due to the finite precission of the computer. However, this maximum of the likelihood function is not valid since the parameter estimates are not consistent.

The estimated long run growth rates in both regimes are different. The calculated long run growth rate in the first regime is equal to 0.35% per quarter, whereas it is equal to 1.08% per quarter in the second regime. This seems to imply that in the case of Japan accumulated balance of payments deficits boost growth. The restriction that the growth rates are equal cannot, however, be rejected at 5% significance level, although it is just rejected at 10%. The probabilities of being in regime one associated with each point in the data used are shown in Figure 5.9. The probability of being in regime one is higher than the probability of being in regime two if the accumulated current account over output surplus is higher than 0.1089. Therefore, it seems to be the case for Japan large accumulated surpluses do not significantly increase growth or, might even prevent the sort of growth achieved with small accumulated surpluses or small accumulated deficits. Nevertheless, they influence the dynamics of output growth since there is evidence of two different regimes with switches depending on the position of this variable.

It is worth recalling at this point the linear model estimated for Japan in section 5.2.2. In that linear model both, the accumulation of the current account over output and the lags of output growth in the rest of the countries were found to be insignificant in the process of growth. However, it is clear from the switching regressions model estimated here that these variables contribute significantly to understand the growth process in Japan.

Looking at the characteristic roots of the polynomials associated with each of the regimes, it is found that the first regime has two real roots with opposite signs with modulus less than one. This implies that after some initial fluctuations due to the opposite signs of the real roots, output growth will tend towards a mean value. The second regime has a pair of complex roots and one positive real root with an absolute value less than one. Therefore, initially, as the pair of complex roots dominates, the trajectory will oscillate but then it will settle towards a mean value.

This section has presented the switching regressions models of output growth estimated for each of the G7 countries. There was little evidence of a balance of payments constraint on growth in these countries. This hypothesis seems to be supported only in the case of Canada. Nevertheless, there is clear evidence that the accumulation of the current account of the balance of payments over output influences the dynamics of growth for all the countries with the exception of France. This finding is consistent with the model of constrained borrowing presented in chapter 2. Clear evidence was also presented in favour of the inclusion of output growth in the rest of the countries in the growth equations highlighting even more the importance of interactions across countries. It is also important to note that the switching regressions models estimated in this section encompass a particular type of linear model with heteroskedastic errors. It was found that the restrictions imposed by this linear model are not accepted by the data for any of the countries under study here with the exception of France. The linear models estimated in section 5.2.2 are even more restrictive than these, in the sense that they not only impose equality of dynamics for both process but also impose homoskedasticity in the error terms. Those linear models failed to find any effect of both the accumulation of the current account over output for the majority of the countries and the lags of output growth in the rest of the countries for some of the countries. These findings are a direct consequence of the restricted nature of the linear model. It is obvious, therefore, that the use of switching regressions models is essential if a better understanding of the growth process is to be achieved. Additionally, it is also clear that inter-relationships across countries play a very important role in the growth process, and as a result any investigation of output growth needs to take into account these interdependencies.

5.4 Estimated Impulse Response functions for the switching regressions model.

The switching regressions models fitted in the last section suggest asymmetries in the behaviour of output in the sense that the dynamics are different in each regime depending on the position of the accumulation of the current account of the balance of payments over output. In contrast, the linear models of section 5.2.2 cannot capture this asymmetric behaviour. This section analyses the responses of output to shocks by means of impulse responses. In section 5.2.2 the impulse responses of linear models were analysed. The present section concentrates on obtaining the impulse responses of the switching regressions models obtained in section 5.3.3. These impulse responses will also be compared to those of a linear model.

The main difference between the impulse responses of linear and nonlinear models is that, in the case of nonlinear models, the impulse responses are history, shock and composition dependent (see chapter 4). Consequently, the impulse responses are better viewed as whole distributions. In the case of distributions of the impulse response function which are symmetric and not widely spread, the average response and, even sometimes, a particular response will provide a good measure of persistence and, furthermore, a good idea of the response path followed by a series after a shock. If this is not the case, then, the average response might be quite misleading. This point is depicted more clearly in the next two sections.

The focus here is on the distribution of the impulse responses as well as how the responses might vary across the regimes of the switching regression models estimated in the last section in each country and across countries. The impulse responses are obtained using the estimated parameters in section 5.3.3 and randomly drawing from the distribution of the residuals (see chapter 4, section 4.6). The current account of the balance of payments is forecasted using an AR process for the fourth difference of the current account ($\Delta_4 ca_{i,t}$) and $z_{i,t}$.[70] The estimated regressions are shown in Table 5.24. All histories (ω_{t-1}) in the sample are employed and the joint distribution (conditional on the regime) of the innovations at each history is drawn 100 times. Each history together with the random innovation for the current account equation determines the actual regime. The maximum horizon N is set to 20 (5 years) and the average is taken over 200 futures (R). It is further assumed that the rest of the countries continue to grow at the same rate as the last known figures for each particular history.

Section 5.4.1 presents the average impulse responses for each of the G7 countries. In addition, two histories are selected in each regime for each of the countries and compared across regimes and countries. These average and individual impulse responses will illustrate the dangers of looking only at the average behaviour of the series after a shock, or, even worse, the danger of selecting a particular history instead of looking at the whole distribution of the impulse responses.

Section 5.4.2 focuses, then, on the distribution of the impulse responses across the countries under study here. Additionally, the distribution of the impulse responses conditional on each of the regimes will also be presented for each of the countries.

[70]Note that both regressions, for the US and West Germany, have a dummy variable in 1991q1 and the regression for France has a dummy in 1982q2. These dummies are introduced to take care of large residuals in those years but they do not have a significant effect on the parameters of the model.

5.4.1 Average and individual impulse responses.

When dealing with nonlinear models, it is tempting to pick a particular history and calculate an impulse response function based on this history. However, the response of a series to a shock in a nonlinear model is history dependent and, therefore, a particular history gives just a particular example of a possible response to a shock. If there is a lot of variation in the responses, to have a more accurate picture of the behaviour of the series, the whole distribution of the responses needs to be studied rather than a couple of them in isolation. Figures 5.10 to 5.16 show the average response of the impulse response functions of the logarithm of the level of output for each country for a shock of size $+1$.[71] These averages are calculated conditional on each regime. This allows for comparisons in the average responses between the two regimes for any one country and it also allows for comparisons across different countries. The average responses for the UK, West Germany and Italy level off before reaching the horizon considered here. The average responses for the USA, however, reach a peak after 6 periods and from then onwards seem to start decreasing very slowly. The average impulse responses for Canada and Japan do not level off either. For these two countries they seem to still increase slowly with time. It will be clear later why the average response for these three countries does not seem to level off after 20 periods. The first thing to notice in Figures 5.10 to 5.16 is that the average responses for each country are very similar in both regimes, but the average responses across country are markedly different.[72] The

[71] The reason why the impulse responses are calculated only for a shock of size +1 is because in the switching regime model, conditioning on one regime, makes the model linear. Thus, after conditioning on one regime, the responses are symmetric and the response of a size $+n$ shock is exactly the same as n times the response of a size +1 shock.

[72] The case of France is different to the rest of the countries, since the model has the same dynamics in both regimes and the accumulation of the current account over output is not significant. In this case, therefore, the model is essentially linear.

average responses of output for the USA and the UK are very similar in shape, although quantitatively different. The effect of a shock is magnified after the initial shock until it reaches a peak (earlier in the UK than in the USA) and then the responses level off. This behaviour is similar in shape to the behaviour found for all the countries using the linear model of section 5.2.2 (see Figure 5.3).[73] However, even though the shape of the average impulse responses is similar in the linear model and the switching regressions model for both the USA and the UK, the responses are quantitatively different. For the USA, a typical shock in the linear model generates a larger response in output than what a typical shock in the nonlinear model would generate. This is very important since mistakenly using the linear model, it could be concluded that a positive shock increases output by a larger amount than it actually does. The UK is a special case, since the linear model of section 5.2.2 does not have any significant lags of output growth as independent variables.[74] As a consequence, a shock to output of size one, will generate a response equal to one. After the initial shock, output will stay at this level in the absence of additional shocks. In contrast, in the switching regressions model, some lags of output growth are found to be significant. The response of output to a typical shock in this nonlinear model is magnified in the long run and, therefore, it is higher than what the linear model would suggest. The response of Canadian output to a shock in the linear and nonlinear models is also similar, in the sense that the effect of the shock is magnified after the first period. However, even after 20 periods, the average response for Canada in the switching regressions model has not yet reached a peak and level off, that is, it is still increasing although at a decreasing rate. This

[73]In Figure 5.3, there are no impulse responses for the United Kingdom and West Germany since in the linear models for these two countries no lags of output growth were found to be significant. Therefore, the response of output toa shock is equal to the size of the shock.

[74]This is also the case for West Germany.

behaviour contrast with the typical behaviour of an I(1) series. This behaviour will be discussed in more detailed in the next section. In the case of Canada, the response of output to a typical shock in the linear model is a lot smaller than the response found with the switching regressions model. Therefore, once more, the linear model gives misleading responses.

The behaviour of the output series in the rest of the countries is very different from the behaviour found using the linear models in section 5.2.2. Perhaps the most similar is that of Japan. After the initial shock to output, the response in the switching regressions model is dampened for one period and then it starts increasing again. Although the linear model captures the general shape, it fails completely to capture the drop in output after the initial shock. This drop in output and, therefore, a period of negative growth, will have consequences in the short run for the economy which the linear model is unable to predict. In addition, the long run response obtained by the linear model is more than double the long run response obtained by using the switching regressions model.

The case of France is very interesting. For this country it was found in section 5.3 that the appropriate model was linear but with a heteroskedastic pattern in the errors. However, it is clear from the impulse responses obtained from the linear models of section 5.2.2 and that obtained here, that the linear model in section 5.2.2 is unable to capture the periods of negative growth after the initial shock. Instead they are portrayed as periods of constant growth. Thus, even though the model in 5.2.2 seems to be well specified, it is clear that it is unable to capture the full dynamics of output growth.

Similar situations are found for West Germany and Italy. The linear model in section 5.2.2 for West Germany had no lags of output growth. Thus, a shock will generate a response in output equal to the size of the shock and this will be equal to the long run response. It is clear from Figure 5.13, that this is not the case. After the initial shock, there are two periods of negative growth before the response of output increases again. After some fluctuations, the response of output levels off, but after 20 periods, the response is smaller than the initial size of the shock. This contrast with the linear model which depicts a response exactly equal to the size of the shock.

When comparing the impulse responses for output growth in Italy, once more, large differences were found between the linear and nonlinear models. Similar to the case of West Germany, the linear model finds that the response of output to a shock is magnified in the long run. The response of output to a shock when the nonlinear model is used is, however, completely different. The long run response is actually smaller than the size of the shock. Furthermore, the average response of output has large swings before levelling off at the level of the long run response. Once more, the linear model is unable to capture these dynamics giving misleading inferences about the behaviour of output growth after a shock.

So far, the average responses of output for the G7 countries obtained with the switching regressions model have been compared to those obtained with linear models. It has been shown how misleading inferences made based on linear models can be, since the type of dynamics in the growth process which were found present in the nonlinear models of section 5.3.3 cannot be captured by linear models. This is mainly a consequence of the restrictive nature of linear models.

However, in nonlinear models, the average response does not tell the whole story (see chapter 4). Figures 5.17 to 5.22 show two individual responses for each regime for each of the countries. From these graphs it is immediately obvious how different responses can be in switching regressions models depending on the history chosen. As an example, the long run response of a shock of size one for the UK can either be magnified or curtailed. This is also the case of Italy and Canada. Even if both responses are either magnified or curtailed in the long run, both the short run and the long run responses are quantitative very different. One exception to this is West Germany. Although the responses are different, the response after 20 periods is quantitatively more similar for two different histories than it is for the rest of the countries. The two histories selected for each regime for each country are just some extreme examples, and seem to be similar for both regimes. However, it is possible to find examples for which shocks in the first regime are more persistent than in the second and vice versa. Consequently, these graphs illustrate why it is so important to look at the distribution of the responses instead of looking at the average behaviour or, even worse, to just a particular history when dealing with nonlinear models. If the distribution of the impulse response functions is symmetric and not widely spread, then, the average response would provide a good measure of persistence. However this will not be the case if the distribution is very skewed and widely spread.

Next section concentrates on the distribution of the impulse responses in each of the G7 countries and discusses their importance in more detail.

5.4.2 Box plots of the impulse responses.

Figures 5.23 to 5.28 show the box plots at each horizon for each country in the sample. These box plots outline the minimum, first quartile, median, third quartile and maximum response of the generalized impulse response function of the log level of output. The procedure is conducted conditional on the regime as well as unconditionally to give an idea of the differences in responses obtained from different regimes. It will be shown that the average response is not a very good indicator for most of the countries studied here, since the distributions of the impulse responses are widely spread. This highlights the importance of looking at the whole distribution of responses when using switching regressions models.

Most of these figures show the median response levelling off after approximately 8 quarters. This levelling off is typical of I(1) behaviour which is what both tests (ADF and Kwiatkowski *et al.*) in section 5.2.1 suggested.

Figure 5.23 shows the boxplots of the generalized impulse response functions for the USA for shocks fixed at one standard deviation. The dispersion in both regimes is high. The responses lie in the interval $[1.2841, 2.4984]$ after 20 periods in regime 1and in the $[1.1376, 2.2618]$ in regime 2 after 20 periods. Furthermore, the distribution, in the second regime seems to be quite symmetric and the range between the first and the third quartile is $[1.5277, 1.8641]$. The dispersion as measured by the standard deviation of the responses in the second regime increases with the time horizon and it reaches a peak after three time periods. Subsequently, it decreases and after twelve time periods, it stabilises at a value of 0.2331. The distribution of the first regime seems to be skewed to the left. Therefore, the average response in this regime will tend to overestimate the likely response. The range between the first and third quartile is slightly larger $[1.5245, 1.8807]$. The dispersion in this regime reaches a peak after 5 time periods and in the same way as the second regime it stabilises after twelve time periods at a value of 0.2698, higher than the dispersion in the second regime. The median response is similar, 1.6524 in regime 1 and 1.6790 in regime 2 after 20 periods. Thus it can be said that the persistence of shocks in both regimes is similar.

Figure 5.24 presents the boxplots for the UK. The UK is similar to the USA in that the dispersion in both regimes is high, although slightly lower than the dispersion of responses for the USA. After 20 periods the responses lie in the interval $[0.8811, 1.6225]$ for regime 1 and $[0.7955, 1.6591]$ for regime 2. This is interesting, since in the case of the UK, since a shock can increase the logarithm of output by a factor either higher or lower than the initial size of the shock, although it is never completely reversed. The dispersion of both regimes reaches a peak after 3 time periods and then the dispersion stays at this level (0.1865 for regime 1 and 0.2075 for regime 2). The distributions of the responses in both regimes are skewed to the left although the distribution in regime 1 is more asymmetric. The median of the response in regime 1 is slightly lower, 1.1180 compared to 1.1373 in regime 2.

The box plots for West Germany are shown in Figure 5.25. The first difference between West Germany and the countries considered earlier (apart from the shape of the responses) is that the dispersion of the responses after 20 time periods is very small in both regimes. After 20 time periods, the responses lie in the interval $[0.6215, 0.7934]$ in regime 1 and $[0.6160, 0.7537]$ in regime 2 and the range between the first and the third quartile is small: $[0.6804, 0.7132]$ in regime 1 and $[0.6842, 0.7009]$ in regime 2. The dispersion in both regimes reaches a peak after just one time period but then it decreases rapidly to the values 0.0288 and 0.0184 for regimes 1 and 2 respectively. However, even at its peak, the dispersion is still quite low (0.0663 for regime 1 and 0.0594 for regime 2). The distribution of the responses after 20 time periods is almost symmetric in the first regime and slightly skewed to the right in the second regime. It is interesting to note that the distribution of the responses in both regimes after just one time period is very much skewed to the right. The median of the responses after 20 time periods is almost identical for West Germany. The median response in regime 1 is 0.6966 and the median response in regime 2 is 0.6918. In this case it can, therefore, be concluded that the persistence in both regimes is very similar and, even more, since the distribution of the responses in both regimes is quite tight, the average response will give a good indication of the persistence in this case.

Figure 5.26 shows the corresponding boxplots for Italy. In this case, the dispersion in both regimes is very similar in magnitude to the dispersion found for the USA and the UK. The responses in regime 1 are in the interval $[0.4952, 1.4079]$ after 20 periods and in the interval $[0.4905, 1.3544]$ in regime 2. The dispersion in both regimes reaches its peak after five time periods, and it then decreases a bit and stabilises at a value of 0.1943 in the first regime and 0.1799 in the second regime. The distribution of the responses after 20 time periods in both regimes are slightly skewed to the left. Again, the median of the responses is similar in both regimes, 0.8900 in regime 1 and 0.8157 in regime 2 and therefore, the persistence to shocks in both regimes is quite similar.

The boxplots for Canada are shown in Figure 5.27. In this case the dispersion in both regimes is very high and after 20 periods the responses lie in the interval $[0.8492, 3.5182]$ in regime 1 and in the interval $[0.7837, 3.1155]$ in regime 2. Thus, the dispersion is equal to 0.6826 and 0.6581 after 20 periods in regimes 1 and two respectively. Even the dispersion between the first and the third quartile after 20 time periods is very high: the ranges between the first and the third quartiles are $[1.1337, 2.0389]$ and $[1.1389, 2.0141]$ for regime 1 and 2 respectively. The dispersion of the responses in both regimes seem to increase with time, however this is due to some extreme responses.

The distributions in both regimes are skewed to the left, thus the average response will overestimate the likely response in both regimes. Again the median response after 20 time periods is similar in both regimes although slightly higher in the second regime, 1.6266 in regime 1 and 1.6450 in regime 2.

Figure 5.28 presents the boxplots for Japan. The dispersion in both regimes is high. In both regimes the dispersion increases rapidly during the first two periods, but it drops significantly in the third period only to increase again in the fourth time period. After 20 time periods, the dispersion in regime 1 is 0.3121 and 0.3111 in regime 2. The responses lie in the interval $[1.1722, 2.3988]$ in regime 1 and $[1.2275, 2.3337]$ in regime 2 after 20 time periods. The distributions in both regimes are almost symmetric slightly skewed to the left if anything. The median of the distribution is higher in regime 1, 1.7002, than in regime 2, 1.6550 after 20 time periods.

From this analysis it can be concluded that for West Germany,the persistence of shocks is quite similar between both regimes. For The USA, the UK, Italy, Canada and Japan the median of the generalized impulse response functions is quite similar but the distribution is very spread so in these cases the particular initial conditions turn out to be very important in assessing persistence.

So far both regimes within each country have been compared . If attention is restricted to the median of the distribution of the unconditional responses (that is including all the generalized impulse response functions irrespective of the regime they are in at $t = 0$), for all the countries except from Italy and West Germany, the logarithm of output is increased by more than the initial size of the shock. After 20 periods the smallest factor for the median response is 0.6940 (West Germany) and the highest is

1.6987 (Japan). However the overall minimum factor can be as low as 0.4905 (Italy) and the maximum as high as 3.5182 (Canada) depending on the initial conditions.

Therefore, it can be concluded that across the G7 countries, the responses of output to shocks are very different when attention is paid at the median response, but even more different when their distributions are compared.

5.5 Conclusions.

Recent empirical studies on output growth have mainly concentrated on linear models which unfortunately place certain symmetry restrictions that, if wrongly imposed, could bias the measures of persistence. Lately, there has been a switch towards nonlinear models which relax these restrictions and allow for more complicated dynamics than a simple linear model but the focus of the studies has been on output in isolation. Clearly, the introduction of other relevant macroeconomic variables would enhance any growth analysis. In this chapter, the closed economy assumption postulated in most growth studies is relaxed and the effects of current account deficits on growth dynamics and long run growth have been assessed. In the countries studied here, there seems to be little evidence of a long run balance of payments constraint on growth with the exception of Canada. However, there is evidence that past accumulations of current account deficits or surpluses influence the short run dynamics of the growth equation. A possible explanation for this is that a borrowing constraint exists in the international market as shown in chapter 2. It could also be the case that the deficits encountered in this sample are not big enough to trigger the balance of payments constraint or that the accumulation over two years of the current account/output ratio fails to pick up the effect of the constraint that some other accumulation would single out.

On the other hand, there is wide evidence supporting the inclusion of lagged output growth in the rest of the countries in the model. Although this is a very simple method of allowing for feedback effects across countries, it proves how important it is to take into account the behaviour of other countries if any meaningful conclusions about output growth are to be found.

It has also been shown how persistence varies with the initial conditions chosen even for the same regime and same size of shock making use of the Generalized Impulse Response functions of Koop, Pesaran and Potter (1996). They have been adapted to the switching regressions model utilized here. It has also been successfully demonstrated that although the median persistence of shocks to output in both regimes appears to be quite similar, the distribution of the responses is very wide for some countries (an extreme example is Canada). In these cases, the average response is not a good indicator of persistence to the extent that conditioning on a particular regime and shock the response of the logarithm of output to a shock can either be magnified or curtailed. In conclusion, when the impulse responses are dependent on the initial conditions, the focus should be on the distribution of the responses rather than simply examining a couple in isolation.

Clearly, there is still scope for future development in this area. The switching regressions model studied here is piecewise linear, that is, even though the model as a whole is nonlinear, conditioning on a particular regime becomes linear. Thus the model could be further improved by allowing for nonlinearities in each regime. This would lead to different measures of persistence for positive and negative shocks even after conditioning on a regime. It is also worth noting that even though feedbacks across countries have been introduced here in an overly simplified manner they appear to be of importance. Therefore, a more sophisticated treatment of the links between countries will certainly lead towards a more realistic model of growth.

Table 5.1. Summary statistics for $\Delta y_{i,t}$ and $z_{i,t}$. Sample period 1972Q1 to 1994Q4.

	Descriptive statistics for $\Delta y_{i,t}$						
	USA	**UK**	**FRA**	**GER**	**ITA**	**CAN**	**JPN**
Mean	0.0064	0.0049	0.0056	0.0055	0.0066	0.0079	0.0092
Std. Deviation	0.0096	0.0132	0.0073	0.0089	0.0096	0.0101	0.0091
Skewness	-0.5065	0.9375	-0.5725	-0.2781	0.0970	-0.0248	-0.5086
Kurtosis - 3	1.7195	3.6863	0.3987	-0.0650	0.8626	-0.2164	1.6253
Minimum	-0.0260	-0.0268	-0.0177	-0.0196	-0.0217	-0.0145	-0.0260
Maximum	0.0316	0.0595	0.0185	0.0241	0.0343	0.0304	0.0312
	Descriptive statistics for $z_{i,t}$						
	USA	**UK**	**FRA**	**GER**	**ITA**	**CAN**	**JPN**
Mean	-0.0785	-0.0409	-0.0206	0.1001	-0.0465	-0.1315	0.1244
Std. Deviation	0.1010	0.1334	0.0499	0.1428	0.0967	0.1150	0.1120
Skewness	-0.6301	-0.2056	-0.2765	0.4446	0.3070	-0.1535	-0.1352
Kurtosis - 3	-0.8758	-0.9857	0.7184	-0.6859	-0.5579	-0.9428	-0.9904
Minimum	-0.2792	-0.3091	-0.1534	-0.1336	-0.2313	-0.3467	-0.0831
Maximum	0.0561	0.1880	0.0941	0.3847	0.1718	0.0609	0.3164

Table 5.2. Augmented Dickey-Fuller tests for the logarithm of output ($y_{i,t}$) and the sum of the current account/output ratio ($z_{i,t}$).[75]

	$H_0 : y_{i,t} \sim I(1)$		$H_0 : y_{i,t} \sim I(2)$		$H_0 : z_{i,t} \sim I(1)$	
	$\varkappa$	Statistic	$\varkappa$	Statistic	$\varkappa$	Statistic
USA	1	-2.9072	0	-6.7617	4	-1.8029
UK	0	-2.0584	0	-10.1419	3	-2.9532
FRA	2	-2.8926	1	-4.5128	2	-3.6812
GER	4	-3.4927	4	-4.2397	2	-2.7457
ITA	1	-1.8168	0	-7.9934	2	-4.9746
CAN	1	-2.1901	0	-7.2972	1	-1.2992
JPN	4	-2.8523	4	-3.7269	2	-2.7079
95% Critical value		-3.4586		-3.4586		-2.8947

[75] $\varkappa$ stands for the number of augmentations chosen by the Akaike Information Criterion. The ADF tests for $y_{i,t}$ include an intercept and a linear trend and the sample size is 92. The tests for $z_{i,t}$ include an intercept and the sample size is 87.

Table 5.3. Test of the null hypothesis of stationarity around a deterministic trend for the logarithm of output, $y_{i,t}$.

	Truncation parameter (ϱ)					
	0	1	2	3	4	5
USA	0.3038	0.1585	0.1106	0.0874	0.0740	0.0657
UK	0.7128	0.3736	0.2579	0.2002	0.1660	0.1434
FRA	0.7218	0.3754	0.2595	0.2022	0.1685	0.1464
GER	0.6145	0.3200	0.2214	0.1726	0.1441	0.1260
ITA	1.0399	0.5417	0.3786	0.2999	0.2549	0.2263
CAN	1.0452	0.5432	0.3746	0.2911	0.2419	0.2096
JPN	0.3880	0.2073	0.1456	0.1153	0.0978	0.0869
95% **critical** value: **0.146**						

Table 5.4. Test of the null hypothesis of stationarity around a level for the sum of the current account/output ratio ($z_{i,t}$).

	Truncation parameter (ϱ)					
	0	1	2	3	4	5
USA	5.0456	2.5392	1.7058	1.2912	1.0442	0.8810
UK	2.1133	1.0708	0.7246	0.5533	0.4521	0.3859
FRA	0.4397	0.2256	0.1554	0.1214	0.1020	0.0900
GER	1.2939	0.6396	0.4326	0.3303	0.2699	0.2305
ITA	0.5224	0.2687	0.1858	0.1460	0.1236	0.1100
CAN	4.2096	2.1462	1.4569	1.1137	0.9091	0.7740
JPN	4.1791	2.1097	1.4250	1.0869	0.8873	0.7569
95% **critical** value: **0.463**						

Table 5.5. Tests on the number of Cointegrating Vectors based on Johansen's Maximum Likelihood Approach.[76]

H_0	H_1	**Max. Eigenvalue**	**Critical Value** (95%)	**Trace**	**Critical Value** (95%)
$r = 0$	$r = 1$	55.1872	45.63	142.1014	124.62
$r \leq 1$	$r = 2$	33.9007	39.83	86.9141	95.87
$r \leq 2$	$r = 3$	26.8016	33.64	53.0135	70.49
$r \leq 3$	$r = 4$	11.7040	27.42	26.2118	48.88
$r \leq 4$	$r = 5$	6.1650	21.12	14.5079	31.54
$r \leq 5$	$r = 6$	5.3706	14.88	8.3429	17.86
$r \leq 6$	$r = 7$	2.9723	8.07	2.9723	8.07

Table 5.6. Estimated output cointegrating vector subject to a just-identifying restriction and one over-identifying restriction (SE's in brackets).[77]

Country	**CointegratingVector** (just-identifying restriction)	**CointegratingVector** (over-identifying restrictions)
USA	1.0000 (–)	1.0000 (–)
United Kingdom	−0.6936 (0.2396)	−0.5924 (0.1388)
France	1.6733 (0.5933)	1.5031 (0.4145)
West Germany	−1.3699 (0.3572)	−1.3285 (0.2903)
Italy	−1.7908 (0.8121)	−1.3748 (0.3092)
Canada	0.2227 (0.3611)	0.0000 (–)
Japan	0.6219 (0.2866)	0.5285 (0.1893)
Log Likelihood	2230.4	2230.2
LR test of restrictions	–	$\chi^2(1) = 0.55013[0.458]$

[76] Statistics are calculated on the basis of a VAR(2) model with unrestricted intercepts and no trends.

[77] Unrestricted intercepts and no trends in the VAR.

Table 5.7. (continues on the next page) Estimated linear regressions for each of the G7 countries. Standard errors are in round brackets and significance levels in square brackets.[78]

	Country			
Coefficient	USA	UK	FRA	GER
β_0	0.0037 (0.0016)	0.0055 (0.0014)	0.0021 (0.0012)	0.0012 (0.0014)
β_1	0.2497 (0.1086)	– –	0.0004 (0.1108)	– –
β_2	– –	– –	0.1875 (0.0986)	– –
β_3	– –	– –	– –	– –
γ_1	0.5514 (0.1892)	– –	0.4337 (0.1333)	0.4194 (0.1486)
γ_2	−0.3774 (0.1807)	– –	– –	– –
γ_3	– –	– –	– –	– –
δ	0.0011 (0.0092)	0.0145 (0.0103)	0.0211 (0.0141)	0.0133 (0.0062)
F1 stat	5.1535 [0.001]	1.9723 [0.164]	7.1296 [0.000]	7.0651 [0.001]
AIC	302.6470	266.9920	330.4114	308.0458
F2 stat [prob]	0.6482 [0.664]	0.7635 [0.636]	0.1725 [0.972]	0.6277 [0.732]
F3 stat [prob]	4.7242 [0.011]	– –	10.5853 [0.002]	7.9645 [0.006]
	Diagnostic tests			
Serial correlation	1.8144 [0.770]	2.7179 [0.606]	1.6997 [0.791]	4.0487 [0.399]
Functional form	0.2536 [0.615]	0.3318 [0.565]	0.3706 [0.543]	1.3765 [0.241]
Normality	12.4775 [0.002]	56.6378 [0.00]	9.3363 [0.009]	3.8229 [0.148]
Heteroskedasticity	0.1319 [0.716]	1.0166 [0.313]	0.3099 [0.578]	0.0089 [0.925]

[78] The the number of lags in the equations are selected according to the Akaike Information Criterion. $F1$ *stat* refers to the F statistic of the reduction of the model from 4 lags. $F2$ *stat* refers to the F statistic of the joint significance of all $\triangle\bar{y}_{i,t}$'s for country i.

Table 5.7 (cont.).

	Country		
Coefficient	ITA	CAN	JPN
β_0	0.0002 (0.0017)	0.0056 (0.002)	0.0042 (0.0019)
β_1	0.1453 (0.0812)	0.1951 (0.1059)	0.1095 (0.1075)
β_2	– –	– –	0.2167 (0.1057)
β_3	– –	– –	0.2050 (0.1073)
γ_1	0.4025 (0.1516)	0.7208 (0.1736)	– –
γ_2	0.4828 (0.1632)	−0.3612 (0.1798)	– –
γ_3	– –	– –	– –
δ	0.0112 (0.0099)	0.0123 (0.0084)	−0.0005 (0.0082)
F1 stat	15.8578 [0.000]	7.5227 [0.000]	3.3826 [0.0123]
AIC	317.4543	301.8358	303.7652
F2 stat [prob]	0.5978 [0.702]	0.3364 [0.889]	0.5706 [0.722]
F3 stat [prob]	11.0685 [0.000]	8.8055 [0.000]	– –
	Diagnostic tests		
Serial correlation	10.1260 [0.038]	0.7926 [0.939]	3.7107 [0.447]
Functional form	1.4483 [0.229]	3.4878 [0.062]	0.3856 [0.535]
Normality	9.1777 [0.010]	1.5359 [0.464]	13.7091 [0.001]
Heteroskedasticity	0.1658 [0.684]	1.3530 [0.245]	0.0523 [0.819]

Table 5.8. Estimated unrestricted switching regressions model for the USA (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0070 (0.0029)	−0.0789 (0.1875)	−0.0895 (0.2135)	−0.0276 (0.2253)	0.3357 (0.1867)	0.00333 (0.00272)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.9741 (0.3313)	−0.6345 (0.4129)	−0.1358 (0.4658)	−0.8193 (0.3069)	0.0098 (0.0012)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	0.0015 (0.0013)	0.5846 (0.0735)	0.1534 (0.0759)	0.0718 (0.0584)	−0.2175 (0.0653)	0.00820 (0.00137)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	0.3362 (0.1329)	0.3461 (0.1088)	−0.1240 (0.1056)	0.3997 (0.1197)	0.0030 (0.0004)	
Switching Equation	β_{30}	β_{31}				
	−0.1219 (0.2522)	−4.7601 (2.5580)				
Likelihood:				331.57076		

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	7.16	0.2092
AR(1) test for Regime 1	0.45	0.5027
AR(1) test for Regime 2	3.90	0.0482
ARCH(1) test for Regime 1	2.80	0.0946
ARCH(1) test for Regime 2	0.02	0.8888
Omitted Markov Effects	0.02	0.8952
Test for equality of growth rates	-1.58	0.1138

Table 5.9. Estimated switching regressions model for the USA (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0072 (0.0030)	-0.0891 (0.1876)	-0.0896 (0.2172)	-0.0259 (0.2321)	0.3529 (0.1863)	0.00303 (0.00284)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.9862 (0.3355)	-0.6555 (0.4220)	-0.1905 (0.4756)	-0.8312 (0.3069)	0.0098 (0.0013)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	0.0015 (0.0012)	0.5843 (0.0718)	0.1487 (0.0712)	0.0733 (0.0570)	-0.2192 (0.0635)	0.00828 (0.00133)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	0.3457 (0.1328)	-0.3460 (0.1065)	-0.1244 (0.1023)	0.4041 (0.1164)	0.0030 (0.0003)	
Switching Equation	β_{30}	β_{31}				
	– –	-4.0989 (2.0965)				
Likelihood:			331.44932			
LR test of restrictions:			0.2432[0.6219]			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	5.96	0.2026
AR(1) test for Regime 1	0.35	0.5551
AR(1) test for Regime 2	3.18	0.0747
ARCH(1) test for Regime 1	2.29	0.1300
ARCH(1) test for Regime 2	0.00	0.9820
Omitted Markov Effects	-	-
Test for equality of growth rates	-1.73	0.0838
Test of significance of the $\Delta\overline{y}_{i,t-j}$'s	25.73	0.0023
Equal dynamics test	21.30	0.0191

Characteristic roots	
Regime 1	Regime 2
$-0.01 + 0.80i$	$0.69 + 0.31i$
$-0.01 - 0.80i$	$0.69 - 0.31i$
-0.77	$-0.40 + 0.47i$
0.71	$-0.40 - 0.47i$

Table 5.10. Estimated unrestricted switching regressions model for the UK (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0036 (0.0036)	−0.2621 (0.1395)	−0.0003 (0.1662)	0.0332 (0.1470)	−0.0368 (0.1647)	0.00476 (0.00151)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.5866 (0.5070)	−0.0018 (0.5238)	−0.4437 (0.4822)	0.2047 (0.4052)	0.0141 (0.0013)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	−0.0006 (0.0008)	0.7930 (0.0548)	−0.1223 (0.0349)	−0.0266 (0.0470)	0.0338 (0.0428)	0.00345 (0.00177)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	−0.4845 (0.0829)	0.6590 (0.0996)	0.3197 (0.1115)	−0.2430 (0.1175)	0.0018 (0.0003)	
Switching Equation	β_{30}	β_{31}				
	−0.5194 (0.2189)	−2.5936 (1.5553)				
Likelihood:				293.434		

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	13.29	0.0208
AR(1) test for Regime 1	6.70	0.0096
AR(1) test for Regime 2	3.68	0.0551
ARCH(1) test for Regime 1	1.41	0.2359
ARCH(1) test for Regime 2	0.01	0.9200
Omitted Markov Effects	1.43	0.2320
Test for equality of growth rates	0.57	0.5674

Table 5.11. Estimated switching regressions model for the UK (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0056	−0.2173	0.0814	–	–	0.00491
	(0.0022)	(0.1263)	(0.1406)	–	–	(0.00165)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	–	–	–	–	0.0145	
	–	–	–	–	(0.0014)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	−0.0008	0.7884	−0.1246	–	–	0.00444
	(0.0008)	(0.0505)	(0.0306)	–	–	(0.00130)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	−0.4001	0.5945	0.3199	−0.1899	0.0017	
	(0.0944)	(0.0953)	(0.1031)	(0.0786)	(0.0003)	
Switching Equation	β_{30}	β_{31}				
	−0.6298	−3.3882				
	(0.2510)	(1.7204)				
Likelihood:			291.93624			
LR test of restrictions:			2.996[0.9346]			

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	9.17	0.1026
AR(1) test for Regime 1	4.33	0.0375
AR(1) test for Regime 2	2.00	0.1578
ARCH(1) test for Regime 1	1.03	0.3093
ARCH(1) test for Regime 2	0.15	0.7009
Omitted Markov Effects	2.11	0.1465
Test for equality of growth rates	0.22	0.8260
Test of significance of the $\Delta\overline{y}_{i,t-j}$'s	28.55	0.0046
Equal dynamics test	23.71	0.0140

Characteristic roots

Regime 1	Regime 2
-0.41	0.57
0.20	0.22

Table 5.12. Estimated unrestricted switching regressions model for France (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0015 (0.0018)	0.1350 (0.1726)	0.3350 (0.1475)	−0.3765 (0.2227)	−0.3119 (0.1797)	0.00561 (0.00077)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.5426 (0.1896)	−0.1411 (0.2167)	0.7893 (0.2818)	−0.4109 (0.1688)	0.0045 (0.0008)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	0.0032 (0.0019)	−0.0241 (0.1365)	0.2845 (0.1186)	0.1076 (0.1321)	−0.0892 (0.1830)	0.00572 (0.00130)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	0.0104 (0.2481)	−0.0709 (0.2241)	−0.5919 (0.2424)	0.7882 (0.1878)	0.0043 (0.0007)	
Switching Equation	β_{30}	β_{31}				
	−0.0549 (0.3339)	−6.1054 (4.7664)				
Likelihood:				345.92920		

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	16.69	0.0051
AR(1) test for Regime 1	0.14	0.7063
AR(1) test for Regime 2	3.40	0.0653
ARCH(1) test for Regime 1	3.82	0.0507
ARCH(1) test for Regime 2	12.91	0.0003
Omitted Markov Effects	0.86	0.3541
Test for equality of growth rates	-0.07	0.9464

Table 5.13. Estimated switching regressions model for France (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0015 (0.0018)	0.1227 (0.1603)	0.3394 (0.1488)	−0.3656 (0.2053)	−0.3423 (0.1576)	0.00564 (0.00074)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.5266 (0.1831)	−0.1155 (0.2295)	0.8095 (0.2371)	−0.4134 (0.1689)	0.0045 (0.0007)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	0.0025 (0.0015)	−0.0178 (0.1303)	0.2611 (0.1137)	–	–	0.00549 (0.00106)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	0.0238 (0.2086)	−0.0748 (0.2032)	−0.4797 (0.1727)	0.7666 (0.1481)	0.0044 (0.0006)	
Switching Equation	β_{30}	β_{31}				
	–	−6.6168 (4.2968)				
Likelihood:				345.52440		
LR test of restrictions:				0.8095[0.8472]		

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	15.78	0.0033
AR(1) test for Regime 1	0.06	0.8146
AR(1) test for Regime 2	4.36	0.0368
ARCH(1) test for Regime 1	3.48	0.0622
ARCH(1) test for Regime 2	12.76	0.0004
Omitted Markov Effects	-	-
Test for equality of growth rates	0.11	0.9092
Test of significance of the $\Delta\overline{y}_{i,t-j}$'s	25.67	0.0073
Equal dynamics test	7.27	0.6094

Characteristic roots

Regime 1	Regime 2
$0.66 + 0.60i$	-0.52
$0.66 - 0.60i$	0.50
$-0.60 + 0.28i$	
$-0.60 - 0.28i$	

Table 5.14. Estimated switching regressions model with equal dynamics and constant switching probability for France (SE's in brackets).

Regime 1, 2	β_0	β_1	β_2	β_3	β_4		**Growth**
	-0.0008 (0.0006)	-0.1421 (0.0468)	0.3551 (0.0517)	0.0099 (0.0462)	-0.1036 (0.0371)		0.00562 (0.00046)
	γ_1	γ_2	γ_3	γ_4	σ_1	σ_2	
	0.8080 (0.0791)	-0.1165 (0.0595)	0.0060 (0.0656)	0.1342 (0.0694)	0.0080 (0.0008)	0.0010 (0.0003)	
Switching Equation	β_{30}	β_{31}					
	-0.5142 (0.2170)	– –					
Likelihood:			341.62084				
LR test of restrictions:			8.6160[0.5689]				

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	7.40	0.1161
AR(1) test for Regime 1 and 2	3.26	0.0712
ARCH(1) test for Regime 1	1.22	0.2692
ARCH(1) test for Regime 2	2.14	0.1435
Omitted Markov Effects	1.06	0.3031

Characteristic roots
Regime 1 and 2
$-0.54 + 0.25i$
$-0.54 - 0.25i$
$0.47 + 0.28i$
$0.47 - 0.28i$

Table 5.15. Estimated unrestricted switching regressions model for West Germany (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0127 (0.0019)	−0.3491 (0.0472)	0.0196 (0.0591)	−0.0235 (0.0717)	0.0991 (0.0730)	0.01177 (0.00082)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	−0.8836 (0.1589)	1.1933 (0.1670)	−0.0669 (0.2249)	0.0532 (0.1656)	0.0016 (0.0003)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	−0.0010 (0.0018)	−0.0353 (0.1928)	−0.3838 (0.1363)	−0.0476 (0.1388)	0.1004 (0.1194)	0.00389 (0.00077)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	0.7174 (0.3066)	0.2910 (0.2163)	0.1124 (0.2042)	−0.2164 (0.1943)	0.0070 (0.0006)	
Switching Equation	β_{30}	β_{31}				
	1.4381 (0.3840)	−5.0927 (1.6660)				
Likelihood:				332.04456		

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	9.98	0.0759
AR(1) test for Regime 1	5.01	0.0253
AR(1) test for Regime 2	0.02	0.8945
ARCH(1) test for Regime 1	1.74	0.1874
ARCH(1) test for Regime 2	0.00	0.9663
Omitted Markov Effects	2.29	0.1299
Test for equality of growth rates	4.09	0.0000

Table 5.16. Estimated switching regressions model for West Germany (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0135 (0.0011)	−0.3410 (0.0478)	– –	– –	– –	0.01156 (0.00038)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	−0.9391 (0.1108)	1.2291 (0.1577)	– –	– –	0.0018 (0.0003)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	−0.0013 (0.0015)	−0.0558 (0.1873)	−0.3792 (0.1306)	– –	– –	0.00412 (0.00066)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	0.7411 (0.2600)	0.2987 (0.1868)	– –	– –	0.0071 (0.0006)	
Switching Equation	β_{30}	β_{31}				
	1.5023 (0.3763)	−5.1627 (1.6327)				
Likelihood:				328.99016		
LR test of restrictions:				6.109[0.7290]		

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	5.44	0.3645
AR(1) test for Regime 1	0.00	0.9708
AR(1) test for Regime 2	0.02	0.8870
ARCH(1) test for Regime 1	3.66	0.0559
ARCH(1) test for Regime 2	0.05	0.8270
Omitted Markov Effects	1.90	0.1679
Test for equality of growth rates	5.81	0.0000
Test of significance of the $\Delta\overline{y}_{i,t-j}$'s	24.11	0.0022
Equal dynamics test	20.40	0.0856

Characteristic roots

Regime 1	Regime 2
−0.34	$-0.03 + 0.62i$ $-0.03 - 0.62i$

Table 5.17. Estimated switching regressions model with equal dynamics for West Germany (SE's in brackets).

Regime 1, 2	β_0	β_1	β_2	β_3	β_4		**Growth**
	0.0004 (0.0002)	−0.0252 (0.0238)	−0.1490 (0.0232)	– –	– –		0.00491 (0.00013)
	γ_1	γ_2	γ_3	γ_4	σ_1	σ_2	
	0.3517 (0.0351)	0.4252 (0.0322)	– –	– –	0.0091 (0.0007)	0.0004 (0.0001)	
Switching Equation	β_{30}	β_{31}					
	−0.7366 (0.2163)	−5.9328 (2.1895)					
Likelihood:					321.84360		
LR test of restrictions:					20.4000[0.0856]		

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	0.75	0.9456
AR(1) test for Regime 1 and 2	0.29	0.5924
ARCH(1) test for Regime 1	0.10	0.7579
ARCH(1) test for Regime 2	0.29	0.5908
Omitted Markov Effects	0.21	0.6463

Characteristic roots
Regime 1 and 2
$-0.01 + 0.39i$
$-0.01 - 0.39i$

Table 5.18. Estimated unrestricted switching regressions model for Italy (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	-0.0054 (0.0013)	0.2603 (0.0554)	-0.0306 (0.0496)	0.2002 (0.0375)	0.1672 (0.0341)	-0.00046 (0.00311)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.8618 (0.1094)	-0.0296 (0.1345)	-0.2856 (0.1319)	0.2392 (0.1262)	0.0030 (0.0005)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	0.0011 (0.0015)	0.0684 (0.1483)	0.1919 (0.1114)	-0.6860 (0.1213)	-0.2229 (0.1204)	0.00676 (0.00047)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	-0.1130 (0.1680)	0.8948 (0.1755)	0.6458 (0.1807)	0.0670 (0.1705)	0.0048 (0.0005)	
Switching Equation	β_{30}	β_{31}				
	0.0605 (0.1946)	-3.3442 (1.8443)				
Likelihood:				346.28432		

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	24.16	0.0002
AR(1) test for Regime 1	14.09	0.0002
AR(1) test for Regime 2	0.74	0.3891
ARCH(1) test for Regime 1	13.88	0.0002
ARCH(1) test for Regime 2	5.39	0.0203
Omitted Markov Effects	0.03	0.8733
Test for equality of growth rates	-4.04	0.0001

Table 5.19. Estimated switching regressions model for Italy (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	-0.0018	0.4949	-0.1241	0.1858	$-$	0.00503
	(0.0016)	(0.0999)	(0.0680)	(0.0492)	$-$	(0.00248)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.6056	$-$	$-$	$-$	0.0046	
	(0.1351)	$-$	$-$	$-$	(0.0007)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	-0.0012	-0.0621	0.1573	-0.8037	$-$	0.00602
	(0.0016)	(0.1001)	(0.0829)	(0.1196)	$-$	(0.00049)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	-0.0573	1.1167	0.6378	$-$	0.0050	
	(0.1954)	(0.1926)	(0.2046)	$-$	(0.0006)	
Switching Equation	β_{30}	β_{31}				
	$-$	-4.2597				
	$-$	(2.1067)				
Likelihood:				339.82369		
LR test of restrictions:				12.9213[0.0741]		

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	8.55	0.0735
AR(1) test for Regime 1	0.37	0.5450
AR(1) test for Regime 2	2.78	0.0957
ARCH(1) test for Regime 1	2.88	0.0896
ARCH(1) test for Regime 2	3.08	0.0793
Omitted Markov Effects	-	-
Test for equality of growth rates	-0.40	0.6930
Test of significance of the $\Delta\overline{y}_{i,t-j}$'s	51.20	0.0000
Equal dynamics test	44.60	0.0000

Characteristic roots

Regime 1	Regime 2
0.70	-1.01
$-0.10 + 0.51i$	$0.47 + 0.76i$
$-0.10 - 0.51i$	$0.47 - 0.76i$

Table 5.20. Estimated unrestricted switching regressions model for Canada (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0123 (0.0030)	−0.3305 (0.1534)	−0.2683 (0.1606)	−0.2668 (0.1340)	0.1809 (0.1724)	0.01134 (0.00096)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.1870 (0.2819)	0.1922 (0.2552)	−0.3638 (0.2505)	1.0263 (0.4766)	0.0061 (0.0010)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	−0.0013 (0.0012)	0.5317 (0.0823)	0.2092 (0.0848)	0.1058 (0.0859)	−0.0670 (0.0716)	−0.00293 (0.00753)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	0.6649 (0.1680)	−0.1066 (0.1789)	−0.4420 (0.1732)	−0.0174 (0.1412)	0.0045 (0.0006)	
Switching Equation	β_{30}	β_{31}				
	−0.0295 (0.3416)	−4.1921 (1.9587)				
Likelihood:				330.58728		

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	4.81	0.4393
AR(1) test for Regime 1	0.11	0.7363
AR(1) test for Regime 2	2.28	0.1309
ARCH(1) test for Regime 1	0.28	0.5948
ARCH(1) test for Regime 2	0.26	0.6098
Omitted Markov Effects	3.01	0.0830
Test for equality of growth rates	3.38	0.0072

Table 5.21. Estimated switching regressions model for Canada (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	0.0109 (0.0030)	−0.2031 (0.1681)	−0.3914 (0.1614)	−0.1684 (0.1355)	– –	0.01139 (0.00095)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.2550 (0.2965)	0.2019 (0.2768)	−0.2747 (0.2557)	1.2399 (0.4826)	0.0065 (0.0011)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	−0.0013 (0.0012)	0.4909 (0.0809)	0.2621 (0.0877)	– –	– –	−0.00262 (0.00574)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	0.6590 (0.1743)	−0.1056 (0.1509)	−0.4618 (0.1541)	– –	0.0045 (0.0005)	
Switching Equation	β_{30}	β_{31}				
	– –	−4.3498 (1.3983)				
Likelihood:				329.04536		
LR test of restrictions:				3.083[0.6872]		

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	3.58	0.4656
AR(1) test for Regime 1	0.07	0.7902
AR(1) test for Regime 2	1.64	0.2005
ARCH(1) test for Regime 1	1.12	0.2893
ARCH(1) test for Regime 2	0.30	0.5836
Omitted Markov Effects	-	-
Test for equality of growth rates	4.10	0.0000
Test of significance of the $\Delta\overline{y}_{i,t-j}$'s	36.29	0.0003
Equal dynamics test	39.90	0.0000

Characteristic roots	
Regime 1	Regime 2
$0.08 + 0.67i$	0.81
$0.08 - 0.67i$	−0.32
−0.37	

Table 5.22. Estimated unrestricted switching regressions model for Japan (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	−0.0024 (0.0030)	0.0612 (0.2570)	0.4867 (0.2058)	0.0705 (0.1818)	0.1443 (0.1930)	−0.00195 (0.01369)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.9605 (0.4286)	−0.0247 (0.4757)	0.1822 (0.3815)	−0.8231 (0.3108)	0.0071 (0.0010)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	0.0065 (0.0022)	−0.1082 (0.1145)	−0.0264 (0.1210)	0.3242 (0.1266)	0.1465 (0.1156)	0.01072 (0.00147)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	0.0449 (0.1601)	−0.1392 (0.2220)	−0.4280 (0.2607)	0.6173 (0.1887)	0.0055 (0.0007)	
Switching Equation	β_{30}	β_{31}				
	1.1583 (0.4818)	−7.4789 (3.1663)				
Likelihood:				322.89516		
			Misspecification tests			
Test				Statistic		Significance
Joint test for Misspecification				2.09		0.8362
AR(1) test for Regime 1				0.83		0.3630
AR(1) test for Regime 2				0.47		0.4917
ARCH(1) test for Regime 1				0.05		0.8238
ARCH(1) test for Regime 2				0.02		0.8772
Omitted Markov Effects				0.51		0.4768
Test for equality of growth rates				-1.55		0.1207

Table 5.23. Estimated switching regressions model for Japan (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	**Growth**
	−0.0013 (0.0025)	0.0918 (0.1959)	0.5026 (0.1516)	– –	– –	0.00353 (0.004445)
	γ_{11}	γ_{12}	γ_{13}	γ_{14}	σ_1	
	0.6752 (0.3168)	0.0371 (0.3685)	0.3802 (0.2727)	−0.6612 (0.2776)	0.0072 (0.0009)	
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	**Growth**
	0.0094 (0.0026)	−0.2169 (0.1653)	−0.0323 (0.1150)	0.3940 (0.1230)	– –	0.01080 (0.00113)
	γ_{21}	γ_{22}	γ_{23}	γ_{24}	σ_2	
	−0.0030 (0.1592)	0.0442 (0.2106)	−0.8318 (0.3240)	0.7739 (0.1751)	0.0047 (0.0009)	
Switching Equation	β_{30}	β_{31}				
	0.6783 (0.3676)	−6.2296 (2.5187)				
Likelihood:				322.00552		
LR test of restrictions:				1.780[0.6194]		

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	3.33	0.6494
AR(1) test for Regime 1	1.12	0.2886
AR(1) test for Regime 2	0.81	0.3671
ARCH(1) test for Regime 1	0.87	0.3502
ARCH(1) test for Regime 2	0.66	0.4154
Omitted Markov Effects	0.28	0.5944
Test for equality of growth rates	-1.66	0.0969
Test of significance of the $\Delta\overline{y}_{i,t-j}$'s	-	-
Equal dynamics test	-	-

Characteristic roots

Regime 1	Regime 2
0.76	$-0.44 + 0.64i$
-0.66	$-0.44 - 0.64i$
	0.65

Table 5.24 . Estimated current account regression models (SE in brackets).

Dependent variable $\Delta_4 ca_{i,t}$	Country			
Regressor	USA	UK	FRA	GER
int	−1.0892 (0.5277)	−0.1567 (0.1805)	−0.0907 (0.1283)	0.3298 (0.2454)
$\Delta_4 ca_{i,t-1}$	0.9050 (0.0792)	0.4770 (0.0997)	0.6288 (0.0905)	0.6178 (0.0645)
$\Delta_4 ca_{i,t-2}$	−0.1713 (0.0963)	0.1790 (0.1111)	0.0458 (0.0910)	– –
$\Delta_4 ca_{i,t-3}$	0.2732 (0.0940)	0.1694 (0.1108)	0.0687 (0.0909)	– –
$\Delta_4 ca_{i,t-4}$	−0.6662 (0.0964)	−0.3551 (0.1000)	−0.5787 (0.0916)	– –
$\Delta_4 ca_{i,t-5}$	0.4274 (0.0808)	– –	0.2958 (0.0943)	– –
$z_{i,t}$	−6.9043 (4.2174)	−2.5816 (1.3483)	−8.7277 (2.5631)	−2.3474 (1.4297)
	ITA	CAN	JPN	
int	−0.3188 (0.1641)	−0.1883 (0.1179)	0.6123 (0.3760)	
$\Delta_4 ca_{i,t-1}$	0.7285 (0.1004)	0.6800 (0.1047)	0.9031 (0.1037)	
$\Delta_4 ca_{i,t-2}$	0.0334 (0.1233)	0.0617 (0.1152)	0.0560 (0.1424)	
$\Delta_4 ca_{i,t-3}$	−0.0414 (0.1230)	−0.0659 (0.1163)	−0.2430 (0.1030)	
$\Delta_4 ca_{i,t-4}$	−0.3313 (0.1253)	−0.4975 (0.1171)	– –	
$\Delta_4 ca_{i,t-5}$	0.2761 (0.1115)	0.3455 (0.1078)	– –	
$z_{i,t}$	−6.7378 (1.7195)	−0.9667 (0.6810)	−4.0453 (2.2928)	

Figure 5.1. Growth rate of output for each of the G7 countries.

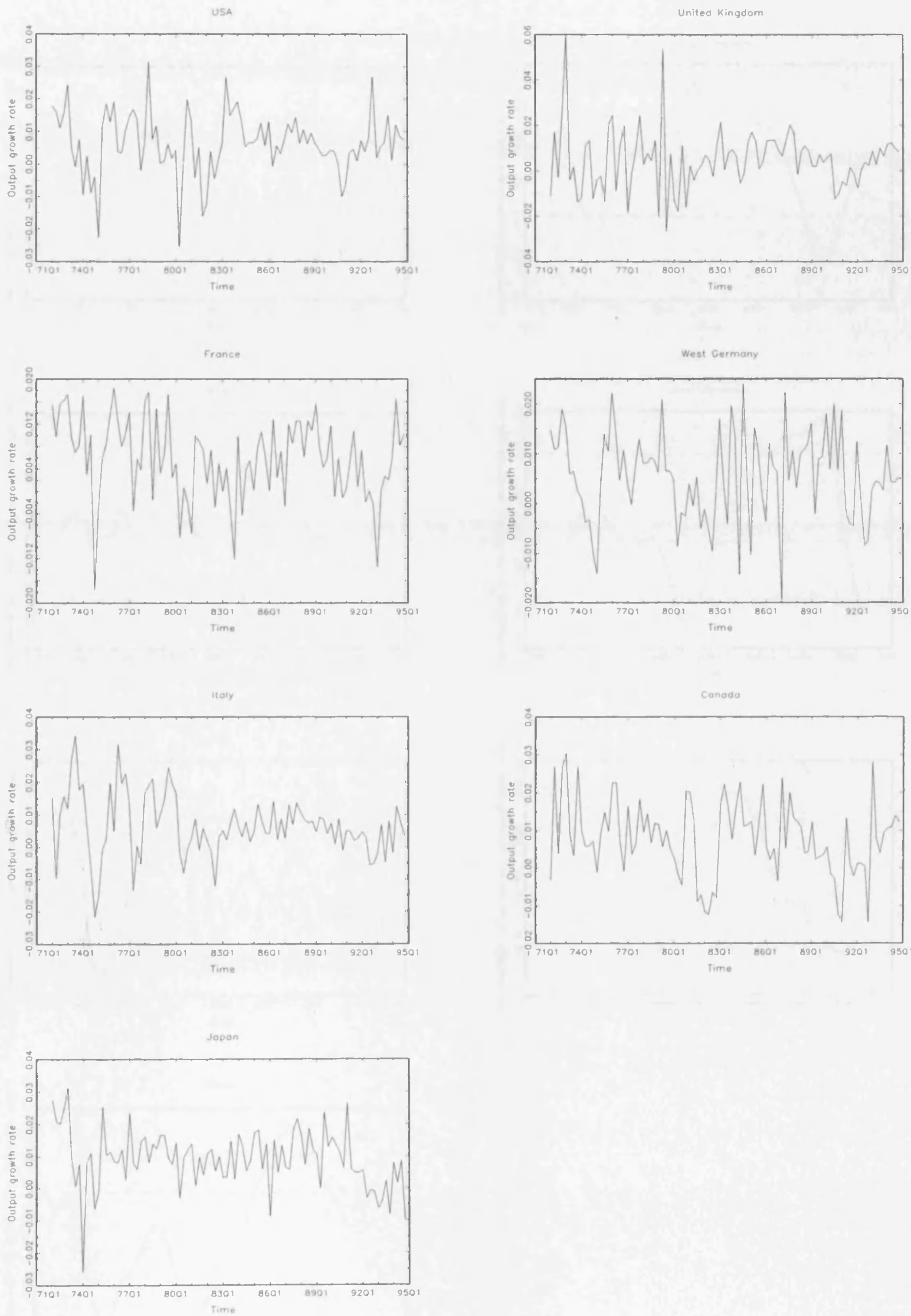


Figure 5.2. Accumulation over two years of the current account/output ratio for the G7 countries.

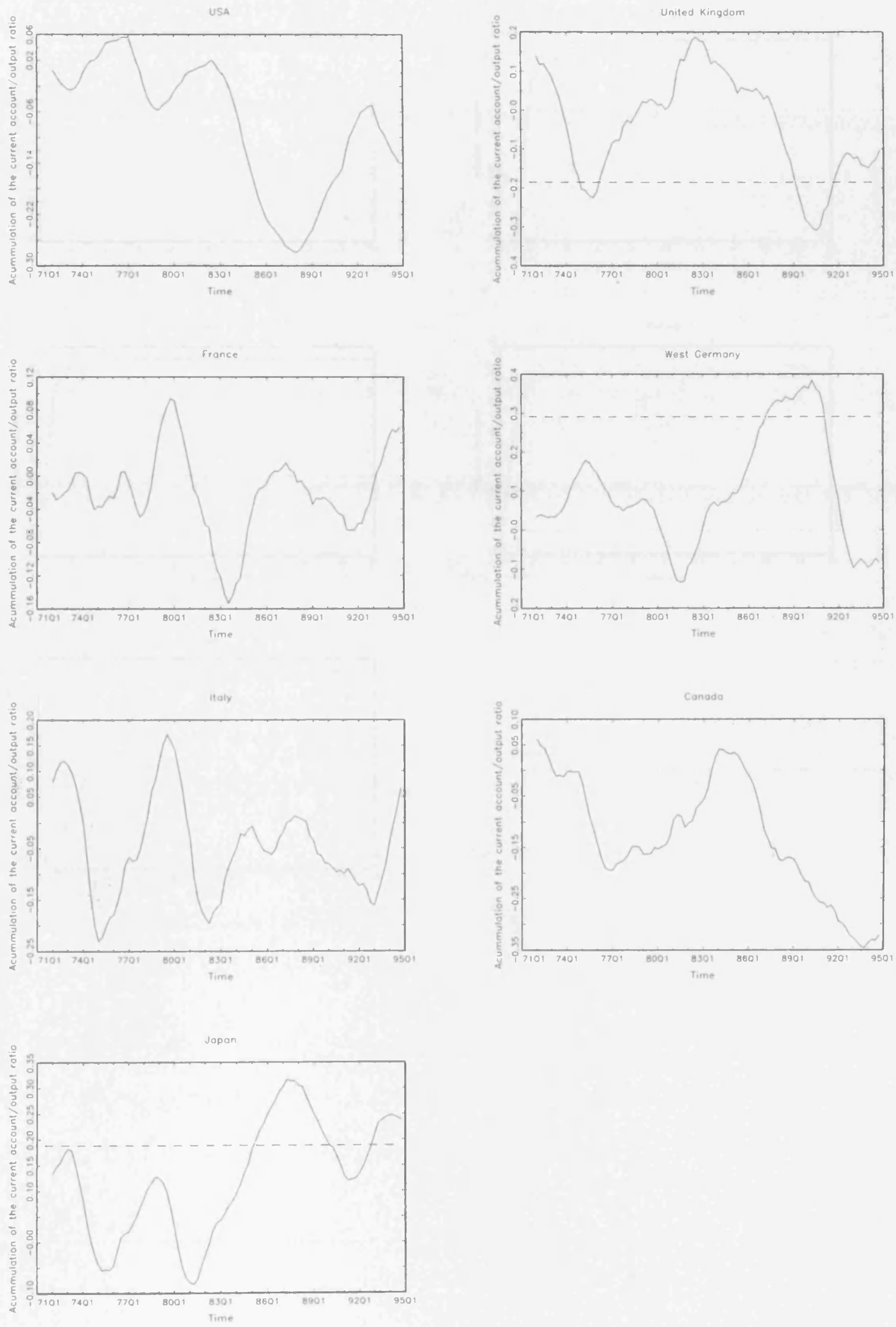


Figure 5.3. Impulse responses for the logarithm of output. Size of the shock equal to one standard deviation.

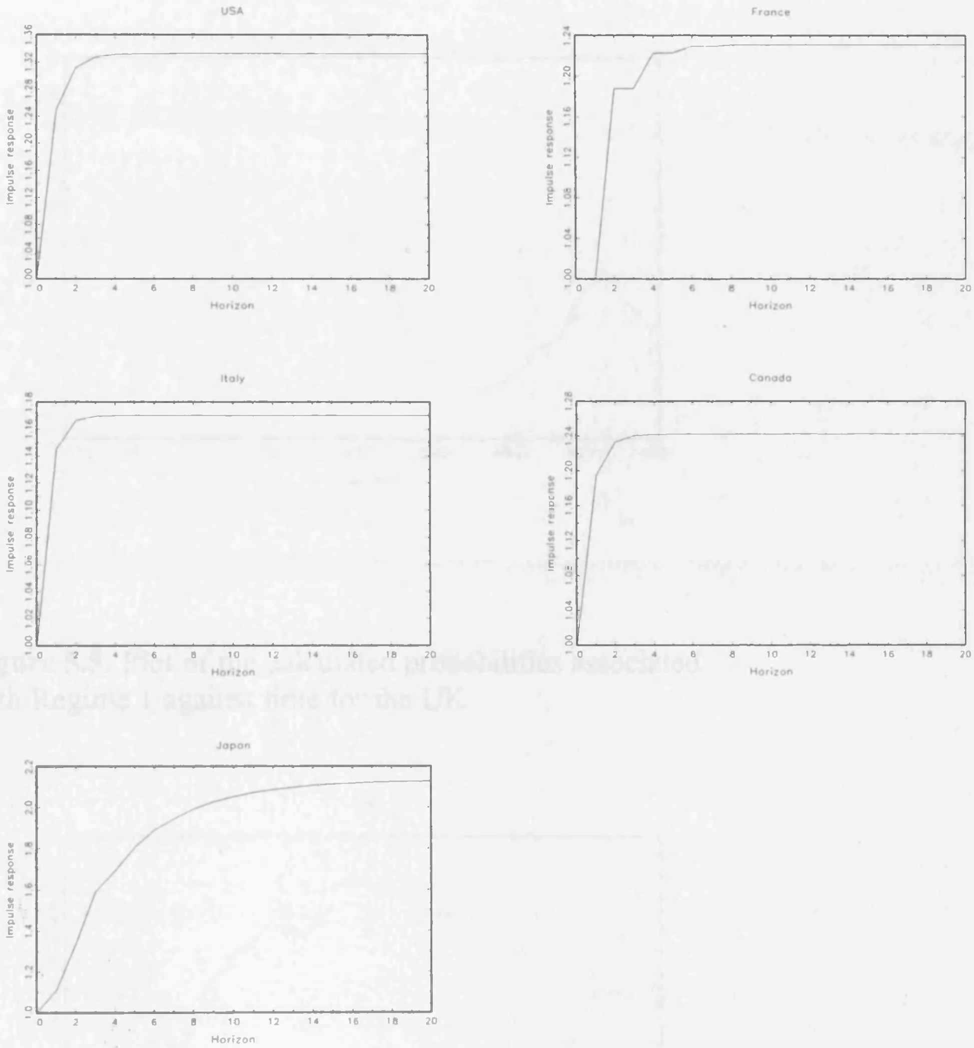


Figure 5.4. Plot of the calculated probabilities associated with Regime 1 against time for the USA.

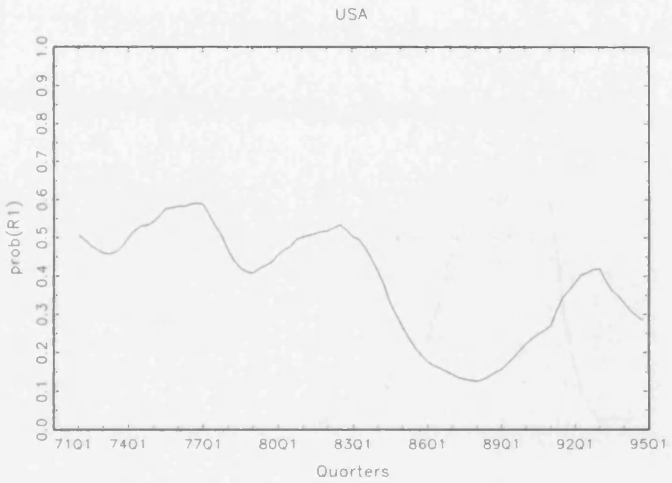


Figure 5.5. Plot of the calculated probabilities associated with Regime 1 against time for the UK.

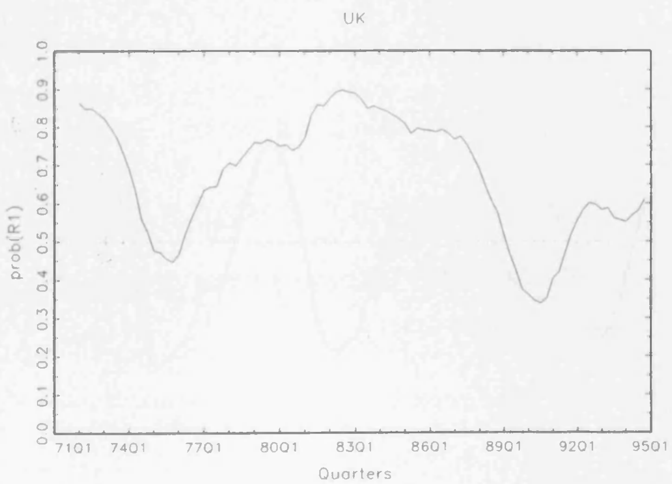


Figure 5.6. Plot of the calculated probabilities associated with Regime 1 against time for West Germany.

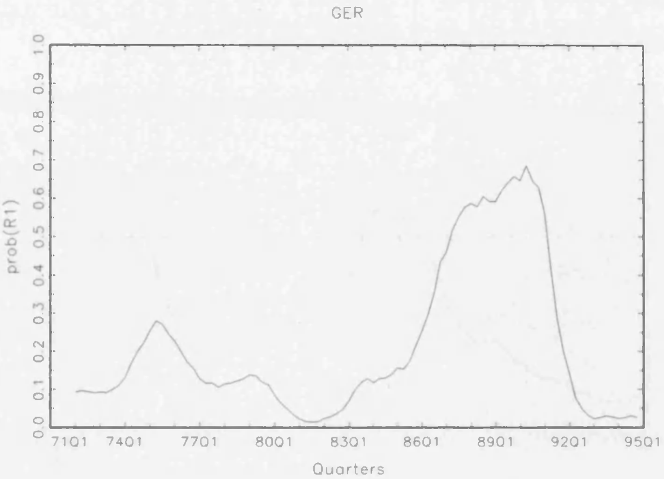


Figure 5.7. Plot of the calculated probabilities associated with Regime 1 against time for Italy.

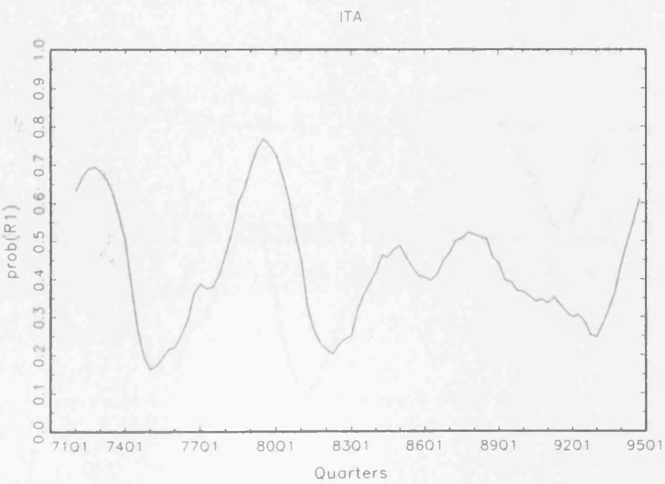


Figure 5.8. Plot of the calculated probabilities associated with Regime 1 against time for Canada.

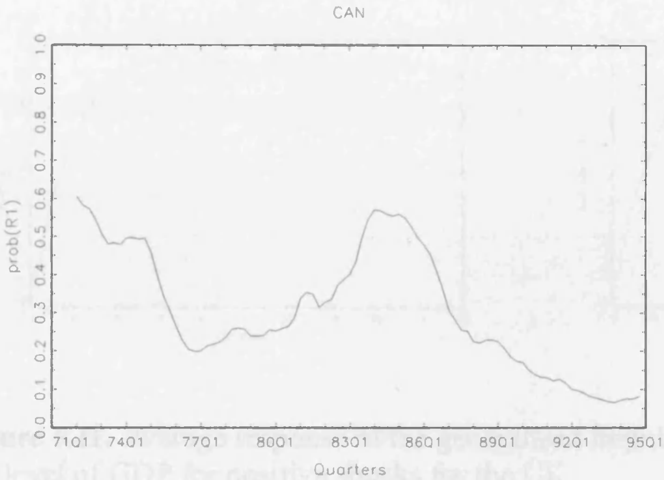


Figure 5.9. Plot of the calculated probabilities associated with Regime 1 against time for Japan.

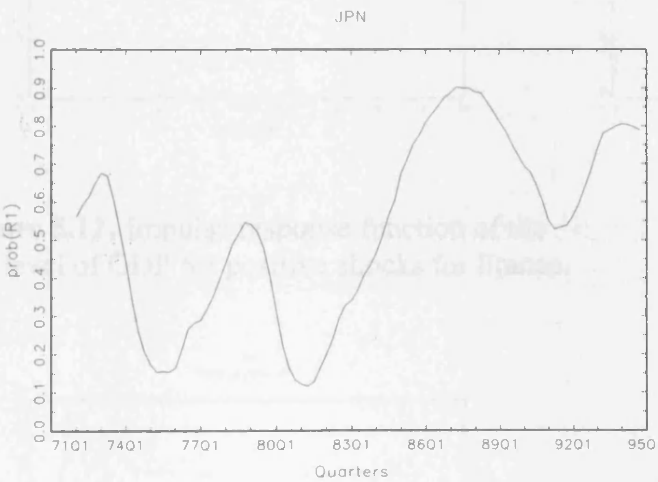


Figure 5.10. Average response of the generalized impulse response functions of the log level of GDP for positive shocks for the USA.

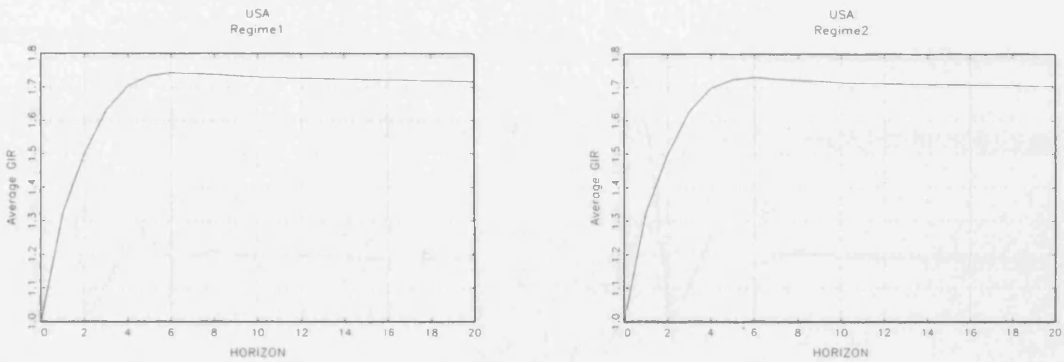


Figure 5.11. Average response of the generalized impulse response functions of the log level of GDP for positive shocks for the UK.

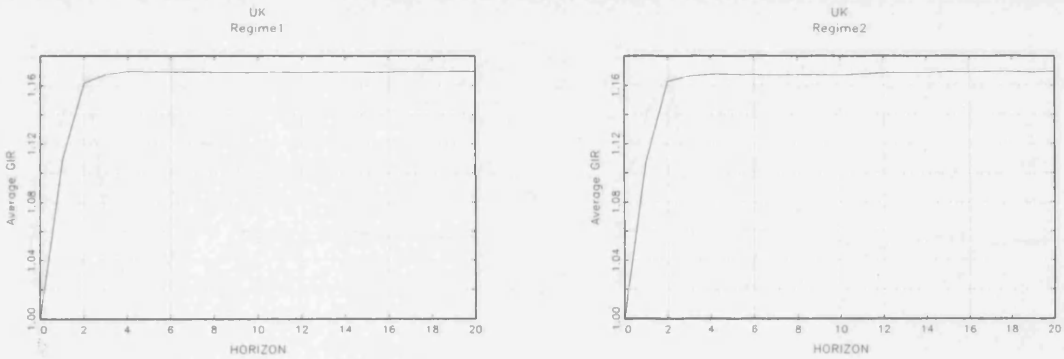


Figure 5.12. Impulse response function of the log level of GDP for positive shocks for France.

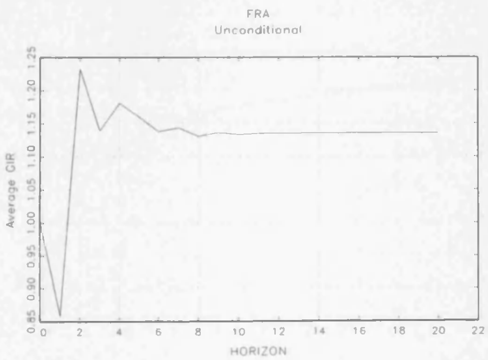


Figure 5.13. Average response of the generalized impulse response functions of the log level of GDP for positive shocks for West Germany.

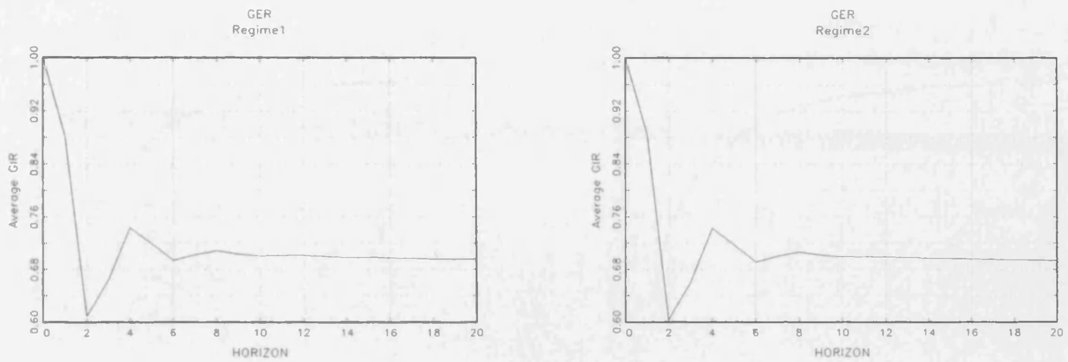


Figure 5.14. Average response of the generalized impulse response functions of the log level of GDP for positive shocks for Italy.

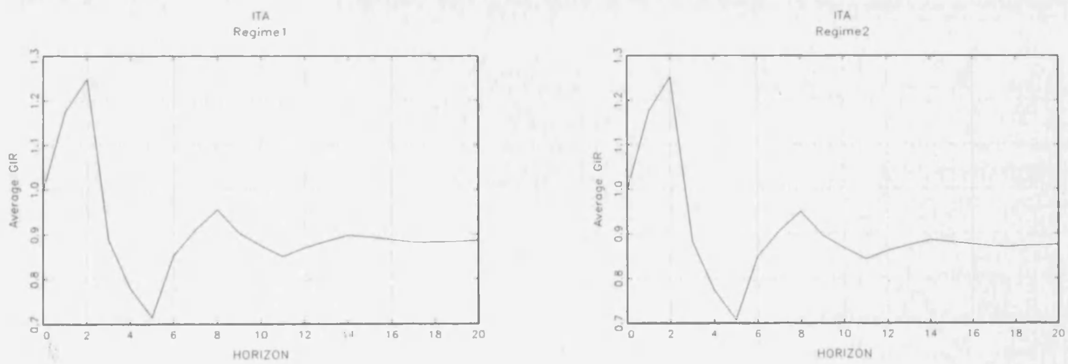


Figure 5.15. Average response of the generalized impulse response functions of the log level of GDP for positive shocks for Canada.

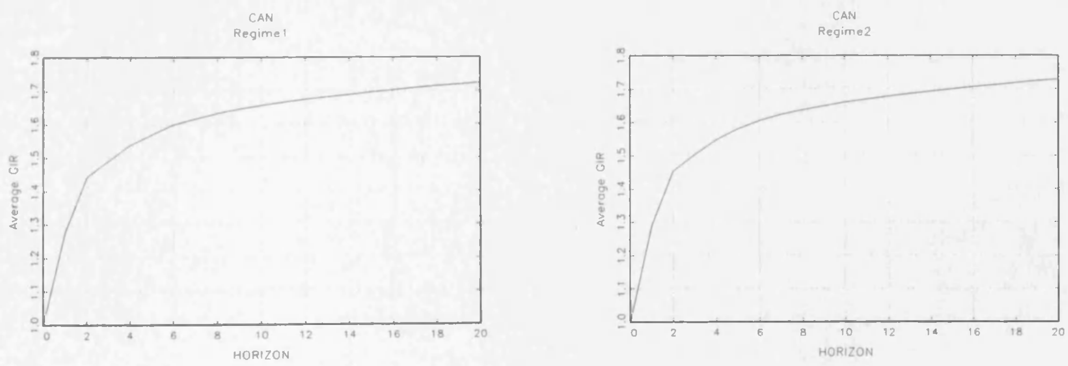


Figure 5.16. Average response of the generalized impulse response functions of the log level of GDP for positive shocks for Japan.

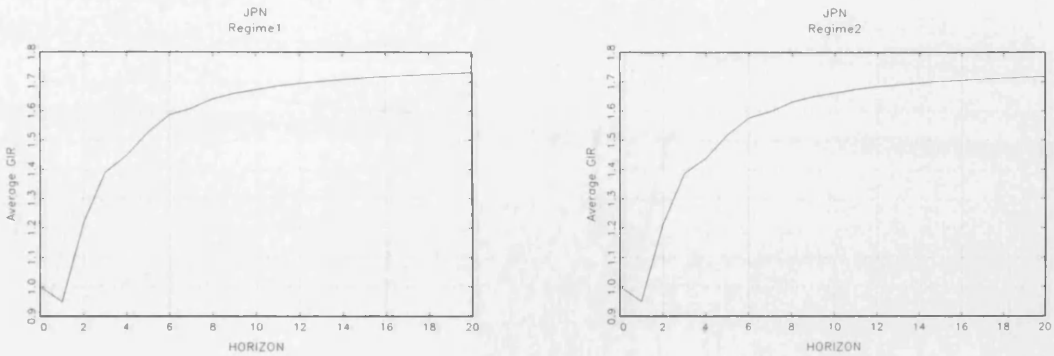


Figure 5.17. Individual responses of the log level of output for 2 different histories in each regime for the USA.

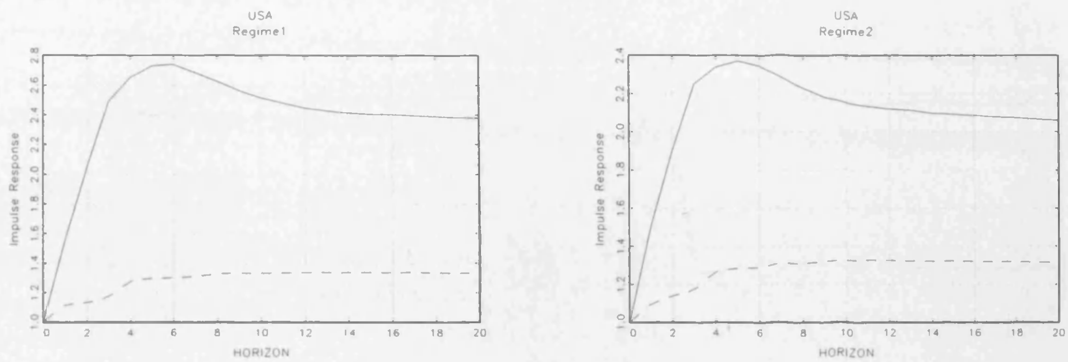


Figure 5.18. Individual responses of the log level of output for 2 different histories in each regime for the UK.

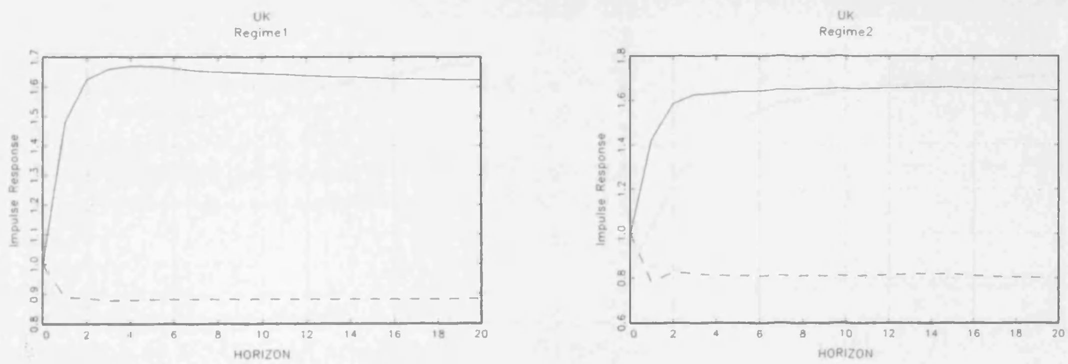


Figure 5.19. Individual responses of the log level of output for 2 different histories in each regime for West Germany.

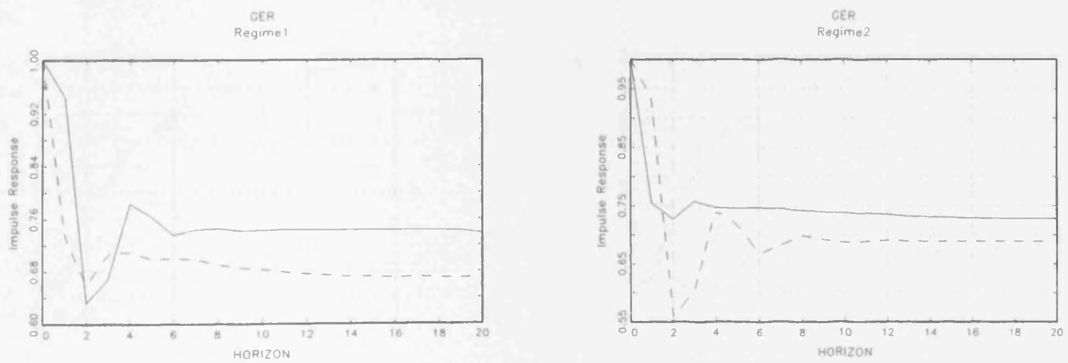


Figure 5.20. Individual responses of the log level of output for 2 different histories in each regime for Italy.

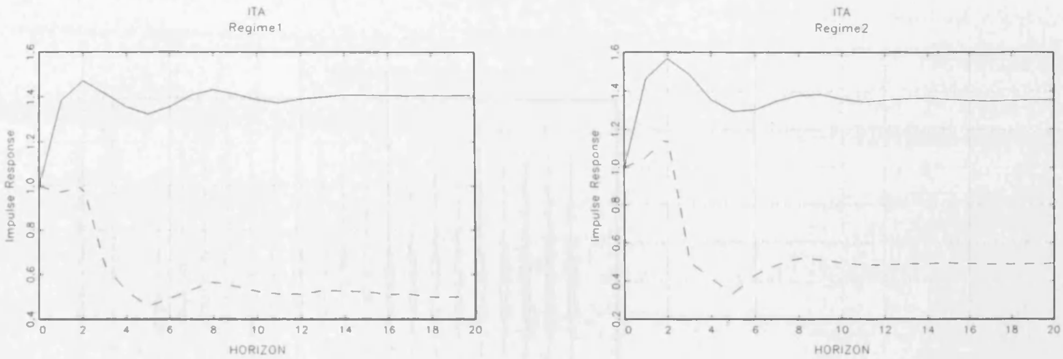


Figure 5.21. Individual responses of the log level of output for 2 different histories in each regime for Canada.

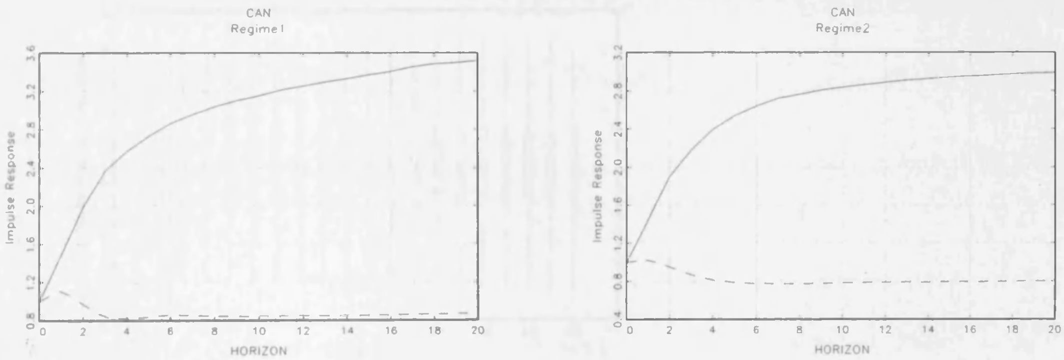


Figure 5.22. Individual responses of the log level of output for 2 different histories in each regime for Japan.

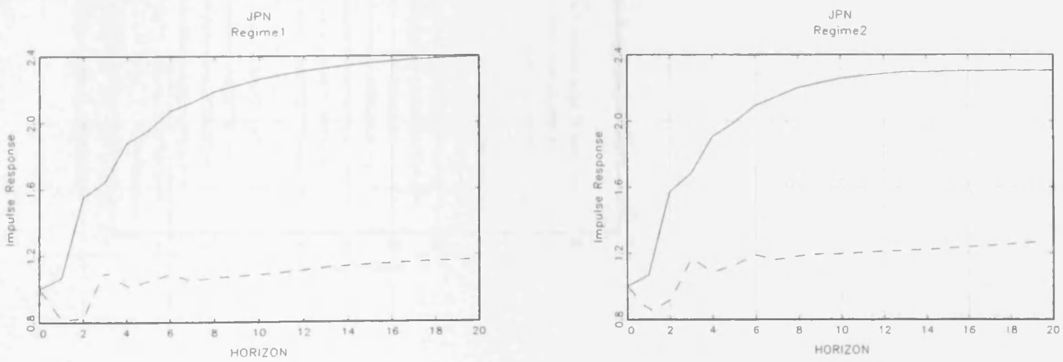
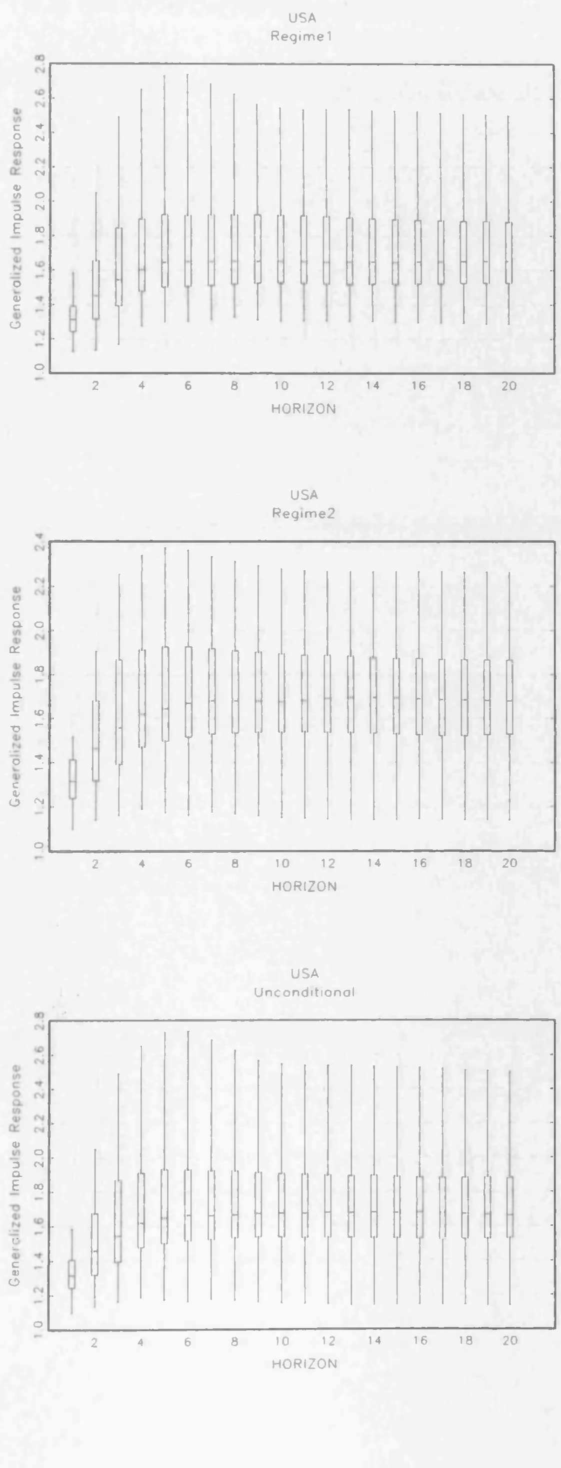
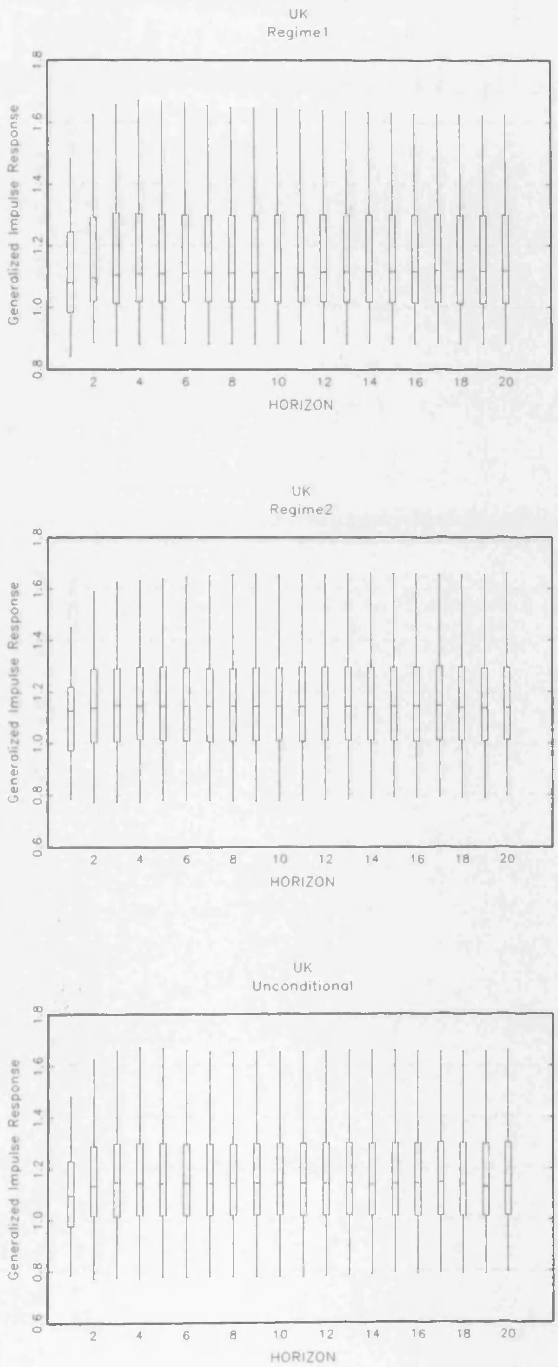


Figure 5.23. Box plot of the generalized impulse response functions of the log level of GDP for positive shocks for the USA[79].



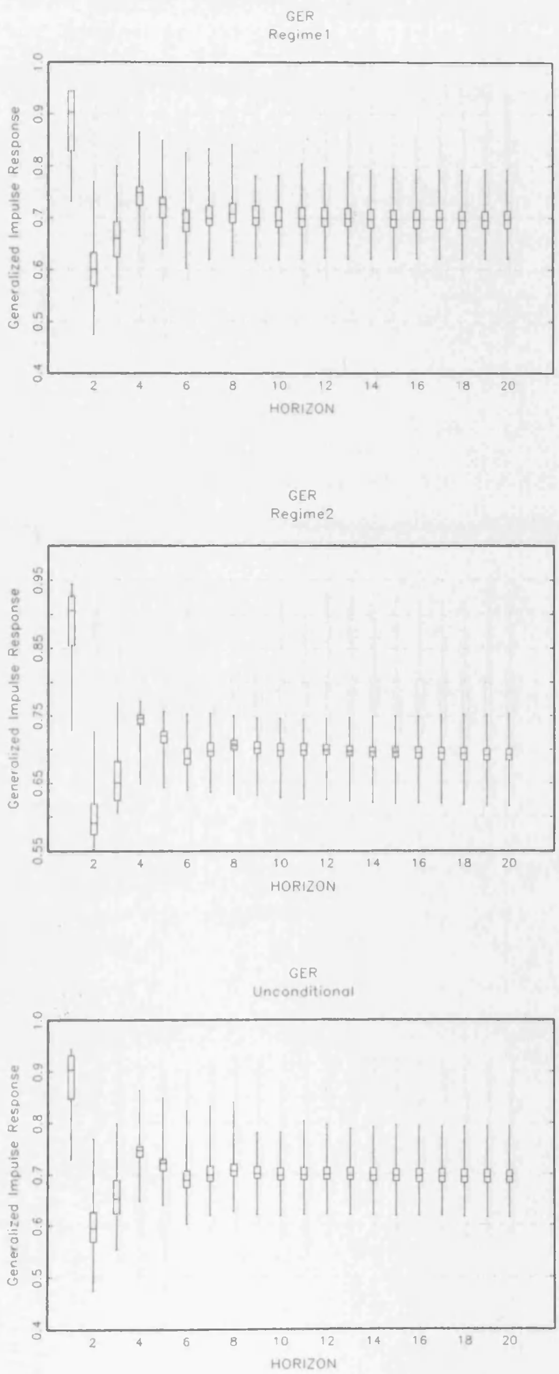
[79]These figures show the boxplot at each horizon for the generalized impulse response functions generated by a shock of size 1 and the associated histories.

Figure 5.24. Box plot of the generalized impulse response functions of the log level of GDP for positive shocks for the UK[80].



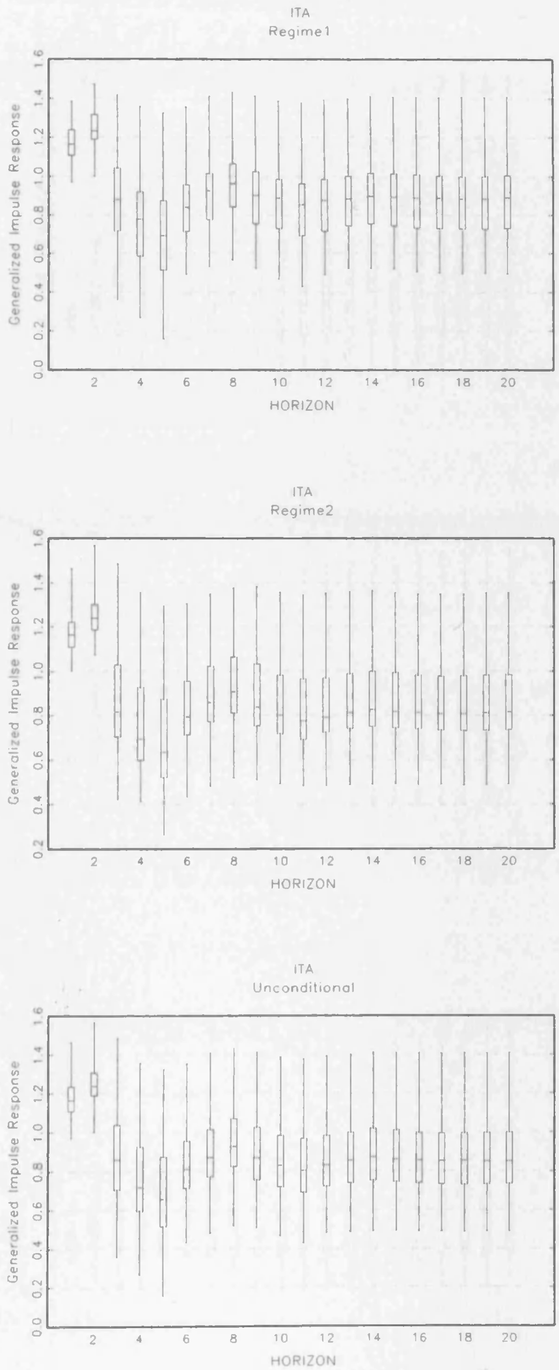
[80]See footnote for Figure 5.23.

Figure 5.25. Box plot of the generalized impulse response functions of the log level of GDP for positive shocks for West Germany[81].



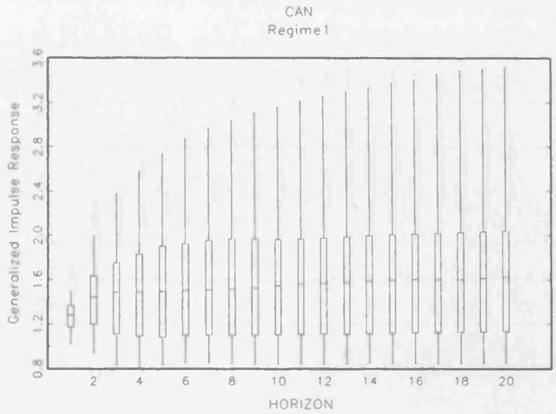
[81]See footnote for Figure 5.23.

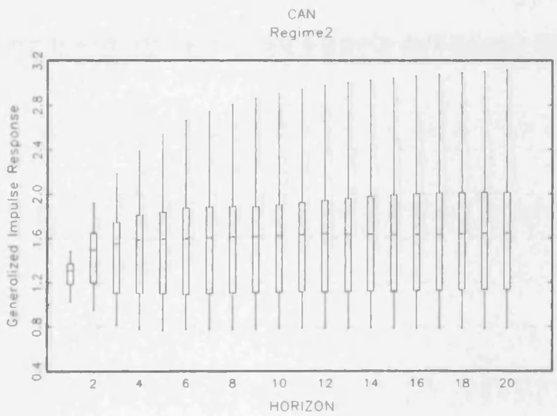
Figure 5.26. Box plot of the generalized impulse response functions of the log level of GDP for positive shocks for Italy[82].

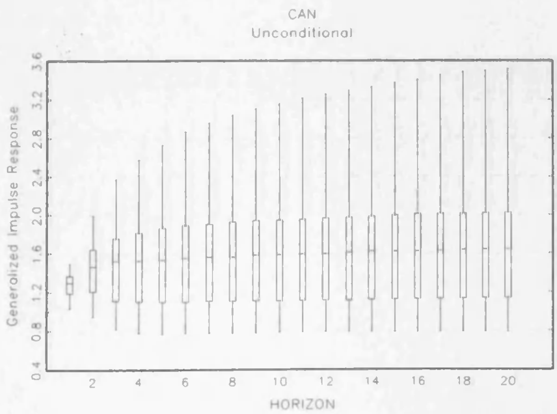


[82]See footnote for Figure 5.23.

Figure 5.27. Box plot of the generalized impulse response functions of the log level of GDP for positive shocks for Canada[83].

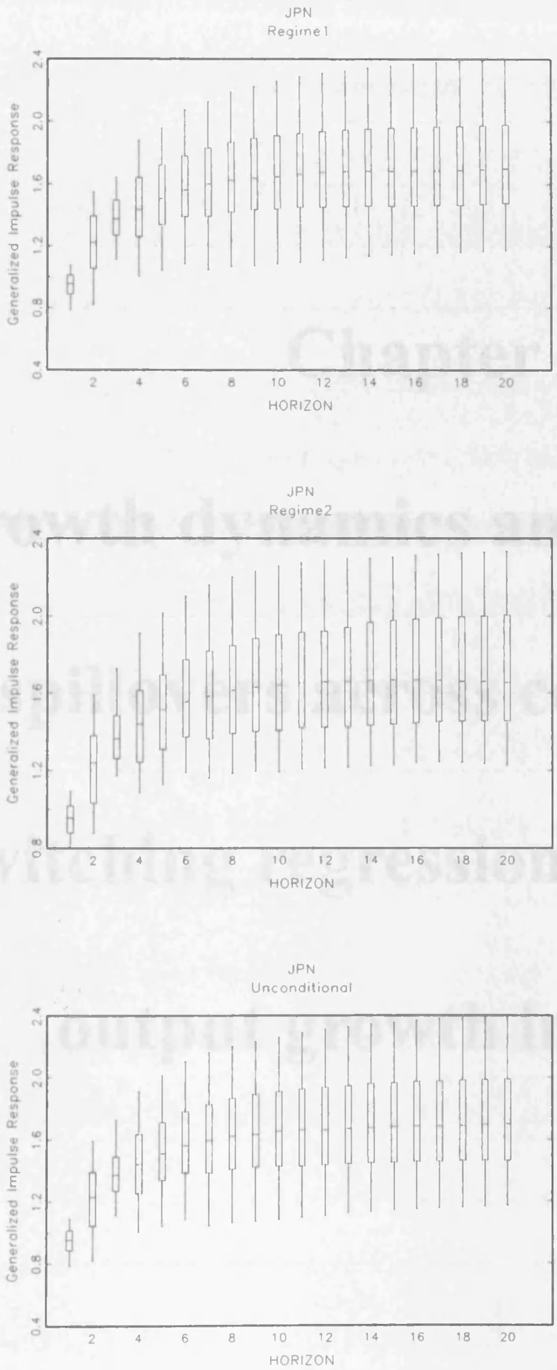






[83]See footnote for Figure 5.23.

Figure 5.28. Box plot of the generalized impulse response functions of the log level of GDP for positive shocks for Japan[84].



[84]See footnote for Figure 5.23.

Chapter 6

Growth dynamics and technology spillovers across countries; a switching regressions analysis of output growth in the G7.

6.1 Introduction.

In chapter 5 attention was restricted to the G7 countries with the aim of studying the process of growth in much more detail than in chapter 3, incorporating interrelationships across countries. Switching regressions models were estimated to test the assumption of a long run balance of payments constraint on growth. In addition, the role of technology spillovers across countries was investigated albeit in a simple form by using for each country the average growth rate in the rest of the countries as a proxy for the level of technology in the rest of the world. Little evidence was found of a long run balance of payments constraint on growth in the G7 countries. However, there was ample evidence supporting the role of technology spillovers across countries.

Based on this initial evidence, the present chapter aims to investigate this issue in more detail. To this avail, the theoretical model illustrated in chapter 2, namely that of section 2.4.2, is estimated using the nonlinear techniques highlighted in chapter 4. Most of the studies relating to technology spillovers concentrate only on modelling technology on the assumption that higher growth of technology will undoubtedly lead to a higher growth rate of output. In this chapter, the process of output growth as opposed to technology growth is directly modelled for the G7 countries with technology spillovers governing this process. The theoretical model in chapter 2 is a closed economy model in which technology is allowed to flow from more technologically advanced countries to less advanced countries; that is, technological advances in one country spill over to the rest of the countries. If this is the case, it was shown in chapter 2 that the growth of output might be generated by two different processes with very different short run dynamics. Countries are assumed to grow at a certain country specific rate which depends on the rate of innovation of the country, which in turn depends on, say, the amount of resources dedicated to research and development, etc. Since imitation is cheaper and quicker than innovation, countries with a relatively low technology level can grow relatively fast while this spillover is taking place. Once countries reach the steady state, the asymptotic long run growth rate of output will be the same as the closed economy model since, ultimately, the long run growth rate of output is determined by the asymptotic growth rate of technology. However, the level of output will be higher for the countries benefiting from spillovers than it would be for a closed economy country. This behaviour can be portrayed well by using a switching regime model of the type described in chapter 4.

The remainder of the chapter is organized as follows. Section 6.2 provides an overview of the data and detailed explanations of the measure of technology used in this chapter. Section 6.3 gives details of the country by country nonlinear econometric analysis of the technology spillovers models for each of the G7 countries. In section 6.4 the responses of output to different shocks are analysed using the adapted Generalized Impulse Response functions which were described in chapter 4. Finally, section 6.5 presents the final conclusions of this chapter.

6.2 Measures of technology and spillovers in the G7 countries; 1970q1-1994q4.

The data analysed in this chapter relates again to the G7 countries matching the sample of countries in chapter 5. It runs from the first quarter of 1965 to the last quarter of 1994. However, only the last 92 observations on each variable are used to maintain the time series comparable to that used in chapter 5. The analysis in this chapter requires both data on output and technology. The measure of technology used in this chapter is Total Factor Productivity (TFP). TFP measures the part of output that cannot be explained by changes in either labour input or capital input and, therefore, that unexplained component of output is attributed to technology. A more detailed account of the sources of the data for both output and technology as well as the computation of TFP is given in Appendix IV. Broadly, the measure of TFP employed here is based on the assumption of a constant returns to scale Cobb-Douglas production function of the form described in chapter 2, equation 2.1 but allowing for the shares of input to change over time. Taking logarithms on both sides of the equation and rearranging, the following measure of TFP ($a_{i,t}$) is obtained for each country i at each point in time t.

$$a_{i,t} = \frac{1}{1-\alpha_{i,t}} y_{i,t} - l_{i,t} - \frac{\alpha_{i,t}}{1-\alpha_{i,t}} k_{i,t}$$

where $y_{i,t}$, $l_{i,t}$ and $k_{i,t}$ denote the logarithms of output, labour input and capital input respectively for country i at time t and $\alpha_{i,t}$ is the share of labour in output for country i at time t.

Figure 6.1 plots these TFP measures for the G7 countries from the first quarter of 1970 to the last quarter of 1994. It is clear that the dispersion of the distribution of this measure of technology across the G7 countries is higher at the beginning of the sample than towards the end. It is also clear that the level of technology grows quicker for those countries which had lower technology levels at the beginning of the sample. This is consistent with the idea of spillovers of technology across countries from the most technologically advanced to those at the bottom of the distribution. There are several features that are apparent when inspecting this plot. Firstly, the measure of TFP for the USA is the highest at the beginning of the sample and it remains so up to the end of 1994. This result is common to many studies dealing with technology spillovers which consider the USA as the technological leader and the rest of the countries as followers (see chapter 2). Therefore, when considering spillovers of technology, these studies concentrate on the evolution of the technological gap between the rest of the countries and the USA. Second, at the other extreme of the distribution of technology across the G7 countries at the beginning of the period is Japan. The level of technology of Japan starts as the lowest of all the G7 countries in 1970 but it increases very rapidly. By the early eighties, Japan's technology level overtakes Italy, the second country from the bottom, converging towards the rest of the G7 countries. This is also very much in accordance with the assumption of technology spillovers. At the beginning of the seventies, the level of technology of Japan was so far behind compared to the rest of the G7 countries that it could have conceivably benefited enormously from technology spillovers from these countries and this could be the reason why the growth rate of its technology is so high at the beginning. The behaviour of the level of technology of Italy does not seem to follow this pattern of convergence so clearly. At the beginning of the sample, Italy's TFP level is very close to that of France and definitely higher than the level of Japan's technology. However, the TFP level of Italy seems to get further away from France, and, obviously, further away from the rest of the countries. In conclusion, this simple plot of the levels of technology across the G7 countries seems to provide some grounds to believe that technology spillovers across the G7 countries might play an important role.

Section 6.3 sets out to test this assumption by using a switching regressions model for output growth for each of the G7 countries in which switches across two different output growth processes depend on the size of the potential technology spillover available for each country at each point in time. For this purpose, a variable measuring the size of the potential spillover in each time period for each country has to be defined. The technical derivations of what follows here can be found in chapter 2, section 2.4.2 together with the derivations in Appendix I. Nevertheless, to help the discussion, the aspects relevant to the arguments presented here are reproduced below. To derive the variable measuring the size of the potential spillover of technology for each country i at time t, some distributional assumptions are needed with regards to technology. In chapter 2, it is assumed that technology at time t across countries follows a logistic distribution uniquely defined by two parameters, η_t and γ_t. These two parameters define the mean and the variance of the distribution of technology at time t. The mean is equal to $-\eta_t/\gamma_t$ and the variance is equal to $\pi^2\left(3\gamma_t^2\right)^{-1}$. It is also assumed that technology in each country evolves over time according to the following process

$$\Delta a_{i,t+1} = g_i + \lambda_i SPILL_{i,t} + \varepsilon_{i,t} \tag{6.1}$$

where g_i is the asymptotic rate of technology growth of country i, λ_i represents the speed of technology catch up, $SPILL_{i,t}$ measures the amount of technology available in the pool of technology that country i can acquire and $\varepsilon_{i,t}$ is the usual error term. For the purposes of the analysis in section 6.3, the variable of interest is, therefore,

$SPILL_{i,t}$, which was derived in chapter 2 and can be written as

$$SPILL_{i,t} = \frac{\sqrt{3}S_t}{\pi}\left(1+\phi_{i,t}\right)\ln\left(\frac{1+\phi_{i,t}}{\phi_{i,t}}\right)$$

where $\phi_{i,t} = \exp\left(\eta_t + \gamma_t a_{i,t}\right)$ and S_t is the standard deviation of the distribution of technology across countries at time t. An estimate of this variable for each country can be obtained by using sample averages. Figure 6.2 plots this variable for each of the G7 countries from the first quarter of 1970 to the last quarter of 1994. It is immediately obvious from this plot that the variables show a downward trend during the seventies before stabilising. It is evident that using these variables in their present form would pose a serious problem when using them as switching variables, since the model will most likely find one switch between processes at around the end of the seventies or the beginning of the eighties but without capturing the idea of continuous technology spillovers across countries which motivates this chapter. Therefore, an alternative variable is needed which conveys the same information as $SPILL_{i,t}$ but in a way that can be used as a switching variable in a switching regressions model. An obvious candidate for such a role is the area of the distribution of technology at time t that lies above the value of the logarithm of technology of country i at time t. Thus, the probability of a value of the logarithm of technology higher than the actual value for country i, can be defined as

$$z_{i,t}^{*} = \frac{1}{1+\phi_{i,t}} \tag{6.2}$$

that is, one minus the cumulative distribution of the logarithm of technology. Since $z_{i,t}^{*}$ is a probability, its values are restricted to lie between zero and one. A high value of $z^*_{i,t}$ represents a high likelihood of a spillover for country i at time t since the level of technology for that country is falling behind that of the rest of the countries in the distribution. Conversely, a small value of $z^*_{i,t}$ indicates a low probability of a spillover since the country in question is quite high in the distribution of technology at that point in time. Table 6.1 show some descriptive statistics of this variable, $z^*_{i,t}$, across time for each of the G7 countries. In Figure 6.1 it was clear that the technology of the USA was the highest of all the G7 countries throughout the entire sample period and, therefore, its potential for benefiting from spillovers was quite low. This is summarized in Table 6.1 by a low average value of $z^*_{i,t}$ for this country across time coupled with an also small standard deviation. The highest average of $z^*_{i,t}$ across time is that of Italy, closely followed by Japan. During the seventies, the value of $z^*_{i,t}$ fluctuates showing some evidence of technology spillovers taking place. However, for the rest of the sample period, the level of technology in Italy does not seem to follow this pattern, that is, it increases but at a lower rate than the rest of the countries leading to a higher and higher value of $z^*_{i,t}$ as time goes by, and, therefore, a higher probability of a spillover which does not seem to finally materialize in the sample period considered here. Japan has also a very high average value of $z^*_{i,t}$ across time. This is indicative of the potentially high technology spillovers from which this country could take advantage of in the sample considered here. The rest of the G7 countries, that is the UK, France, West Germany and Canada are more or less grouped together with the average values of $z^*_{i,t}$ in between 0.37 to 0.52 with the variable $z^*_{i,t}$ fluctuating across time.

This preliminary overview of the data on technology for the G7 countries seems to give weight to the assumption that the spillovers of technology across countries might be important. In the model of technology spillovers of chapter 2 the presence of spillovers of technology across countries has implications on the growth rate of output of the countries and, therefore, it is important to try to establish in a more rigorous analysis whether spillovers of technology in fact influence the growth process in each of the G7 countries. The next section addresses this issue.

6.3 Output growth analysis in the G7; a nonlinear analysis incorporating technology spillovers.

In this section, the switching regressions model of chapter 4 is estimated. That is, the model corresponding to equation (4.1) under the assumption that the errors follow the joint distribution of equation (4.2). For clarity of exposition, the model is reproduced below.

$$\begin{aligned} \Delta y_{i,t} &= \begin{cases} \beta_{i,1,0} + \sum_{j=1}^{k_{i,1}} \beta_{i,1,j} \Delta y_{i,t-j} + \varepsilon_{i,1,t} & \text{if } y_{i,t}^{*} < 0 \\ \beta_{i,2,0} + \sum_{j=1}^{k_{i,2}} \beta_{i,2,j} \Delta y_{i,t-j} + \varepsilon_{i,2,t} & \text{if } y_{i,t}^{*} \geq 0 \end{cases} \\ & \text{and} \\ y_{i,t}^{*} &= \beta_{i,3,0} + \beta_{i,3,1} z_{i,t-1} + \varepsilon_{i,3,t} \end{aligned} \tag{6.3}$$

where $y_{i,t}$ is the logarithm of output for country i at time t and $z_{i,t}$ is defined as the variable $z_{i,t}^{*}$ of last section demeaned across time for each country. Thus, output growth in this model is generated by two different processes or regimes, one corresponding to the case when each country's output is growing at their own pace, and the other corresponding to the case of technology spillovers from the rest of the countries. This model allows for different dynamics in both regimes and also allows for different growth rates under both regimes. Once the model is estimated, the equality of both the dynamics and the growth rates in both regimes will be tested explicitly.

The lower the level of the logarithm of technology for a country with respect to the rest of the countries in the distribution, the larger the potential spillover. As a result, the probability of the economy being in the regime in which the spillovers take place (a high growth rate regime), is higher. Similarly, the higher the level of the logarithm of technology of a country with respect to the rest of the countries in the distribution, the smaller the potential spillover. If this is the case, the country is likely to be growing as a result of its own efforts and therefore being in the low growth regime. Assuming that $\beta_{3,1} < 0$ and disregarding the value of $\beta_{3,0}$ for the moment, a potential spillover higher than the mean across time for a country will translate on the probability of being in the first regime,$\Phi\left(-\beta_{3,1} z_{i,t-1}\right)$, higher than 0.5. Therefore, the regime with a higher growth rate should be the first, while the second regime would be the one with a normal growth rate.

Sections 6.3.1 to 6.3.7 present a detailed account of the estimated switching regressions models of equation (6.3) for each of the G7 countries. The mechanics of estimating these models have already been highlighted in section 5.3 together with a thorough discussion of the problems involved. The models here are somewhat easier to estimate because of the reduced number of parameters, only 14 compared to 22 in chapter 5.

The estimated switching regressions models in section 6.3.1 to 6.3.7 corroborate the hypothesis of technology spillovers across the G7 countries. The variable measuring the potential spillover is found to be significant in five out of seven countries. It is only insignificant for the USA and Japan. For these two countries, however, the measure used here to indicate the potential for spillovers of technology across countries might not be an adequate choice. Since the measure of technology adopted in this chapter (following the bulk of the literature) gives a measure of technology for the USA that is higher than the rest of the countries for all the sample period,[85] it is evident that the model will not be able to pick up technological spillovers from the rest of the G7 countries. This is not to say that the USA does not benefit from technology spillovers from other countries. Some industries will obviously benefit from them, but the technology measure utilised here is an aggregate measure, which will not pick up small spillovers at a disaggregate level. Similarly, it is found that for Japan, the measure of technology spillovers might not serve its intended purpose. At the beginning of the sample the technology measure for Japan was by far the lowest amongst the G7 countries, but it started to grow at a very fast pace during the sample period. Since the level of technology of Japan was so far behind the rest of the countries in this sample, it might be the case that spillovers of technology from economies outside the G7 countries were important. This will help explain the high growth of Japan during the sample period considered here. In these circumstances, a measure of the potential spillover based only on the G7 countries will not perform well. In addition, it was also found that for Canada the effect of spillovers from the USA outweighed the effect of the spillovers from the rest of the G7 countries. Thus it is quite likely that because of their proximity, Canada followed the technological evolution of the USA more closely.

The following sections present all the details of the country by country estimation of the switching regressions models for the G7 countries. First, the unrestricted models for each country are estimated. The maximum lag of output growth for the regime regressions was set to four in the unrestricted models since the data are quarterly. These unrestricted switching regressions models are tested against a certain linear alternative.

[85]Even though the effect of natural resources in the measure of technology was tried to be kept to a minimum by adjusting output through the elimination of the value added of the mining and quarrying industry when the technology measure was calculated (see the data Appendix to this Chapter).

The linear alternative selected for this purpose is nested within the switching regressions model. It portrays two regimes with identical dynamics but with different variances. In effect, this is just a linear model with a heteroskedastic error term which ought to be rejected if the extra dynamics provided by the unrestricted switching regressions model are of importance. After some preliminary conclusions, the number of lags in the preferred specification is selected for each country according to AIC if there is no misspecification in the model. A set of tests is then carried out as detailed in chapter 4, comprising diagnostic tests for the presence of AR and ARCH effects in each individual regime and Markov effects in the switching equation as well as a general misspecification test for the switching regressions model. A further test for the equality of calculated growth rates in both regimes is also included since different dynamics in both regimes do not necessarily lead to statistically different growth rates.

Before studying the impulse responses of these models which are the subject matter of section 6.4, the convergence properties of the time paths of the two different regimes of the switching regressions models are also established by examining the roots of the characteristic polynomials associated with each of the regimes.

For clarity of exposition, the country subscript, i, will only be used from this point onwards in cases for which misinterpretation is likely to arise.

6.3.1 USA.

Table 6.2 presents the parameter estimates of the unrestricted switching regressions model for the USA. The model is well specified, but the estimated coefficient of the switching variable (z_t) is not significant. Another problem with this model is that the estimated growth rate in the high growth regime, that is, regime 1, is actually smaller than that of the second regime, although they are not significantly different from each other. Although the model seems to be well specified in terms of the diagnostic tests, it is signalling that the measure of technology spillovers employed here is not adequate for the USA. To ensure the most parsimonious model for the USA, this unrestricted switching regressions model is tested against a model with equal dynamics in both regimes but different variances. The joint test of the five restrictions from the unrestricted model of Table 6.2 cannot be rejected with a LR test of 7.8394 and an associated probability equal to 0.1653. It was also found that the AIC value of this model was higher than that of the unrestricted regime switching with a value of 301.1542 compared to 300.0739 for the unrestricted switching regressions. This shows that the appropriate model for the USA is in fact linear, but with a especial kind of heteroskedasticity in the error terms: the errors in the first regime have higher variance than the errors in the second regime but the means are the same in both regimes. Out of all the different specifications with equal dynamics, AIC chooses a model with just one lag of output growth and without the inclusion of the switching variable with a AIC value of 303.67344 (the LR test from the unrestricted switching regressions model has a value of 10.8010 [0.2896]). However there is evidence of ARCH effects of order 1 in the first regime with a test statistic equal to 4.1329 and an associated probability equal to 0.0421. Thus, the final model for the USA is shown in Table 6.3 with two lags of the growth rate. The model is well specified and the LR test of the restrictions from the switching regressions model with four lags is equal to 9.6099 [0.2935]. The long run growth rate is equal to 0.0071 per quarter and the probability of being in regime 1 is constant and equal to 0.4769, that is, there is almost an equal probability of being in each regime. The characteristic polynomial of this model has two real roots with modulus less than one (see Table 6.22) and therefore the model is stable, that is the time path is convergent.

In conclusion, the measure of technology spillovers used here does not seem to be adequate in the case of the USA. Since the level of technology of this country was above the levels of technology of the rest of the G7 countries throughout the whole sample period, the measure used here will show very small potential for spillovers from the rest of the G7 economies. However, this does not necessarily mean that at an industry level there are no spillovers of technology from the rest of the countries towards the USA it just happens that the aggregate measure employed in this chapter cannot capture these effects occurring at a disaggregate level.

6.3.2 United Kingdom.

The parameter estimates of the unrestricted switching regressions model for the UK are shown in Table 6.4. The model passes the misspecification test, however, there is evidence of omitted markov effects with a test statistic and significance equal to 6.7245 and 0.0095 respectively. The coefficient of the switching variable (z_t) is significant; the coefficient is equal to -10.5306 and its standard error is 3.7205. The hypothesis of equal dynamics in both regimes but different volatility is also tested. The LR statistic is 17.5803 [0.0035] which clearly rejects such a model in favour of the switching regressions model in Table 6.4.

Table 6.5 shows the estimates of the specification chosen by AIC; the LR test accepts the restrictions needed from the unrestricted model with a value of 5.5134 [0.3565]. This model is well specified and, again, the coefficient of the switching variable is significant with a value of -12.5458 and standard error equal to 4.7170. The two regimes have different dynamics, in fact, the second regime has no lags of the output growth rate whereas the first regime has all four lags of the dependent variable. The long run growth rate in the first and the second regimes are estimated to be 0.912% and 0.310% per quarter respectively. However, the hypothesis of equality of growth rates cannot be rejected at 5% significance level, although it is rejected at 10%. In the case of the UK, the probability of being in the first regime is higher than 0.5 if the value of the switching variable is bigger than zero. Therefore, the probability of being in the first regime is higher than the probability of being in the second regime if the probability of a value of the logarithm of technology higher than the actual value is bigger than the mean probability across time, which in this case is equal to 0.3721. The plot of the calculated probabilities for each observation of being in the first regime can be seen in Figure 6.3. From this figure, it is clear that during the early seventies, the level of technology in the UK was quite high in the distribution across the G7 countries. However, towards the end of the seventies, the level of technology in the UK fell behind the rest of the G7 countries, so much that in 1977 the probability of being in the spillover regime started to rise dramatically. This probability is quite high during the eighties, so that in this period, almost certainly, spillovers of technology were taking place in the UK. By the early nineties, even though the probability of being in the first regime is generally higher than the probability of being in the second regime and, consequently, there is a high probability of the UK absorbing technology from the rest of the G7 countries, the contribution of the second regime is higher than what it was in the eighties.

The time path of the first regime is convergent: its characteristic polynomial has four roots, a pair of complex roots and two real roots which are less than one in absolute value. The second regime is obviously stable since the growth rate of output is a constant.

To conclude, in this section evidence is found of technology spillovers from the rest of the G7 countries towards the UK using a switching regressions model. These spillovers of technology were more likely to have occurred during the eighties, although they were also probable during the early nineties.

6.3.3 France.

Table 6.6 shows the unrestricted switching regressions model estimated for France. The model is well specified passing not only the joint misspecification test but also all the individual misspecification tests. Both regimes have different growth rates in a manner consistent with the theory and the coefficient on the switching variable is significant (the coefficient is equal to -9.2991 and its standard error equals 3.5823). The dynamics of the two regimes seem to differ and, in fact, the restriction of equal dynamics but different variance in both regimes is openly rejected; the LR test has a value of 13.3132 [0.0206].

The final model chosen by AIC is presented in Table 6.7. This model has one lag of output growth in the first regime and two in the second regime. The restrictions from the unrestricted model cannot be rejected with a LR test equal to 3.8261 [0.7002]. The model is well specified, the null hypotheses of all the misspecification tests cannot be rejected at any standard significance level. Again the coefficient of the switching variable is significant with a value of -6.6460 and standard error equal to 2.4072. The growth rates in both regimes are different and in accordance to the theory. The first regime shows a growth rate equal to 1.267% per quarter whereas the second regime has a growth rate of 0.095% per quarter. The hypothesis of equal growth rates in both regimes is clearly rejected with a value of 3.5014 and significance 0.0002. Once more, since the constant in the switching equation is not significant, the probability of being in the first regime will be higher than the probability of being in the second regime if the probability of a value of the logarithm of technology higher than the actual value is bigger than the mean probability across time, which in this case is equal to 0.4987. Figure 6.4 plots the calculated probabilities of being in the first regime for each observation in the sample. At the beginning of the seventies, the probability of being in the spillovers regime was very high and increasing but in 1974 this probability started to decline, so much, that by the beginning of 1979 the probability of being in either of the regimes was roughly equal. This pattern continues until the firs quarter of 1983 in which the probability of being in the spillovers regime is at its lowest. From this point onwards, the level of technology in France seems to slowly fall behind in the distribution of technologies. This is depicted in Figure 6.4 by a steady rise of the probability of being in the first regime. Even thought this probability is increasing during the eighties is small compared to the probability of being in the second regime. Therefore, it is more likely that it was during the seventies that France was taking advantage of the comparatively higher levels of technology in the G7 countries.

Both regimes in this model are dynamically stable. The characteristic polynomial of the first regime has obviously only one real root with modulus less than one. Both characteristic roots of the second regime are real an less than one in absolute value.

Summing up, using a switching regressions model, there is evidence of spillovers of technology from the G7 countries to France. Unlike the case of the UK in which technology spillovers were more likely during the eighties and perhaps the early nineties, in France these spillover effects were more likely to have occurred during the early and mid seventies.

6.3.4 West Germany.

Table 6.8 presents the parameter estimates of the unrestricted switching regressions model for West Germany. The model is well specified; it passes all the misspecification tests. The coefficient of the switching variable, -12.280, is just significant at 5% (its standard error is equal to 6.2651). The hypothesis of equal dynamics but different variance in both regimes is also tested. This restricted model is shown in Table 6.10. The hypothesis cannot be rejected with a LR equal to 5.2055 and significance 0.3913. However, the estimated coefficient of the switching regressions variable is very different from the estimate obtained in Table 6.8. Another worrying feature of this model with equal dynamics is the fact that the estimated standard deviation for the first regime is almost 10 times the standard deviation of the second regime which seems to indicate convergence to a different maximum tending towards an unbounded solution. In fact, trying to estimate an unrestricted switching regressions model in the neighborhood of this point results in the usual unbounded solution. With this in mind the unrestricted switching regressions model is kept for further analysis.

The model chosen by AIC is shown in Table 6.9. The LR of the restrictions from the switching regressions model in Table 6.8 is equal to 0.5376 [0.4634]. This model has three lags of output growth in the first regime and four in the second. However, the model rejects the null hypotheses of the joint test for misspecification, therefore, the final model selected for West Germany is the unrestricted model of Table 6.10. The estimated growth rates in the two regimes are consistent with the theory, that is, the long run growth in the first regime (0.557% per quarter) is higher than that of the second regime (0.367% per quarter). However, the hypothesis of equality of growth rates in both regimes cannot be rejected at standard significance levels with a test statistic equal to 0.4423 [0.3291]. This seems to be a consequence of the uncertainty surrounding the calculated growth of output in the second regime since its standard error is higher than the calculated value.

The probability of being in the high growth regime, that is, the first regime, will be higher than 0.5 when the probability of a value of the logarithm of technology higher than the actual value is bigger than 0.3319. Figure 6.5 shows the probabilities of being in regime 1 for each observation in the sample. At the beginning of the seventies, the probability of being in the spillovers regime for West Germany was very high and consequently, in this period West Germany was taking advantage of beneficial technology spillovers from the rest of the G7 countries with high probability. Furthermore, almost throughout the whole sample period, the probability of being in regime 1 was higher than the probability of being in regime 2. Therefore, it is very likely that West Germany received spillovers of technology during the whole sample period. This was only relatively less likely to have happened during the late eighties and early nineties.

The time paths of both regimes of the switching regressions model selected for West Germany are stable. The first regime has four roots; a pair of complex roots and two real roots with modulus less than one. The second regime has also four roots; a pair of complex roots and two real roots less than one in absolute value.

In conclusion, it appears that spillovers of technology from the rest of the G7 countries play an important role in the case of West Germany. The spillover variable is significant, uncovering two distinct regimes. However, the calculated growth rate of output in the spillovers regime is not statistically significantly different from that calculated for the regime with no spillovers. This is a consequence of the uncertainty around the calculated growth rate for the regime with no spillovers. Even though it might seem that the spillovers of technology did not affect the growth rate of output, care should be taken before reaching this conclusion since for most of the period West Germany was highly likely to have been in the spillovers regime. Consequently, the uncertainty surrounding the output growth rate in the second regime might just be a result of this.

6.3.5 Italy.

Table 6.11 shows the estimated unrestricted switching regressions model for Italy. The estimated coefficient of the switching variable, -11.1275, is significant (the corresponding standard error is equal to 4.0875). However, the growth rate in the second regime is higher than the growth rate of the first regime, which seems to be inconsistent with the theory. Nevertheless the restriction that the growth rates are the same in both regimes cannot be rejected with a value of -0.3656. This model fails the joint test for misspecification which seems to be due to the presence of AR effects of order one in the first regime. However, the fourth lag of output growth is significant in both regimes. Thus, this unrestricted model is expanded further by adding four extra lags of output growth in each of the regimes, making a total of eight lags of output growth in both regimes. However, this expanded model is still misspecified due to the presence of AR effects of order 1 in the first regime. A finding that is more puzzling is the fact that all the extra lags of the output rate of growth are significant. To avoid the problem of overfitting, the model with four lags of output growth of Table 6.11 is kept for further analysis. The LR test of the hypothesis of equal dynamics but different variances in each regime is rejected with a test statistic of 11.7527 [0.0383].

In the case of Italy, the probability of being in the first regime is higher than 0.5 if the probability of a value of technology higher than the actual value is bigger than 0.8039. Figure 6.6 shows the probabilities of being in the spillover regime for each observation in the sample. The differences of Italy with the rest of the countries already studied is immediately obvious. At the beginning of the sample, the probability of being in the first regime is very low, but during 1974 and 1975, this probability experiences an important increase. There is then a period in which it is very likely that Italy is benefiting from spillovers and, therefore, this probability oscillates. First, the level of technology of Italy falls lower in the distribution, pressure mounts and the probability of being in the spillover regime increases. Once the level of technology rises in the distribution, the probability of being in the first regime decreases. However, this is not the case in the eighties and early nineties. Since the beginning of the eighties, the probability of being in the spillover regime increases steadily with time, so much that in the last quarter of 1994 this probability is over 0.95. However, this mounting pressure does not translate in a statistically significant higher growth rate in the spillovers regime so that Italy seems to fall further behind in the distribution of technology across countries. This could be the reason why the switching regressions model is misspecified with AR effects of order one in the first regime.

In conjunction with the rest of the countries, the time paths of the two regimes of the switching regressions model for Italy are stable. Both regimes have four characteristic roots. The first regime has a pair of complex roots and two real roots with modulus less than one. The second regime has two pairs of complex roots.

To conclude, the analysis of the role of technology spillovers in the process of output growth in Italy shows evidence of two regimes with different dynamics. Which regime is in place at each point in time is determined by the measure of the potential spillover defined in section 6.2. However, the calculated growth rates of output emerging from these two regimes are not statistically significantly different. It appears that during the eighties and nineties pressure had been building up which might have resulted in a higher growth rate of output after the end of the sample considered in this chapter. However, this hypothesis cannot be put to the test unless a larger set of data is employed.

6.3.6 Canada.

Table 6.12 presents the parameter estimates of the unrestricted switching regressions model for Canada. The null hypothesis of no misspecification in the model cannot be rejected with a value of (significance in square brackets) 7.5112 [0.1853]. However, there is evidence of ARCH effects of order one in the second regime; test statistic and significance equal to 4.4052 [0.0358]. The coefficient of the switching variable (z_t) is significant with a value of -8.1166 and standard error equal to 3.3374.The LR test of the hypothesis of equal dynamics but different variance in the two regimes is rejected with a test statistic value equal to 32.3007 [0.0000]. However, Canada differs from the other countries in that the long run growth rates for the two different regimes do not follow the theory. The growth rate in the first regime which should be the high growth regime, is calculated to be negative, whereas the growth rate in regime 2 is positive. Since this contradicts the theory, it should be further investigated. As a first thought, it might be the case that Canada gets most of the spillovers from the USA due to its physical proximity, whereas spillovers from the rest of the G7 countries are not that common. Therefore, it could be that the technology in Canada follows more the technological innovations in the USA. This will be investigated further later on but, first, the specification chosen by AIC will be presented and analysed.

Table 6.13 presents the estimates of the model for Canada chosen by AIC. The restrictions needed from the unrestricted model cannot be rejected with a LR statistic and significance equal to 7.5638 and 0.3726 respectively. This restricted model has one lag of output growth in both regimes and passes all the misspecification tests. Nevertheless, the long run growth rate in the first regime is negative (-2.73% per quarter) and in the second regime is positive (1.08% per quarter), and the hypothesis of the growth rate in the first regime being higher or equal than the growth rate in the second regime is rejected at 5% significance level; the statistic and significance being equal to -1.7346 [0.0414]. The probability of being in the first regime, will be higher than 0.5 when the probability of a value of the logarithm of technology higher than the actual value is bigger than 0.5155.

It was pointed out before that it might be the case that Canada follows more closely the technological innovations in the USA rather than the technological innovations in the rest of the G7 countries. To verify this hypothesis a new switching regressions model is estimated but the switching variable is defined this time as the difference between the logarithm of technology in the USA and the logarithm of technology in Canada, again demeaned over time. If the difference between the logarithms of technology between the USA and Canada is higher than the average gap, that is, the newly defined switching variable is positive, then Canada would benefit from a spillover from the USA. Assuming that $\beta_{3,1} < 0$, this will make the first regime the high growth regime. The results from this new switching regressions model are shown in Table 6.14. The model passes the general misspecification test but there is some evidence of ARCH effects of order one in the first regime. The estimated coefficient of the switching variable (-11.9993) is significant with a standard error equal to 4.1786. The hypothesis of equal dynamics but different variance in the two regimes is also tested here; the LR test gives a statistic equal to 30.0953 and its significance is 0.0000, therefore this hypothesis is openly rejected. Table 6.15 presents the parameter estimates of the model selected by AIC. The LR of the restrictions from the unrestricted model in Table 6.14 is equal to 3.8636 and its significance is equal to 0.5692. This restricted model has two lags of output growth in each regime and it is well specified; that is, none of the null hypotheses of the misspecification tests can be rejected. The long run growth rates in the first and second regime are estimated to be 1.25% and -1.23% per quarter respectively, which is consistent with the theory. Furthermore, the hypothesis that the two long run growth rates are the same is clearly rejected with a statistic equal to 3.9586 and 0.0000 significance.

In the case of Canada, the probability of being in the high growth regime, that is, the first regime, will be higher than 0.5 when the gap between the logarithms of technology of the USA and Canada is higher than the mean value over the sample which is 0.6717. A plot of the calculated probabilities of being in the first regime for each observation is shown in Figure 6.7. This plot shows that up to the late seventies, Canada was more likely to be in the regime in which spillovers from the USA were taking place. However, in the eighties and nineties, it was generally slightly more likely that the output growth process for Canada was generated by the second regime and, thus, spillovers were more unlikely.

These two regimes are stable; the first regime has a pair of complex roots whereas the second regime has two real roots less than one in absolute value.

Summing up, the measure of the potential spillover used for the rest of the countries in this chapter significantly defines two different regimes in the process of output growth in Canada. However, the output growth rates of these two regimes seem to contradict the theory that spillovers of technology increase the rate of growth of output while they are taking place. It was hypothesised that because of their proximity, Canada might receive most of the spillovers of technology from the USA. This hypothesis is substantiated by the switching regressions model with the potential spillover of technology measured with respect to the USA.

6.3.7 Japan.

The parameter estimates of the unrestricted switching regressions model for Japan are shown in Table 6.16. The model is well specified since the null hypotheses of not only the joint misspecification test, but all of the individual misspecification tests cannot be rejected. The coefficient of the switching variable is equal to -5.0117, but it is not significant with a standard error equal to 4.1196.

Table 6.17 depicts the parameter estimates of the model chosen by AIC. The LR test of the restrictions from the unrestricted model is equal to 6.0787 [0.5306], therefore, the restrictions cannot be rejected at normal levels of significance. The first regime for this model has no lags of output growth whereas the second regime has two. This restricted model, the same as the unrestricted model, is well specified, not failing any of the misspecification tests. The long run growth rates for the two regimes are consistent with the theory, that is the growth rate in the first regime is higher than that of the second. The calculated long run growth rate in the first regime is equal to 1.14% per quarter, whereas the calculated long run growth rate for the second regime is equal to -0.41% per quarter. However, the hypothesis that the two long run growth rates are the same cannot be rejected; the calculated statistic is equal to 0.8524 and its significance is 0.1970. In this case, the probability of being in the first regime is higher than 0.5 if the value of the switching variable is bigger than zero. Therefore, the probability of being in the first regime is higher than the probability of being in the second regime if the probability of a value of the logarithm of technology higher than the actual value is bigger than the mean probability across time, which in this case is equal to 0.8564. It will be interesting to compare the unrestricted model of Table 6.16 with a model of equal dynamics with four lags of output growth but allowing for different variance in both regimes. When trying to estimate this model problems were found to make the algorithm converge. The problem seems to be the presence of a constant term in the switching equation (when the model converged eventually it was found that the correlation between the two parameters in the switching equation was too high; -0.989). Since the constant term in the switching equation was not significant anyway, the model with equal dynamics was estimated but without a constant in the switching equation (see Table 6.18). The restrictions from the unrestricted model cannot be rejected, with a LR statistic equal to 6.1110 and associated probability equal to 0.4109. This model also had an AIC value higher than that of the unrestricted model; 303.2180 compared to 300.2735 for the unrestricted switching regressions model. This, in fact, seems to point out to the fact that the appropriate model for Japan is linear but with a especial kind of heteroskedasticity in the error terms of both regimes: the errors in the first regime have a higher variance than the errors in the second regime, but the means of both regimes are the same. AIC actually chooses a model with three lags of output growth with an AIC value of 303.4027 (the LR test from the unrestricted switching regressions model is 7.7416 [0.3559]). However, there is evidence of ARCH effects of order one in the first regime with this specification (test statistic and significance equal to 4.0788 and 0.0434 respectively), so the model chosen is the one in Table 6.18 with four lags of output growth. This model is well specified and the long run growth rate of output is equal to 0.82% per quarter. However, the switching variable is not significant; the coefficient and standard error are -5.1137 and 8.7292 respectively. Therefore, a switching regression model with equal dynamics but different volatility in each regime and with constant switching probability is estimated (see Table 6.19). This model has a higher AIC (304.80383) than the model presented in Table 6.18. The LR test of the restrictions needed from the unrestricted switching regressions model is equal to 2.9393 and its significance is 0.8164. The model is well specified and the long run growth rate of output is equal to 0.965% per quarter. The probability of being in regime 1 is constant an equal to 0.9618, therefore the majority of the observations are concentrated on the first regime.

This model is stable, the characteristic polynomial has four roots; a pair of complex roots and two real roots with modulus less than one. The dispersion in the first regime appears to be higher than that of the second regime.

The model for Japan seems, therefore, to be linear but with heteroskedastic errors. The first regime is by far the most likely. The second regime comes in effect, randomly it seems, since the switching equation only has a constant term and Markov effects do not seem present in this equation. However, failure to find a significant effect of the spillover measure used here for Japan does not inevitably imply that spillovers of technology are not important in Japan. Bearing in mind that the level of technology in Japan was very low compared to the rest of the G7 countries at the beginning of the sample, it is easy to conceive that Japan might have received spillovers of technology from other countries outside the G7 economies. If this is the case, the measure selected here based on the G7 countries will not perform well. Based on this measure, it would be tempting to conclude that spillovers of technology played no role in the growth rate of Japan. In fact, without further investigation of this issue, it can only be asserted that spillovers of technology from the G7 countries do not seem to be important in the growth process of Japan in the sample period considered here.

In conclusion, there seems to be evidence of technology spillovers across countries. The variable measuring the potential spillovers is significant in five out of seven countries. It is only insignificant for the USA and Japan. For reasons already explained, the measure of the potential spillover of technology used in this chapter may have not been an adequate one for these two countries. Therefore, further investigation of this issue is required before any further conclusions are drawn for these two countries. It was also found that for Canada the technological advances in the USA played a more significant role than those in the rest of the G7 countries. Next section is devoted to the analysis of the impulse response functions of the models estimated in the present section.

6.4 Estimated Impulse Response functions for the switching regressions model.

The switching regressions models fitted in the last section suggest asymmetries in the behaviour of output growth in the sense that the dynamics are different in both regimes. The focus here is on the distribution of the impulse responses of the logarithm of output as well as how the responses might vary across the regimes of the switching regression models estimated in the last section.

The impulse responses are obtained using the estimated parameters obtained in section 6.3 and randomly drawing from the distribution of the residuals. Since technology is assumed to be exogenous, it needs to be forecasted to be able to compute the impulse responses. There are several assumptions that are needed for this. Firstly, when calculating the impulse response function for a country, the level of technology is kept fixed for the rest of the countries. Although this is an assumption unlikely to hold since advances in the country for which the impulse response function is being calculated will obviously push up the distribution of technology across countries, this is a way of separating the effects of the shocks. To be able to take into account this shift in the technology distribution across countries, all countries should have been estimated simultaneously. The evolution of technology for each country is calculated according to equation (6.1) so that technology in each country grows at a constant rate over time but, depending on the size of the potential spillover, technology can grow temporarily quicker while the spillover is taking place. The estimated regressions for each country are shown in Table 6.23. All the histories (ω_{t-1}) in the sample used to estimate the models are also used here to generate the impulse responses. For each history, 100 random errors are drawn from the joint distribution of the innovations conditional on the regime. The actual regime is then determined by each history and the random innovation drawn for the regression for the evolution of technology. The maximum horizon for the impulse responses , N, is set to 20, that is a total of 5 years and the average is taken over 200 futures, R.

It was shown in chapter 5 how the responses for a nonlinear model of the type used here are history dependent. If there is a lot of variation in the responses, the only way of getting an accurate picture of the behaviour of the series is by looking at the distribution of the responses, rather than looking at just one particular history. The cases of the USA and Japan are, however, different to the rest of the countries in the sense that these two models are actually linear even though the errors in the model have a special kind of heteroskedastic pattern, as it was pointed out in the last section. For these two countries, therefore, the impulse responses will not be history dependent and in these cases, it will be enough to use just one history to generate the impulse response function.

The results of the generalized impulse responses for the logarithm of output are shown in Figures 6.8 to 6.14. Figures 6.8 and 6.9 show the unconditional impulse responses for the USA and Japan respectively. Figures 6.10 to 6.14 show the box plots for the rest of the G7 countries at each horizon, that is, the minimum, first quartile, median, third quartile and maximum response. The shocks are fixed at +1 and the procedure is conducted conditional on the regime and also unconditionally. It is also worth mentioning that the responses of shocks conditional on a particular regime are symmetric and proportional to the size of a shock. The reason is that conditioning on a regime makes the model linear.

A detailed description of the individual impulse response functions for each of the G7 countries is given below. However, the main conclusion arising from these figures is that the responses of the logarithm of output to shocks in the different countries are markedly different, both in the general shape of the median response and in the ultimate level of the response. In addition, the responses of the logarithm of output in the two regimes are also very different for those countries with two distinct regimes, that is, all the G7 countries with the exception of the USA and Japan. The median response across countries can either be curtailed (although not completely reversed) after 20 periods as in the case of Canada or it can be magnified. The extreme case is West Germany for which the median response after 20 periods is over three times the size of the initial shock. In general, the median response after 20 periods is not very different between the two different regimes for each country. Obviously, for the two countries with equal dynamics, the USA and Japan, there is only one measure of persistence since the specification of the regimes are the same apart from allowing for different variances in each of them. If attention is restricted to the remaining five countries in the sample, the path followed by the logarithm of output is very different for different countries and even for different histories for the same country. This fact translates in distributions that are very spread around the median response for all the countries with the only exception of the UK.

The remainder of this section is dedicated to the analysis of these impulse responses for the logarithm of output in more detail.

Figure 6.8 shows the impulse response for the USA for a shock equal to +1. After 20 periods the response to the shock is equal to 1.6134, that is higher than the initial size of the shock.

Figure 6.9 shows the impulse response for Japan, again for a fixed shock of +1. The shocking feature about this figure is that the response has not levelled off even after 20 periods, and the response after these 20 periods is equal to 2.6273 which is more than twice the size of the initial shock.

Figure 6.10 shows the boxplots of the generalized impulse response functions for the United Kingdom for shocks fixed at one standard deviation. The dispersion in both regimes is very small, although it is higher in Regime 1 (the responses lie in the interval $[0.9214, 1.0884]$ after 20 periods) than in regime 2 (the responses lie in the interval $[0.9977, 1.0699]$ after 20 periods). Furthermore, the distribution in the first regime seems to be quite symmetric and the range between the first and the third quartile is $[1.0000, 1.002]$ after 20 periods, therefore, the response after 20 periods is basically equal to the size of the initial shock. The distribution of the second regime seems to be skewed to the left and the range between the first and third quartile is a bit larger $[1.0023, 1.0138]$. The median response is essentially the same, 1.0000 in regime 1 and 1.0070 in regime 2 after 20 periods. In this case it can be concluded that the persistence to shocks in both regimes is quite similar and the dispersion of the responses for each history is quite small.

The boxplots for France are shown in Figure 6.11. In this case, the dispersion in both regimes is higher than the dispersion for the United Kingdom after the responses level off. After 20 periods, the responses lie in the interval [1.4144, 1.7142] in regime 1 and in the interval [1.3243, 1.7930] in regime 2. Therefore, the dispersion in the second regime seems to be higher than in the first regime. The distribution of the responses in both regimes are skewed to the left, although the distribution in regime 1 seems to be more asymmetric than the distribution in regime 2. Thus, the average response will tend to overestimate the likely response in both regimes. The median of the response in regime 2 is slightly higher, 1.4710, than in regime 1, 1.4907.

Figure 6.12 presents the boxplots for West Germany. The impulse response functions of West Germany are very different to the United Kingdom or France. The difference is that even after 20 periods the responses have not levelled off yet and this is not a feature of just a couple of extreme histories though, but of the whole distribution. The dispersion in both regimes is very high and after 20 periods the responses lie in the interval [1.3344, 3.5139] for regime 1 and [0.4540, 4.8556] for regime 2. Therefore, the dispersion in the second regime is even higher than in the first regime. Both distributions seem to be skewed to the right, specially the distribution of the impulse responses in the first regime. Therefore, the average response will tend to underestimate the likely response, specially in regime 1. It is interesting to note that while in the first regime a shock increases the logarithm of output by a factor higher than the initial size of the shock, this is not the case in the second regime. In this regime, after 20 periods the logarithm of output can increase by a factor that is either higher or lower than the size of the initial shock, although it is obviously never completely reversed. The median response in the first regime after 20 periods is equal to 3.2580 and in the second regime is equal to 3.4209, that is higher in regime 2.

Figure 6.13 shows the boxplots for Italy. The dispersion is again high, although not as high as for West Germany. For both regimes, the impulse responses seem to increase rapidly during the first three time periods and then they decrease also rapidly during the next four time periods, only to increase again. The responses seem to level off after 16 time periods. This is quite a large number of periods compared to for example, the United Kingdom or France. After 20 periods the responses in the first regime lie in the interval [0.6726, 1.7734] and in the interval [1.1751, 1.5576] in the second regime. During the first three time periods the distribution seems to be skewed to the right, that is, while the responses are increasing. When the responses start to decrease, the distribution tends to be skewed to the left. Therefore, this seems to indicate that the likely response is much more extreme than what the average response would suggest. It is also important to note that in the first regime a shock can increase the logarithm of output by a factor either higher of lower than the size of the initial shock. The median response in regime 1 after 20 time periods is equal to 1.3191 and in the second regime is equal to 1.3565 and therefore, similar.

The boxplots for Canada are shown in Figure 6.14. In this case the responses level off after approximately 10 periods. The dispersion in both regimes is again high and it seems to be higher in the first regime than in the second. After 20 periods the responses lie in the interval [0.6789, 4.3419] in the first regime and in the interval [0.7408, 3.3906] in the second regime. In both regimes, the shock can increase the logarithm of output by a factor either higher or lower than the initial size of the shock. The distributions in both regimes are skewed to the left. Thus, the average response will overestimate the likely response in both regimes. Again, the median response is practically the same in both regimes after 20 time periods; 0.8645 in the first regime and 0.8713 in the second regime.

From this analysis it can be concluded that for the United Kingdom the persistence of shocks is quite similar between both regimes. Obviously, for the USA and Japan the persistence is the same since there is only one measure due to the linearity of the models for these two countries. For France, West Germany, Italy and Canada the median of the impulse response functions is quite similar but the distribution is very spread so in this case the particular initial conditions turn out to be very important in assessing persistence.

So far both regimes within each country have been compared . If attention is restricted to the median of the distribution of the unconditional responses (that is including all the impulse response functions irrespective of the regime they are in at $t = 0$), after twenty periods for all the countries except from Canada, the logarithm of output is increased by more than the initial size of the shock. After 20 periods the smallest factor for the median response is 0.8685 (Canada) and the highest is 3.2777 (West Germany). However, the overall minimum factor can be as low as 0.4540 (West Germany) and the maximum as high as 4.8556 (West Germany) depending on the initial conditions. If West Germany is excluded, the overall minimum factor can be as low as 0.6726 (Italy) and the maximum as high as 4.3419 (Canada).

6.5 Conclusions.

In chapter 2 a theoretical model for output growth was illustrated based on a closed economy where technology was allowed to flow from the more technologically advanced economies to the less advanced. In this chapter, the importance of technology spillovers across the G7 countries has been studied, using a switching regressions model. There seems to be supporting evidence of the spillover effects of the type proposed in chapter 2. The proposed variable taken as a measure of the potential spillover is significant in five of the countries; the United Kingdom, France, West Germany, Italy and Canada. It is only insignificant for the USA and Japan. The technology measure adopted in this chapter which is in line with the literature gives a technology measure for the USA which is the highest of all the countries throughout the sample. In this case, it will be difficult to find any evidence of technology spillovers at an aggregate level, although they could be singled out at a disaggregate level. The problem in the case of Japan might be similar in the sense that the measure of the potential spillover selected here is perhaps not adequate. The reason being the low level of technology of Japan at the beginning of the sample with respect to the rest of the G7 countries. This low level of technology might make spillovers from other economies outside the G7 countries much more likely than those from the G7 countries. In this case, the measured used here will not perform well. For Canada, a different spillover variable was defined to assess the importance of spillovers of technology from the USA to Canada. It was found that there is evidence of technology spillovers from the USA to Canada. There was evidence of two distinct regimes with growth rates consistent with the theory for the United Kingdom, France and West Germany, although for this last country the calculated output growth rates in both regimes were not statistically different. For Italy the spillover variable was significant but the growth rates of output in both regimes were not significantly different. This model was found to have autoregressive effects of order one in the first regime which could be a consequence of an increasing pressure of a spillover of technology towards the end of the sample which does not yet materialized inside the sample considered in this chapter.

Overall, there is evidence of spillover effects across countries which consequently generates two distinct regimes in output growth with different short run dynamics and different growth rates in each regime. Therefore, it is important in any analysis of output growth to allow for this kind of feedback effects across countries.

It was also shown that the persistence of shocks varies with the initial conditions chosen even after conditioning on a particular regime. Although the median response of output in both regimes is similar after 20 periods for each country in turn, the distribution of the responses is quite wide for some countries and it tends to be skewed. This is the case of France, West Germany, Italy and Canada. In these cases, a single response or even the average response could be misleading when looking at the persistence of shocks. The wide differences of the responses of output to shocks across countries were also illustrated in this chapter. For the majority of the countries considered here the median response after 20 periods is higher than the initial size of the shock, although for some of the individual responses this is not the case. In addition, the paths followed by the logarithm of output in the different countries show distinct differences.

Table 6.1. Descriptive statistics for $1/(1+\phi_{i,t})$. Sample period 1972Q1 to 1994Q4

	Country			
	USA	**UK**	**France**	**West Germany**
Maximum	0.0424	0.3737	0.4963	0.3811
Minimum	0.0029	0.0828	0.1024	0.0600
Mean	0.1609	-0.7689	0.5135	0.9230
Std. Deviation	-0.4101	-0.6230	-0.6519	-0.4312
Skewness	0.0359	0.2010	0.3362	0.2994
Kurtosis-3	0.0498	0.5022	0.7004	0.5182
	Country			
	Italy	**Canada**	**Japan**	
Mean	0.8841	0.5237	0.8537	
Std. Deviation	0.0729	0.0572	0.0677	
Skewness	-1.3530	-0.6898	-0.4264	
Kurtosis-3	0.8319	-0.2577	-0.8245	
Minimum	0.6925	0.3759	0.6917	
Maximum	0.9557	0.6064	0.9475	

Table 6.2. Estimated unrestricted switching regressions model for the USA (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0032 (0.0024)	0.3336 (0.1910)	−0.1031 (0.1897)	−0.1660 (0.1968)	0.2224 (0.1862)	0.0114 (0.0014)	0.00444 (0.00262)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0042 (0.0011)	0.2893 (0.1030)	0.2002 (0.0879)	0.1883 (0.0876)	−0.2044 (0.0652)	0.0031 (0.0006)	0.00793 (0.00138)
Switching	β_{30}	β_{31}					
Equation	−0.0502 (0.2542)	−10.7667 (64.5195)					
Likelihood:				314.07393			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	3.2945	0.6547
AR(1) test for Regime 1	2.4097	0.1206
AR(1) test for Regime 2	0.0449	0.8321
ARCH(1) test for Regime 1	0.0026	0.9594
ARCH(1) test for Regime 2	0.3233	0.5696
Omitted Markov Effects	0.1483	0.7002
Test for equality of growth rates	-1.1093	0.1336
Equality of dynamics test	7.8394	0.1653

Table 6.3. Estimated switching regressions model with equal dynamics for the USA (SE's in brackets).

Regime	β_0	β_1	β_2	β_3	β_4	σ_1	σ_2	**Growth**
	0.0044 (0.0010)	0.2698 (0.0951)	0.1104 (0.0958)	– –	– –	0.0123 (0.0021)	0.0042 (0.0011)	0.00707 (0.00130)
Switching	β_{30}	β_{31}						
Equation	0.0578 (0.4379)	– –						
Likelihood:				309.26898				
LR test of restrictions:				9.6099[0.2935]				

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	3.7769	0.4370
AR(1) test	0.5282	0.4674
ARCH(1) test for Regime 1	3.1558	0.0757
ARCH(1) test for Regime 2	0.2845	0.5938
Omitted Markov Effects	0.0120	0.9128

Characteristic roots

Regimes 1 and 2
-0.2237
0.4935

Table 6.4. Estimated unrestricted switching regressions model for the UK (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0043 (0.0013)	0.3207 (0.1190)	0.0368 (0.0766)	−0.1862 (0.0822)	0.3504 (0.0852)	0.0051 (0.0007)	0.00903 (0.00226)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0055 (0.0035)	−0.2132 (0.1649)	−0.0066 (0.2075)	0.1338 (0.1959)	−0.3043 (0.1760)	0.0172 (0.0022)	0.00394 (0.00222)
Switching	β_{30}	β_{31}					
Equation	−0.2692 (0.2791)	−10.5306 (3.7205)					
Likelihood:				312.27814			

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	9.7883	0.0815
AR(1) test for Regime 1	0.6725	0.4122
AR(1) test for Regime 2	0.2379	0.6257
ARCH(1) test for Regime 1	3.4615	0.0628
ARCH(1) test for Regime 2	2.7546	0.0970
Omitted Markov Effects	6.7245	0.0095
Test for equality of growth rates	1.5428	0.0614
Equality of dynamics test	17.5803	0.0035

Table 6.5. Estimated switching regressions model for the UK (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0043 (0.0014)	0.3374 (0.1231)	0.0549 (0.0869)	−0.1843 (0.0887)	0.3223 (0.0909)	0.0049 (0.0007)	0.00912 (0.00239)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0031 (0.0029)	– –	– –	– –	– –	0.0178 (0.0021)	0.00310 (0.00285)
Switching	β_{30}	β_{31}					
Equation	– –	−12.5458 (4.7170)					

Likelihood:	289.52146
LR test of restrictions:	5.5134[0.3565]

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	6.0863	0.1928
AR(1) test for Regime 1	0.9528	0.3290
AR(1) test for Regime 2	1.1517	0.2832
ARCH(1) test for Regime 1	3.6264	0.0569
ARCH(1) test for Regime 2	2.2617	0.1326
Omitted Markov Effects	-	-
Test for equality of growth rates	1.5528	0.0602

Characteristic roots

Regime 1	Regime 2
-0.7659	-
0.7814	
0.1610+0.7160i	
0.1610-0.7160i	

Table 6.6. Estimated unrestricted switching regressions model for France (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0082	0.5134	−0.0458	−0.0272	−0.0361	0.0034	0.01376
	(0.0017)	(0.1132)	(0.1357)	(0.1332)	(0.1099)	(0.0006)	(0.00205)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0024	0.0221	0.3391	0.0487	−0.1696	0.0065	0.00320
	(0.0014)	(0.1356)	(0.1310)	(0.1367)	(0.1514)	(0.0006)	(0.00145)
Switching	β_{30}	β_{31}					
Equation	0.7331	−9.2991					
	(0.5011)	(3.5823)					
Likelihood:				338.38046			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	4.0381	0.5439
AR(1) test for Regime 1	0.2511	0.6163
AR(1) test for Regime 2	1.3984	0.2370
ARCH(1) test for Regime 1	0.1969	0.6573
ARCH(1) test for Regime 2	0.0109	0.9169
Omitted Markov Effects	2.7717	0.0959
Test for equality of growth rates	3.0372	0.0012
Equal dynamics test	13.3132	0.0206

Table 6.7. Estimated switching regressions model for France (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0072 (0.0011)	0.4291 (0.1092)	– –	– –	– –	0.0040 (0.0006)	0.01267 (0.00178)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0007 (0.0016)	−0.0482 (0.1622)	0.3600 (0.1662)	– –	– –	0.0065 (0.0008)	0.00095 (0.0021)
Switching	β_{30}	β_{31}					
Equation	– –	−6.6460 (2.4072)					
Likelihood:				336.46743			
LR test of restrictions:				3.8261[0.7002]			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	1.1827	0.8809
AR(1) test for Regime 1	0.9443	0.3312
AR(1) test for Regime 2	0.1270	0.7215
ARCH(1) test for Regime 1	0.0211	0.8844
ARCH(1) test for Regime 2	0.0449	0.8323
Omitted Markov Effects	-	-
Test for equality of growth rates	3.5014	0.0002

Characteristic roots	
Regime 1	Regime 2
0.4291	0.5764 -0.6246

Table 6.8. Estimated unrestricted switching regressions model for West Germany (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0041 (0.0018)	0.0629 (0.1505)	−0.0335 (0.1606)	0.3342 (0.1514)	−0.1083 (0.1493)	0.0094 (0.0009)	0.00557 (0.00174)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0010 (0.0014)	0.2267 (0.1112)	0.2333 (0.0898)	−0.2212 (0.0982)	0.4983 (0.1017)	0.0034 (0.0007)	0.00367 (0.00379)
Switching	β_{30}	β_{31}					
Equation	−0.5830 (0.3309)	−12.2801 (6.2651)					
Likelihood:				314.10228			

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	8.7350	0.1201
AR(1) test for Regime 1	0.3803	0.5374
AR(1) test for Regime 2	0.3124	0.5762
ARCH(1) test for Regime 1	2.7252	0.0988
ARCH(1) test for Regime 2	0.6560	0.4180
Omitted Markov Effects	1.4863	0.2228
Test for equality of growth rates	0.4423	0.3291
Equality of dynamics test	5.2055	0.3913

Characteristic roots

Regime 1	Regime 2
0.4513	0.8988
0.4289	-0.9221
-0.4087+0.6264i	0.1250+0.7653i
-0.4087-0.6264i	0.1250-0.7653i

Table 6.9. Estimated switching regressions model for West Germany (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0038 (0.0017)	0.0492 (0.1494)	−0.0462 (0.1582)	0.3195 (0.1512)	– –	0.0094 (0.0009)	0.00565 (0.00189)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0009 (0.0015)	0.2366 (0.1188)	0.2389 (0.0935)	−0.2209 (0.1029)	0.4902 (0.1073)	0.0033 (0.0007)	0.00353 (0.00408)
Switching	β_{30}	β_{31}					
Equation	−0.6160 (0.3392)	−12.0051 (6.4152)					
Likelihood:				313.83349			
LR test of restrictions:				0.5376[0.4634]			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	11.8822	0.0364
AR(1) test for Regime 1	2.1276	0.1447
AR(1) test for Regime 2	0.3512	0.5534
ARCH(1) test for Regime 1	2.6968	0.1005
ARCH(1) test for Regime 2	0.5429	0.4612
Omitted Markov Effects	1.4893	0.2223
Test for equality of growth rates	0.4568	0.3239

Table 6.10. Estimated switching regressions model with equal dynamics for West Germany (SE's in brackets).

Regime	β_0	β_1	β_2	β_3	β_4	σ_1	σ_2	**Growth**
	0.0042 (0.0005)	0.0794 (0.0414)	0.1498 (0.0421)	−0.0311 (0.0325)	0.1468 (0.0320)	0.0097 (0.0008)	0.0010 (0.0003)	0.00640 (0.00068)
Switching	β_{30}	β_{31}						
Equation	−1.1043 (0.2984)	−9.7032 (5.2130)						
Likelihood:				311.49954				
LR test of restrictions:				5.2055[0.3913]				

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	6.1079	0.1912
AR(1) test	3.1532	0.0758
ARCH(1) test for Regime 1	2.0014	0.1572
ARCH(1) test for Regime 2	0.0008	0.9781
Omitted Markov Effects	0.0573	0.8108

Table 6.11. Estimated unrestricted switching regressions model for Italy (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0023 (0.0010)	0.3632 (0.0665)	0.2663 (0.0676)	−0.2156 (0.0995)	0.2405 (0.0737)	0.0049 (0.0005)	0.00676 (0.00197)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0071 (0.0045)	0.8752 (0.2193)	−0.3444 (0.1915)	−0.0018 (0.1141)	−0.3597 (0.1746)	0.0104 (0.0019)	0.00860 (0.00410)
Switching	β_{30}	β_{31}					
Equation	−0.8611 (0.3315)	−11.1275 (4.0875)					
Likelihood:				328.72144			

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	12.4094	0.0296
AR(1) test for Regime 1	6.8666	0.0088
AR(1) test for Regime 2	1.4744	0.2247
ARCH(1) test for Regime 1	1.6647	0.1970
ARCH(1) test for Regime 2	1.1571	0.2821
Omitted Markov Effects	1.2029	0.2727
Test for equality of growth rates	-0.3656	0.3573
Equal dynamics test	11.7527	0.0383

Characteristic roots

Regime 1	Regime 2
-0.7942	0.7575+0.5921i
0.8126	0.7575-0.5921i
0.1724+0.5856i	-0.3199+0.5356i
0.1724-0.5856i	-0.3199-0.5356i

Table 6.12. Estimated unrestricted switching regressions model for Canada (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	−0.0015 (0.0012)	0.8925 (0.1308)	−0.0442 (0.1263)	0.1735 (0.0929)	−0.0752 (0.0934)	0.0052 (0.0007)	−0.02860 (0.09313)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0146 (0.0024)	−0.4463 (0.1312)	−0.0319 (0.1244)	−0.1448 (0.1536)	0.2651 (0.1552)	0.0071 (0.0009)	0.01075 (0.00106)
Switching	β_{30}	β_{31}					
Equation	−0.1465 (0.2481)	−8.1166 (3.3374)					
Likelihood:				316.38129			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	7.5112	0.1853
AR(1) test for Regime 1	3.0001	0.0833
AR(1) test for Regime 2	0.2741	0.6006
ARCH(1) test for Regime 1	0.1642	0.6853
ARCH(1) test for Regime 2	4.4052	0.0358
Omitted Markov Effects	0.0033	0.9544
Test for equality of growth rates	-1.6986	0.0447
Equal dynamics test	32.3007	0.0000

Table 6.13. Estimated switching regressions model for Canada (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	−0.0015 (0.0010)	0.9434 (0.0999)	–	–	–	0.0052 (0.0007)	−0.02725 (0.05717)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	0.0145 (0.0018)	−0.3458 (0.1184)	–	–	–	0.0076 (0.0010)	0.01081 (0.00105)
Switching	β_{30}	β_{31}					
Equation	–	−7.3737 (3.4396)					
Likelihood:				312.59941			
LR test of restrictions:				7.5638[0.3726]			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	5.7907	0.2153
AR(1) test for Regime 1	3.8340	0.0502
AR(1) test for Regime 2	0.1635	0.6860
ARCH(1) test for Regime 1	0.2889	0.5910
ARCH(1) test for Regime 2	0.3762	0.5396
Omitted Markov Effects	-	-
Test for equality of growth rates	-1.7346	0.0414

Table 6.14. Estimated switching regressions model for Canada (SE's in brackets)[86].

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0174 (0.0025)	−0.1043 (0.1449)	−0.2561 (0.1443)	−0.0521 (0.1403)	0.0465 (0.1399)	0.0084 (0.0010)	0.01276 (0.00125)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	−0.0026 (0.0010)	0.5420 (0.0810)	0.2186 (0.0839)	0.1441 (0.0782)	−0.0720 (0.0796)	0.0045 (0.0006)	−0.01533 (0.01562)
Switching	β_{30}	β_{31}					
Equation	0.1348 (0.2350)	−11.9993 (4.1786)					
Likelihood:				318.06436			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	8.2367	0.1437
AR(1) test for Regime 1	0.0473	0.8278
AR(1) test for Regime 2	2.1250	0.1449
ARCH(1) test for Regime 1	4.4457	0.0350
ARCH(1) test for Regime 2	0.3576	0.5499
Omitted Markov Effects	0.7423	0.3889
Test for equality of growth rates	3.4309	0.0003
Equal dynamics test	30.0953	0.0000

[86]The switching variable in this model is defined as the logarithm of the level of technology in the USA minus the logarithm of the level of technology in Canada demeaned over time.

Table 6.15. Estimated switching regressions model for Canada (SE's in brackets).[87]

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0169 (0.0022)	−0.0982 (0.1365)	−0.2512 (0.1351)	–	–	0.0084 (0.0010)	0.01254 (0.00110)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	−0.0025 (0.0009)	0.5522 (0.0833)	0.2415 (0.0851)	–	–	0.0046 (0.0006)	−0.01235 (0.00917)
Switching	β_{30}	β_{31}					
Equation	– –	−12.3645 (4.2889)					
Likelihood:			316.13245				
LR test of restrictions:			3.8636[0.5692]				

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	5.9291	0.2045
AR(1) test for Regime 1	0.0005	0.9828
AR(1) test for Regime 2	1.4375	0.2305
ARCH(1) test for Regime 1	2.5799	0.1082
ARCH(1) test for Regime 2	0.9019	0.3423
Omitted Markov Effects	-	-
Test for equality of growth rates	3.9586	0.0000

Characteristic roots	
Regime 1	Regime 2
-0.0491+0.4988i	-0.2876
-0.0491-0.4988i	0.8398

[87]See footnote from Table 6.14.

Table 6.16. Estimated unrestricted switching regressions model for Japan (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0103 (0.0035)	0.0274 (0.1187)	−0.1457 (0.1658)	−0.0383 (0.1650)	0.2166 (0.1314)	0.0056 (0.0014)	0.01094 (0.00145)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	−0.0008 (0.0031)	0.0440 (0.2451)	0.7392 (0.3682)	0.2993 (0.1674)	−0.1123 (0.2471)	0.0086 (0.0011)	−0.02821 (0.3350)
Switching	β_{30}	β_{31}					
Equation	0.0065 (0.6577)	−5.0117 (4.1196)					
Likelihood:			314.27348				

Misspecification tests

Test	Statistic	Significance
Joint test for Misspecification	5.3595	0.3736
AR(1) test for Regime 1	1.5146	0.2184
AR(1) test for Regime 2	0.4063	0.5238
ARCH(1) test for Regime 1	1.0119	0.3145
ARCH(1) test for Regime 2	0.9674	0.3253
Omitted Markov Effects	1.4932	0.2217
Test for equality of growth rates	0.6387	0.2615

Table 6.17. Estimated switching regressions model for Japan (SE's in brackets).

Regime 1	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	σ_1	**Growth**
	0.0114 (0.0019)	– –	– –	– –	– –	0.0064 (0.0012)	0.01141 (0.00187)
Regime 2	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	σ_2	**Growth**
	−0.0006 (0.0025)	0.2050 (0.2110)	0.6538 (0.2764)	– –	– –	0.0089 (0.0012)	−0.00412 (0.02401)
Switching	β_{30}	β_{31}					
Equation	– –	−6.6152 (5.0788)					
Likelihood:				311.23412			
LR test of restrictions:				6.0787[0.5306]			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	4.5850	0.3326
AR(1) test for Regime 1	0.0687	0.7933
AR(1) test for Regime 2	2.7681	0.0962
ARCH(1) test for Regime 1	0.9943	0.3187
ARCH(1) test for Regime 2	0.3599	0.5486
Omitted Markov Effects	-	-
Test for equality of growth rates	0.8524	0.1970

Table 6.18. Estimated switching regressions model with equal dynamics for Japan (SE's in brackets).

Regime	β_0	β_1	β_2	β_3	β_4	σ_1	σ_2	**Growth**
	0.0031 (0.0018)	0.0999 (0.1025)	0.2125 (0.1096)	0.1802 (0.1024)	0.1347 (0.1055)	0.0105 (0.0013)	0.0059 (0.0011)	0.00823 (0.00249)
Switching	β_{30}	β_{31}						
Equation	– –	−5.1137 (8.7292)						
Likelihood:					311.21797			
LR test of restrictions:					6.1110[0.4109]			

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	3.7750	0.2868
AR(1) test	0.7012	0.4024
ARCH(1) test for Regime 1	3.5865	0.0582
ARCH(1) test for Regime 2	0.0484	0.8258

Table 6.19. Estimated switching regressions model with equal dynamics for Japan (SE's in brackets).

Regime	β_0	β_1	β_2	β_3	β_4	σ_1	σ_2	**Growth**
	0.0038 (0.0016)	0.0734 (0.0960)	0.1373 (0.0987)	0.1957 (0.0947)	0.1989 (0.1012)	0.0073 (0.0007)	0.0233 (0.0142)	0.00965 (0.00204)
Switching	β_{30}	β_{31}						
Equation	−1.7719 (0.6391)	– –						
Likelihood:				312.80383				
LR test of restrictions:				2.9393[0.8164]				

Misspecification tests		
Test	Statistic	Significance
Joint test for Misspecification	3.8030	0.4333
AR(1) test	3.0324	0.0816
ARCH(1) test for Regime 1	0.3502	0.5540
ARCH(1) test for Regime 2	1.3696	0.2419
Omitted Markov Effects	2.3641	0.1242

Characteristic roots
Regimes 1 and 2
0.8432
-0.5845
-0.0926+0.6284i
-0.0926-0.6284i

Table 6.20: Estimated regressions for the evolution of technology (SE in brackets)[88]

Dependent variable $\Delta a_{i,t}$	**Regressor**	
Country	intercept	$spill_{t-1}$
USA	0.0013 (0.0011)	0.0136 (0.0293)
UK	0.0030 (0.0024)	0.1298 (0.0701)
France	0.0093 (0.0014)	0.1027 (0.0160)
West Germany	0.0080 (0.0014)	0.1336 (0.0233)
Italy	0.0032 (0.0020)	0.0982 (0.0482)
Canada	0.0035 (0.0016)	0.0857 (0.0440)
Japan	0.0152 (0.0018)	0.0776 (0.0102)

[88]The variable $spill_{t-1}$ is calculated as $\frac{\sqrt{3}S_t[a_{i,t}]}{\pi}\left(1+\phi_{i,t}\right)\ln\left(\frac{1+\phi_{i,t}}{\phi_{i,t}}\right)$, where $\phi_{i,t} = \exp\left[\frac{\pi}{\sqrt{3}}\frac{a_{i,t}-E_t[a_{i,t}]}{S_t[a_{i,t}]}\right]$ demeaned over time (see Chapter 4, Section 4.4 for a detailed explanation of this variable).

Figure 6.1. Plot of Total Factor Productivity for the G7 countries (1970q1-1994q4)

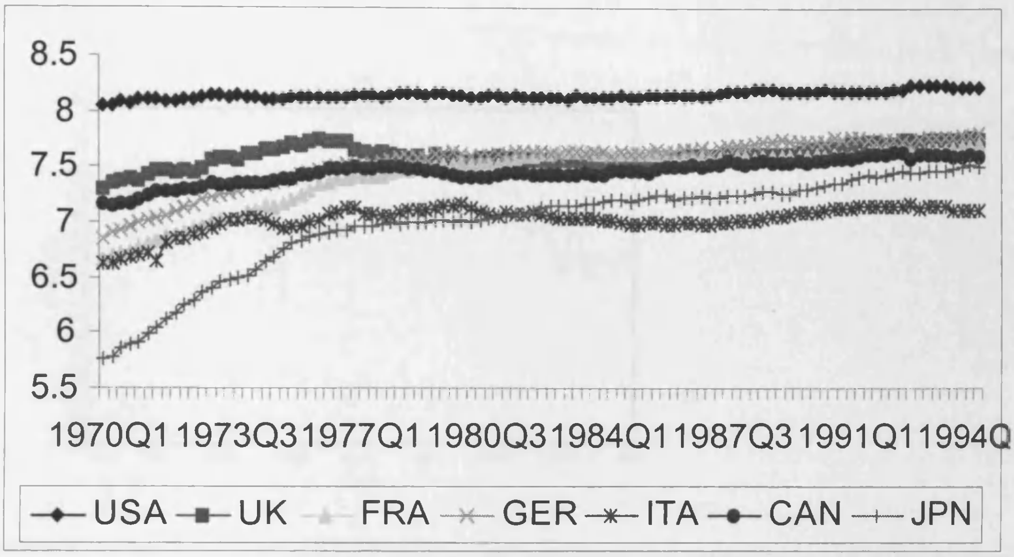


Figure 6.2. Plot of the spillover variable for each of the G7 countries across time.

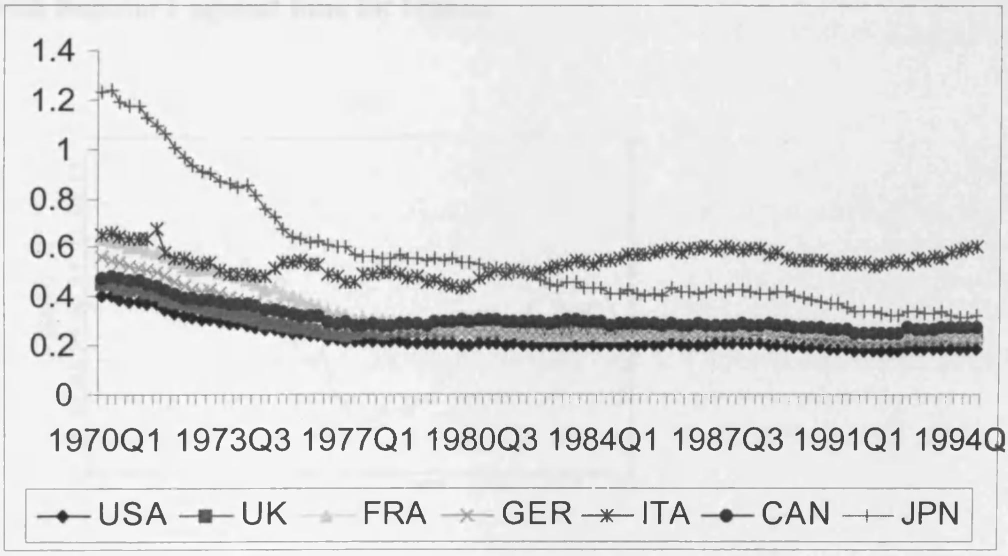


Figure 6.3. Plot of the calculated probabilities associated with Regime 1 against time for the United Kingdom.

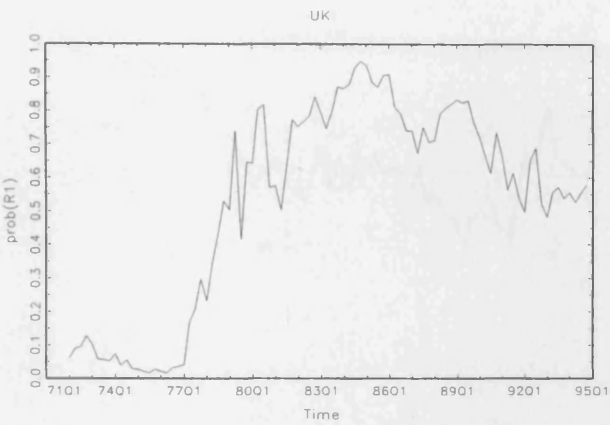


Figure 6.4. Plot of the calculated probabilities associated with Regime 1 against time for France.

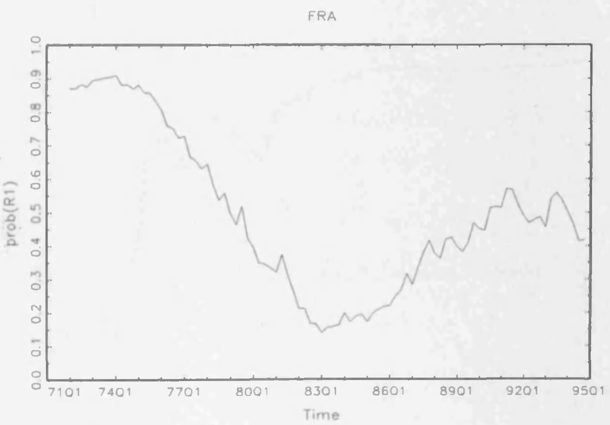


Figure 6.5. Plot of the calculated probabilities associated with Regime 1 against time for West Germany.

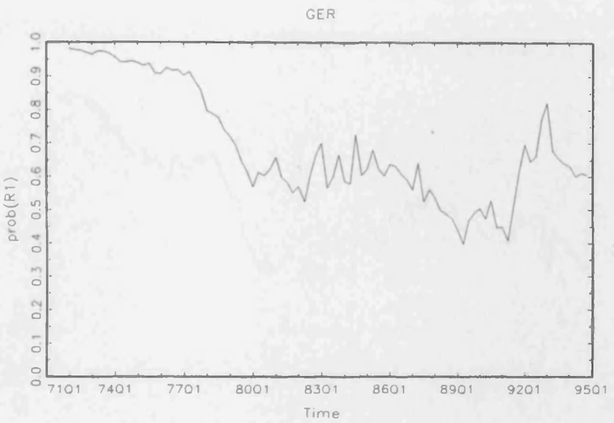


Figure 6.6. Plot of the calculated probabilities associated with Regime 1 against time for Italy.

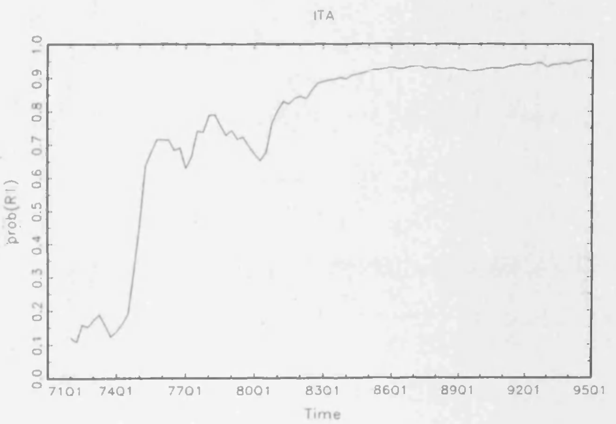


Figure 6.7. Plot of the calculated probabilities associated with Regime 1 against time for Canada.

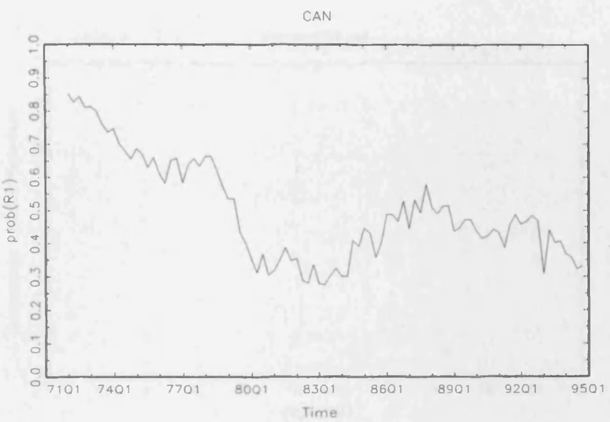


Figure 6.8. Plot of the impulse response function of the log level of GDP for positive shocks for the USA.

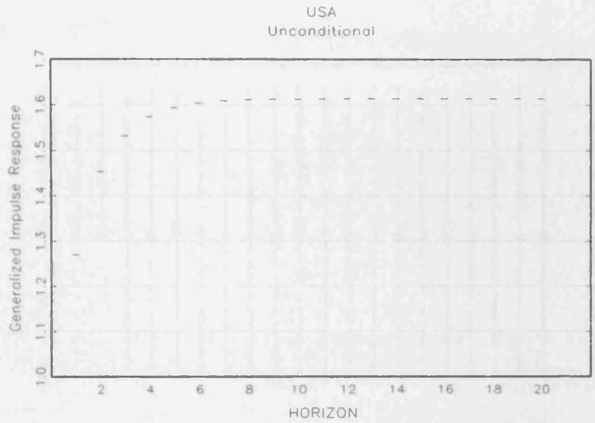


Figure 6.9. Plot of the impulse response function of the log level of GDP for positive shocks for the Japan.

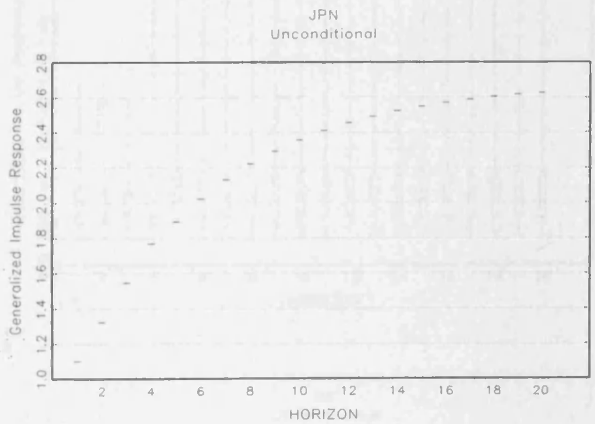
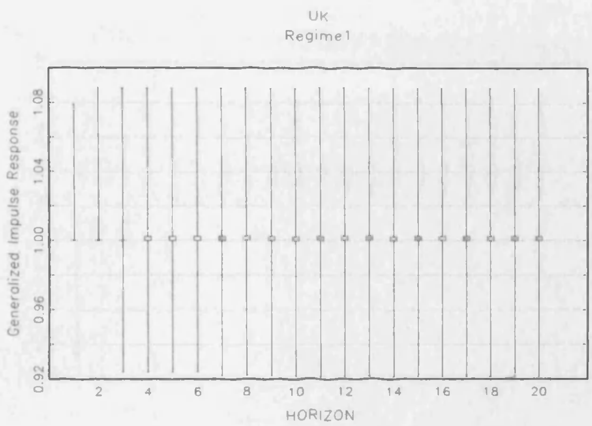
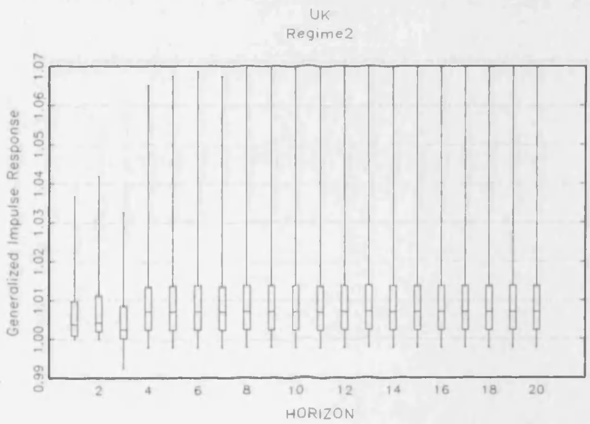
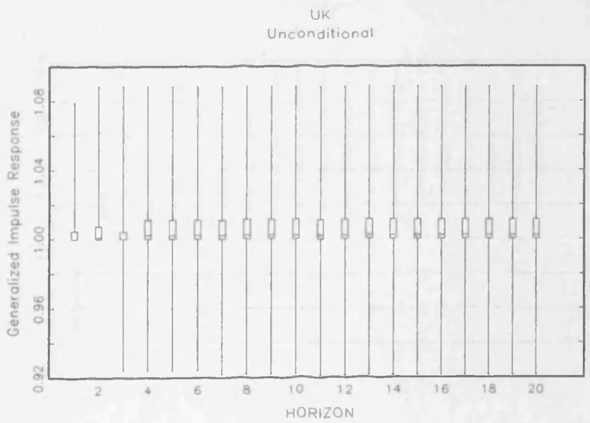


Figure 6.10. Box plot of the impulse response functions of the log level of GDP for positive shocks for the United Kingdom.[89]

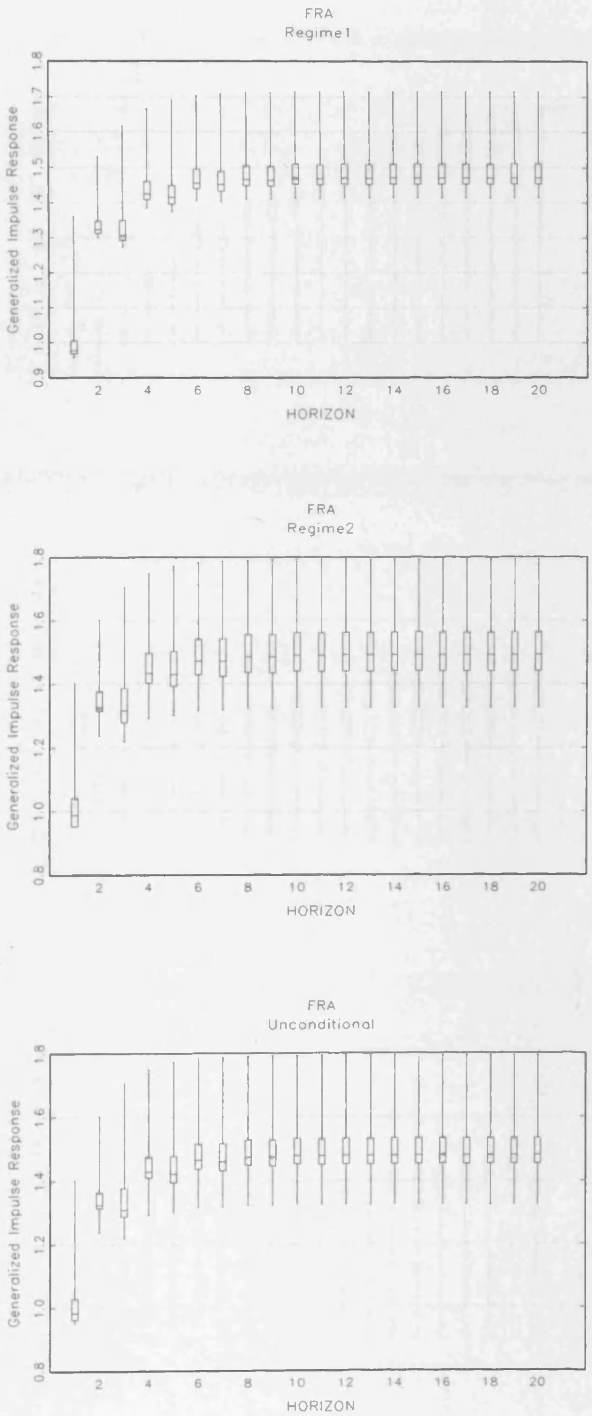






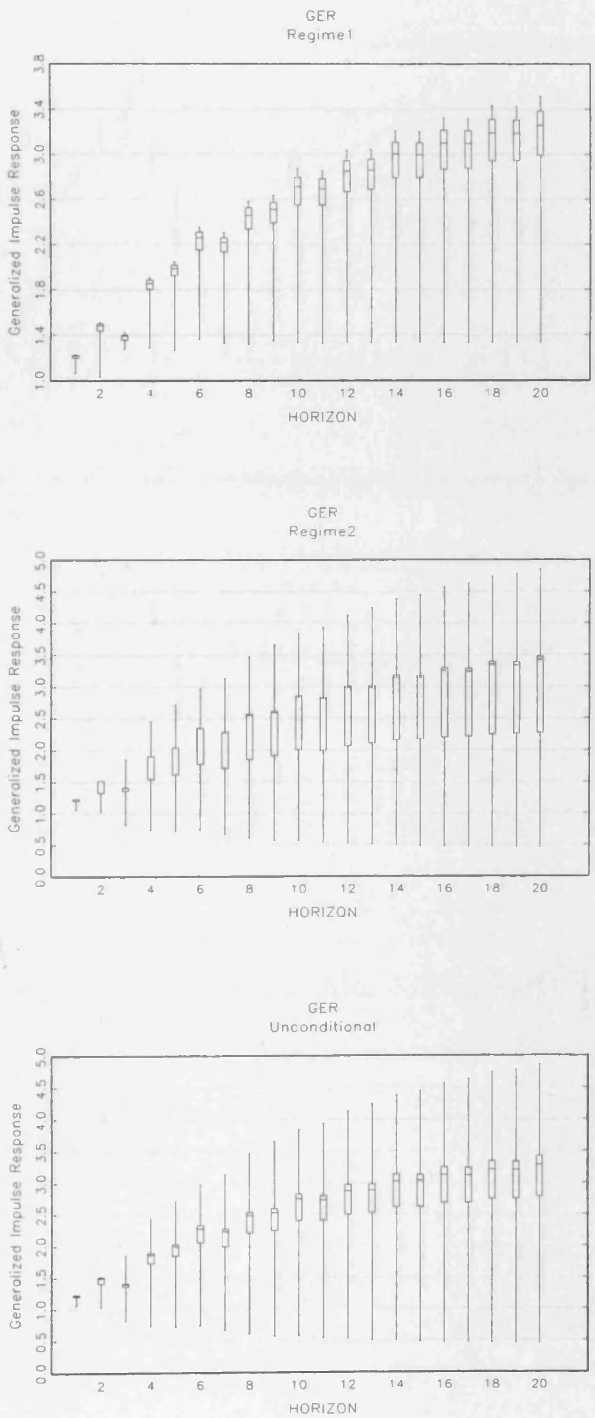
[89]These figures show the boxplot at each horizon for the generalized impulse response functions generated by a shock of size 1 and the associated histories.

Figure 6.11. Box plot of the impulse response functions of the log level of GDP for positive shocks for France.[90]



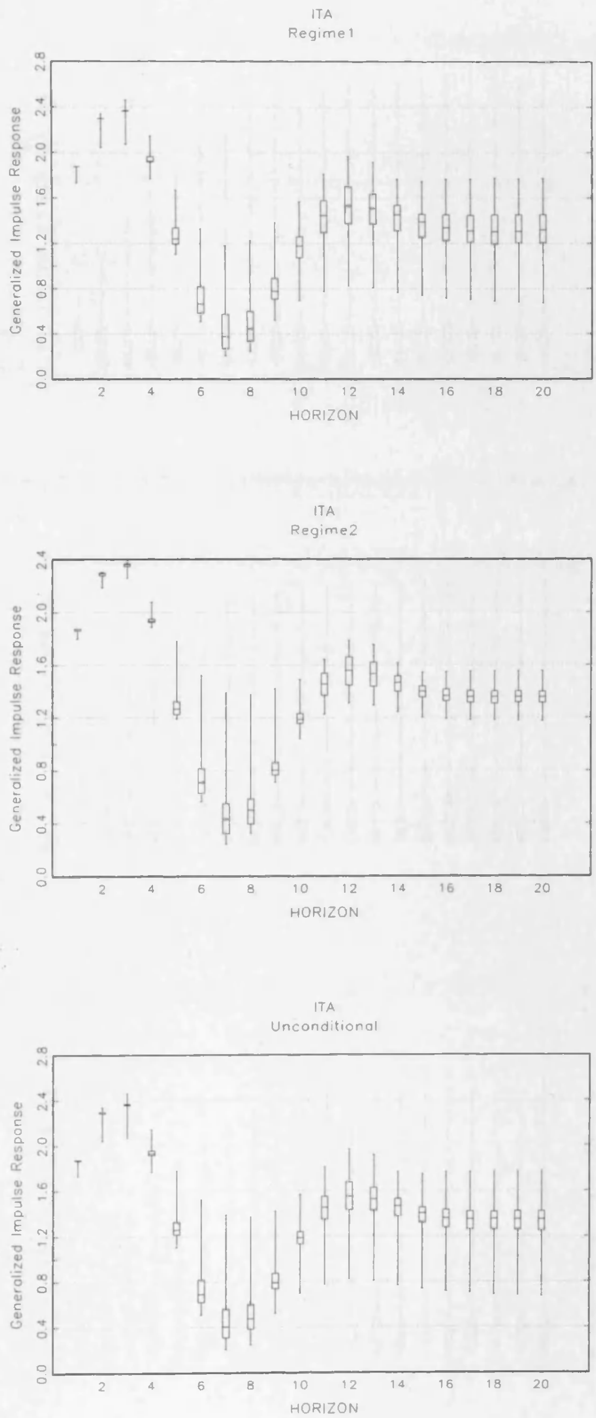
90See footnote to Figure 6.10.

Figure 6.12. Box plot of the impulse response functions of the log level of GDP for positive shocks for West Germany.[91]



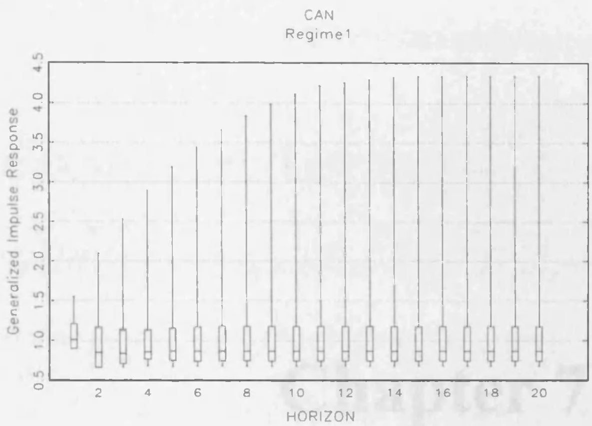
[91]See footnote to Figure 6.10.

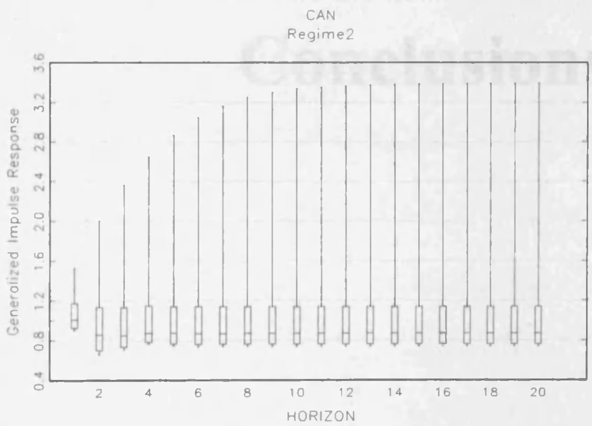
Figure 6.13. Box plot of the impulse response functions of the log level of GDP for positive shocks for Italy.[92]

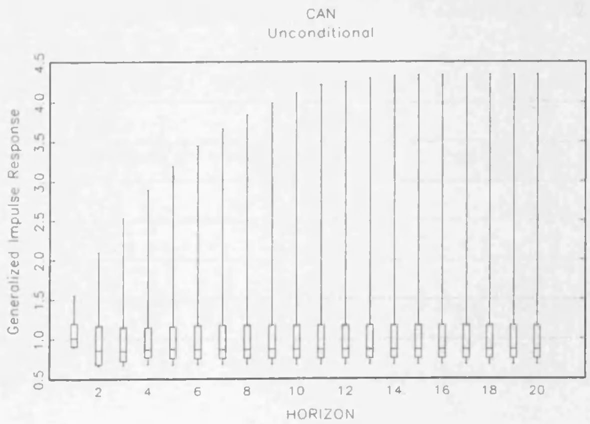


[92]See footnote to Figure 6.10.

Figure 6.14. Box plot of the impulse response functions of the log level of GDP for positive shocks for Canada.[93]







[93]See footnote to Figure 6.10.

Chapter 7

Conclusions.

This final chapter provides an overview of the findings in this thesis together with some suggestions for future research.

7.1 Summary and main contribution.

The objective of this thesis was to study the dynamics of output growth using time series econometric analysis applied to international output data.

It has compared empirically two neoclassical growth models and formally derived two exact discrete econometric models from their continuous time theoretical models. The advantage of this approach in terms of avoiding biases has been demonstrated.

The thesis has combined the techniques of switching regressions models, Simulated Annealing, testing procedures and adapted Generalised Impulse Responses into a coherent econometric framework and applied this in two empirical analyses.

The first part of this thesis concentrates on the issue of convergence of per capita output across a large group of countries. This subject has been studied at length in the empirical literature generally on the basis of a theoretical Solow type growth model. Chapter 2 provides an overview of such studies and illustrates the conflict between results. Alternative theoretical models are discussed and used as the basis for subsequent empirical chapters.

The issue of convergence of per capita output is dealt with in chapter 3. In this chapter the analysis of this issue deviates from the recent literature in two important respects. First, the analysis is carried out under the more sophisticated framework provided by the Ramsey-Cass-Koopmans model as well as the Solow growth model typically used in the applied literature. Second, exact discrete econometric models are formally derived from these two continuous time theoretical growth models. No evidence is found of convergence across the groups of countries considered here. Furthermore, the importance of both, deriving formally an econometric model from the theoretical model and the use of more sophisticated frameworks is highlighted by these results.

The above analysis assumes away any inter-relationships across the processes of growth in different countries. However, it is very difficult to conceive countries in complete isolation whose actions in terms of for example capital flows or technological developments have no consequences for the rest of the countries. As a result, this thesis evolves towards the study of the role of interactions across countries in the process of output growth. To this avail, the empirical work concentrates on a much smaller group of countries, the G7 countries. The rest of the thesis focuses on the incorporation of country inter-relations to the process of growth. In particular, two issues are examined here, the role of the current account of the balance of payments in output growth and the role of technology spillovers across countries. For this purpose, a more flexible econometric modelling framework is needed which is provided by the switching regressions model discussed in chapter 4. The advantages of a less rigid econometric approach come at a cost of additional theoretical and practical difficulties in estimation which are not always addressed in the literature. These problems were discussed at length and a global optimisation algorithm, the Simulated Annealing was proposed to overcome these difficulties when maximising the likelihood of the switching regressions model. Like in any other modelling framework, post-estimation model evaluation tests and computation of the impulse responses is a very important part of the modelling and as such they are also addressed here.

In chapter 5, these econometric tools are applied to the study of the balance of payments constraints in the process of growth based on the theoretical model in chapter 2. The growth of output is assumed to be generated by one of two different processes at each point in time depending on the extent of accumulations of the value of the current account over output in the near past. First, the usefulness of the Simulated Annealing algorithm in this type of situations is illustrated by comparing it with traditional optimisation algorithms when estimating this particular switching regime model. Even though as a general rule, a run of this algorithm takes longer to converge to the final parameter estimates, the problem of unboundness of the likelihood function is shown to be very common in switching regime models. In fact, it seems to be more widespread than what the literature would lead us to believe. Furthermore, multimodality of the likelihood function is a recurrent problem which arises from the highly nonlinear nature of the likelihood function. In this instances, it is shown how useful the Simulated Annealing algorithm is in overcoming these two problems. First, the output of the algorithm enables us to learn about the shape of the likelihood function and this is very useful in identifying its unbounded regions. The unbounded region can then be excluded from the search and as a result, attention can be restricted to the bounded region. Second, the algorithm is not dependent on the initial conditions which is a shortcoming of traditional algorithms. This useful property ensures that the global maximum will always be found. In contrast, the use of traditional algorithms require repeated maximization of the likelihood function using different initial conditions to ensure convergence towards the global maximum. This obviously it is ultimately more time consuming than maximizing the likelihood function with the simulated annealing algorithm.

Subsequently, attention is restricted to the analysis of the results obtained from the estimation of this switching regressions model. The hypothesis put forward was that prolonged current account deficits in the past might, on one hand, affect the dynamics of output growth and, on the other hand, they might inhibit the rate of growth in the long run. In the countries which are analysed here, the G7 countries, current account deficits do not generally to affect the long run growth rate. Nevertheless, there is evidence that it does affect the short run dynamics of output growth.

Additionally, making use of the adapted Generalized Impulse Response functions, it is illustrated how the persistence of shocks varies across countries, varies across regimes and varies with the initial conditions. For some countries, the distribution of the responses is very wide and also skewed. In cases like this, the average response will not be a good indicator of the persistence of output shocks. Therefore, this needs to be taken into account when studying persistence and the focus of attention should be on the distribution of the responses rather than simply examining one or two in isolation.

Chapter 6 concentrates on studying the role of technology spillovers across countries in the process of growth. Again, it draws on the empirical framework discussed in chapter 4 and the theoretical model developed in chapter 2. The switching regressions model is used to uncover the effects of technology spillovers across countries in output growth. Output growth is generated by one of two processes with potentially different parameters. The switches between regimes are governed by a third regression which has a measure of the size of the potential spillover from which each particular country can benefit at each point in time. Overall, there is evidence of spillover effects across the group of countries under study here. Output growth is therefore, generated generally by two different processes with different short run dynamics and different long run growth rates in each regime for the countries studied here. Thus, this chapter highlights the importance of analysing output growth allowing for this type of effects across countries.

7.2 Implications for further research.

The empirical techniques which have been demonstrated in this thesis are reliant upon substantial computing power. It is therefore, not surprising that their use is not widespread. The advantages of these techniques and solutions to their possible drawbacks have been demonstrated in the context of growth analysis. However, such techniques are applicable to a wide range of economic issues and with access to more powerful computers, analysis of this type may become more widespread.

The thesis also highlights the inherent bias of commonly employed discrete time derivations of continuous time models. Such derivations ought to be treated with caution and the feasibility of continuous time derivations examined.

The empirical chapters highlight the importance of a broader range of factors that have not received detailed attention in the convergence literature. In particular the role of spillovers of technology across countries and their influence in the process of growth. Future research could focus on this issue with particular emphasis on the relationships of technology across all countries rather than the relative crude method which focuses on technology leaders, particularly the USA.

APPENDIX I Appendix to chapter 2.

Section I.i of this appendix derives the spillover at time t assuming that the logarithm of technologies ($\mathbf{a}_t$) across countries have a logistic distribution. In section I.ii the expected value and the variance of the spillover are obtained. Section I.iii derives the evolution of the logarithm of technology over time.

I.i. Derivation of the spillover.

The spillover for country i at time t depicted in equation (2.47) can be written as follows

$$SPILL_{i,t} = E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right] = E\left[\mathbf{a}_t \mid \mathbf{a}_t > a_{i,t}\right] - a_{i,t} \tag{I.1}$$

Assuming that technology across countries has a logistic distribution at each time t, the distribution of $\mathbf{a}_t$ at time t is given by

$$f_t\left(\mathbf{a}_t\right) = \frac{\gamma_t \exp\left(\eta_t + \gamma_t \mathbf{a}_t\right)}{\left[1 + \exp\left(\eta_t + \gamma_t \mathbf{a}_t\right)\right]^2} \qquad \gamma_t > 0 \tag{I.2}$$

and its cumulative distribution at time t is

$$F_t\left(\mathbf{a}_t\right) = \frac{\exp\left(\eta_t + \gamma_t \mathbf{a}_t\right)}{1 + \exp\left(\eta_t + \gamma_t \mathbf{a}_t\right)} \tag{I.3}$$

In this case, the expected value and the variance of the logarithm of technology across countries at time t are given by the following expressions

$$E\left[\mathbf{a}_t\right] = \frac{-\eta_t}{\gamma_t} \tag{I.4}$$

$$S^2\left[\mathbf{a}_t\right] = \frac{\pi^2}{3\gamma_t^2} \tag{I.5}$$

The first step towards the calculation of the spillover in equation (I.1) is to find the value of $E[\mathbf{a}_t \mid \mathbf{a}_t > a_{i,t}]$, which by definition is given by

$$\begin{aligned} E[\mathbf{a}_t \mid \mathbf{a}_t > a_{i,t}] &= \frac{\int_{a_{i,t}}^{\infty} a_t^* f_t(a_t^*)\, da_t^*}{1 - F_t(a_{i,t})} \\ &= \frac{\int_{-\infty}^{\infty} a_t^* f_t(a_t^*)\, da_t^* - \int_{-\infty}^{a_{i,t}} a_t^* f_t(a_t^*)\, da_t^*}{1 - F_t(a_t)} \\ &= \frac{E[\mathbf{a}_t] - \int_{-\infty}^{a_{i,t}} a_t^* f_t(a_t^*)\, da_t^*}{1 - F_t(a_{i,t})} \end{aligned} \tag{I.6}$$

To calculate this expected value we need the following three results

1) Let $\phi_{i,t} = \exp(\eta_t + \gamma_t a_{i,t})$. Using equations (I.4) and (I.5), $\phi_{i,t}$ can be written in the following way

$$\phi_{i,t} = \exp(\eta_t + \gamma_t a_{i,t}) = \exp\left[\frac{\pi}{\sqrt{3}} \frac{a_{i,t} - E[\mathbf{a}_t]}{S[\mathbf{a}_t]}\right] \tag{I.7}$$

2) Integrating by parts we get the following result:

$$\int_{-\infty}^{a_{i,t}} a_t^* f_t(a_t^*)\, da_t^* = a_{i,t} F_t(a_{i,t}) - \frac{1}{\gamma_t} \ln[1 + \exp(\eta_t + \gamma_t a_{i,t})] \tag{I.8}$$

which using equation (I.5) gives

$$\int_{-\infty}^{a_{i,t}} a_t^* f_t(a_t^*)\, da_t^* = a_{i,t} F_t(a_{i,t}) - \frac{\sqrt{3} S[\mathbf{a}_t]}{\pi} \ln(1 + \phi_{i,t}) \tag{I.9}$$

3) Equation (I.3) can be rearranged to give

$$1 - F_t(a_{i,t}) = \frac{1}{1 + \exp(\eta_t + \gamma_t a_{i,t})}$$

and using equation (I.7) this can be written as

$$1 - F_t\left(a_{i,t}\right) = \frac{1}{1 + \phi_{i,t}} \tag{I.10}$$

Now, using equations (I.7), (I.9) and (I.10), the expected value in equation (I.6) can be written in terms of $a_{i,t}$, $E\left[\mathbf{a}_t\right]$ and $S\left[\mathbf{a}_t\right]$ as

$$E\left[\mathbf{a}_t \mid \mathbf{a}_t > a_{i,t}\right] = \left(1 + \phi_{i,t}\right) E\left[\mathbf{a}_t\right] - \phi_{i,t} a_{i,t} + \frac{\sqrt{3} S\left[\mathbf{a}_t\right]}{\pi} \left(1 + \phi_{i,t}\right) \ln\left(1 + \phi_{i,t}\right)$$

and therefore, after some rearranging, the spillover in equation (I.1) can be written as

$$\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right] = \lambda \frac{\sqrt{3} S\left[\mathbf{a}_t\right]}{\pi} \left(1 + \phi_{i,t}\right) \ln\left(\frac{1 + \phi_{i,t}}{\phi_{i,t}}\right) \tag{I.11}$$

I.ii. The expected value and the variance of the spillover.

In this section, the expected value and the variance of the spillover variable which was derived in the last section are obtained.

Expected value of the spillover.

Using equations (I.11), (I.2) and (I.7) the expected value of the spillover can be written as

$$E\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] = \int_{-\infty}^{+\infty} \lambda \ln\left(\frac{1 + \phi_{i,t}}{\phi_{i,t}}\right) \frac{\phi_{i,t}}{\left(1 + \phi_{i,t}\right)} da_{i,t}$$

The obvious step to simplify this integral is to change the integration variable from $a_{i,t}$ to $\phi_{i,t}$. Since from equation (I.7), $\phi_{i,t} = \exp\left[\frac{\pi}{\sqrt{3}}\frac{a_{i,t}-E[\mathbf{a}_t]}{S[\mathbf{a}_t]}\right]$, then

$$d\phi_{i,t} = \frac{\pi}{\sqrt{3}S\left[\mathbf{a}_t\right]}\exp\left[\frac{\pi}{\sqrt{3}}\frac{a_{i,t}-E\left[\mathbf{a}_t\right]}{S\left[\mathbf{a}_t\right]}\right]da_{i,t}.$$

Rearranging this equation the following equality is obtained

$$da_{i,t} = \frac{\sqrt{3}S\left[\mathbf{a}_t\right]}{\pi}\frac{1}{\phi_{i,t}}d\phi_{i,t}$$

and consequently the new limits of the integral are

when $a_{i,t} = +\infty \Longrightarrow \phi_{i,t} = +\infty$

when $a_{i,t} = -\infty \Longrightarrow \phi_{i,t} = 0.$

Thus, the expected value of the spillover across countries can be written as

$$\begin{aligned}
& E\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] = \\
= {} & \lambda\frac{\sqrt{3}S\left[\mathbf{a}_t\right]}{\pi}\int_0^{+\infty}\frac{1}{\left(1+\phi_{i,t}\right)}\ln\left(\frac{1+\phi_{i,t}}{\phi_{i,t}}\right)d\phi_{i,t} \\
= {} & \lambda\frac{\sqrt{3}S\left[\mathbf{a}_t\right]}{\pi}\frac{1}{6}\pi^2
\end{aligned}$$

and therefore, simplifying

$$E\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] = \frac{\lambda\pi}{2\sqrt{3}}S\left[\mathbf{a}_t\right] \tag{I.12}$$

Thus, the expected value of the spillover at time t depends on the speed of catch-up and on the standard deviation of the distribution of $\mathbf{a}_t$. That is, the higher λ, the higher the value of the expected spillover, other things being equal. Also, the wider the distribution at time t, that is, the higher the value of $S\left[\mathbf{a}_t\right]$, the higher the expected value of the spillover because the distances between the technologies of countries are bigger and therefore the size of the spillover is bigger.

Variance of the spillover.

By definition the variance of the spillover is equal to

$$\begin{aligned} & S^2\left[\lambda E\left[\mathbf{a}_t - a_t \mid \mathbf{a}_t > a_{i,t}\right]\right] \\ = \quad & E\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right]^2 - \left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right]^2 \end{aligned}$$

Using equations (I.11), (I.2) and (I.7) the variance of the spillover can be expressed as

$$\begin{aligned} & S^2\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] \\ = \quad & \left(\lambda \frac{\sqrt{3} S\left[\mathbf{a}_t\right]}{\pi}\right)^2 \int_0^{+\infty} \left(\ln\left(\frac{1+\phi_{i,t}}{\phi_{i,t}}\right)\right)^2 d\phi_{i,t} - \left(\frac{\lambda \pi}{2\sqrt{3}} S\left[\mathbf{a}_t\right]\right)^2 \\ = \quad & \left(\lambda \frac{\sqrt{3} S\left[\mathbf{a}_t\right]}{\pi}\right)^2 \frac{\pi^2}{3} - \left(\frac{\lambda \pi}{2\sqrt{3}} S\left[\mathbf{a}_t\right]\right)^2 \end{aligned}$$

Simplifying, the following result is obtained

$$S^2\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] = \lambda^2 \left(1 - \frac{\pi^2}{12}\right) S^2\left[\mathbf{a}_t\right]$$

As is the case for the expected value of the spillover at time t, the variance of the spillover also depends on the speed of catch-up and on the variance of the distribution of $\mathbf{a}_t$. That is, the higher λ, the higher the variance of the spillover, other things being equal. Also, the wider the distribution at time t, that is, the higher the value of $S\left[\mathbf{a}_t\right]$, the higher the expected value of the spillover.

I.iii. Dynamic properties of the spillover.

From equation (2.51) for each country i, the logarithm of technology at time $t+1$ evolves according to

$$a_{i,t+1} = g + a_{i,t} + \lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right] + \varepsilon_{i,t+1}$$

Therefore, taking expectations across countries on both sides and using the fact that $E\left[\varepsilon_{i,t+1}\right] = 0$, the expected value of the distribution of the logarithm of technologies at time $t+1$ is given by

$$\begin{aligned} E\left[\mathbf{a}_{t+1}\right] &= g + E\left[\mathbf{a}_t\right] + E\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] \\ &= g + E\left[\mathbf{a}_t\right] + \frac{\lambda\pi}{2\sqrt{3}} S\left[\mathbf{a}_t\right] \end{aligned}$$

On the other hand, the variance of the distribution of the logarithm of technologies at time $t+1$ is given by

$$\begin{aligned} S^2\left[\mathbf{a}_{t+1}\right] &= S^2\left[g + a_{i,t} + \lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t^* > a_{i,t}\right] + \varepsilon_{i,t+1}\right] \qquad \text{(I.13)} \\ &= S^2\left[\mathbf{a}_t\right] + S^2\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] + S^2\left[\varepsilon_{i,t+1}\right] \\ &\quad +2Cov\left[a_{i,t}, \lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] \end{aligned}$$

Hence, it is necessary to find the covariance between the spillover and the logarithm of technology at time t which by definition can be written as

$$\begin{aligned} & Cov\left[a_{i,t}, \lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] \\ = \;& E\left[\lambda a_{i,t} E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] - E\left[\mathbf{a}_t\right] E\left[\lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] \end{aligned}$$

Substituting the result in equation (I.12), the following is obtained

$$
\begin{aligned}
& Cov\left[a_{i,t}, \lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] \\
= & \int_0^{\infty} \lambda\left(\frac{\sqrt{3}S\left[\mathbf{a}_t\right]}{\pi}\ln\phi_{i,t} + E\left[\mathbf{a}_t\right]\right)\frac{\sqrt{3}S\left[\mathbf{a}_t\right]}{\pi}\ln\left(\frac{1+\phi_{i,t}}{\phi_{i,t}}\right)\frac{1}{\left(1+\phi_{i,t}\right)}d\phi_{i,t} \\
& -\frac{\lambda\pi}{2\sqrt{3}}S\left[\mathbf{a}_t\right]E\left[\mathbf{a}_t\right]
\end{aligned}
$$

Rearranging and separating the terms in the integral gives

$$
\begin{aligned}
& Cov\left[a_{i,t}, \lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] \\
= & \lambda\frac{\sqrt{3}S\left[\mathbf{a}_t\right]}{\pi}\left(\frac{\sqrt{3}S\left[\mathbf{a}_t\right]}{\pi}\int_0^{\infty}\ln\left(\frac{1+\phi_{i,t}}{\phi_{i,t}}\right)\frac{\ln\phi_{i,t}}{\left(1+\phi_{i,t}\right)}d\phi_{i,t}\right. \\
& \left. +E\left[\mathbf{a}_t\right]\int_0^{\infty}\ln\left(\frac{1+\phi_{i,t}}{\phi_{i,t}}\right)\frac{d\phi_{i,t}}{\left(1+\phi_{i,t}\right)}\right) - \frac{\lambda\pi}{2\sqrt{3}}S\left[\mathbf{a}_t\right]E\left[\mathbf{a}_t\right] \\
= & \lambda\frac{3S^2\left[\mathbf{a}_t\right]}{\pi^2}(-\Omega) + \lambda\frac{\sqrt{3}S\left[\mathbf{a}_t\right]}{\pi}E\left[\mathbf{a}_t\right]\frac{\pi^2}{6} - \frac{\lambda\pi}{2\sqrt{3}}S\left[\mathbf{a}_t\right]E\left[\mathbf{a}_t\right]
\end{aligned}
$$

where $\Omega = 1.202056903$. Thus

$$
Cov\left[a_{i,t}, \lambda E\left[\mathbf{a}_t - a_{i,t} \mid \mathbf{a}_t > a_{i,t}\right]\right] = \lambda\frac{3S^2\left[\mathbf{a}_t\right]}{\pi^2}(-\Omega)
$$

Note that the integral $\int_0^{\infty}\ln\left(\frac{1+\phi_{i,t}}{\phi_{i,t}}\right)\frac{\ln\phi_{i,t}}{\left(1+\phi_{i,t}\right)}d\phi_{i,t}$ has been calculated by numerical integration.

Now, using this result into equation (I.13), the variance of the logarithm of technology at time $t+1$ can be written as follows

$$
\begin{aligned}
S^2\left[\mathbf{a}_{t+1}\right] &= S^2\left[\mathbf{a}_t\right] + \lambda^2\left(1-\frac{\pi^2}{12}\right)S^2\left[\mathbf{a}_t\right] + \sigma^2 - 2\lambda\frac{3S^2\left[\mathbf{a}_t\right]}{\pi^2}\Omega \\
&= \sigma^2 + S^2\left[\mathbf{a}_t\right]\left[1+\lambda^2\left(1-\frac{\pi^2}{12}\right) - \frac{6\lambda}{\pi^2}\Omega\right]
\end{aligned}
$$

This is a first order difference equation on $S^2[\mathbf{a}_t]$. Since $1+\lambda^2\left(1-\frac{\pi^2}{12}\right)-\frac{6\lambda}{\pi^2}\Omega$ is less than one for values $0<\lambda\leq 1$, this variance tends over time to a constant given by

$$S^2[\mathbf{a}_{ss}] = \frac{\sigma^2}{\frac{6\lambda}{\pi^2}\Omega - \lambda^2\left(1-\frac{\pi^2}{12}\right)} \tag{I.14}$$

In particular, if $\lambda = 0$, that is, there are no spillovers, then the coefficient of $S^2[\mathbf{a}_t]$ in equation (I.14) is equal to 1 and the variance of the distribution of the logarithm of technologies increases with time. The amount it rises by in each period depending on the size of the shocks to the logarithm of technology. For values $0<\lambda\leq 1$ the coefficient of $S^2[\mathbf{a}_t]$ is positive and less than one. As lambda increases the coefficient decreases and in particular when $\lambda = 1$ this coefficient is equal to 0.44677.

APPENDIX II Appendix to chapter 3.

II.i. Linearization of the continuous time Solow Growth model.

This section outlines the steps needed in the linearization of the nonlinear differential equation obtained from the continuous time derivation of the Solow growth model, that is, the following equation

$$\frac{dk(t)}{dt} = sk(t)^{\alpha} - (n+g+\delta)\,k(t) \tag{II.1}$$

The Cob-Douglas production function $Y(t) = [A(t)\,L(t)]^{(1-\alpha)}\,K(t)^{\alpha}$ can be written in effective units as $y(t) = k(t)^{\alpha}$. Differentiating this expression with respect to t, we get

$$\frac{dy(t)}{dt} = \alpha \frac{dk(t)}{dt} k(t)^{(\alpha-1)}$$

Thus, substituting equation (II.1) in this expression and using the production function in effective units, we get the following nonlinear differential equation in $y(t)$

$$\begin{aligned} \frac{dy(t)}{dt} &= \alpha k(t)^{\alpha}\left[sk(t)^{(\alpha-1)} - (n+g+\delta)\right] \\ &= \alpha y(t)\left[sy(t)^{\frac{(\alpha-1)}{\alpha}} - (n+g+\delta)\right] \end{aligned} \tag{II.2}$$

The steady state of $y(t)$, y^{*}, is the point at which $\frac{dy(t)}{dt} = 0$ and it is therefore[94]

$$y^{*} = \left(\frac{s}{n+g+\delta}\right)^{\frac{\alpha}{1-\alpha}}$$

[94]There is obviously another solution to $\frac{dy(t)}{dt} = 0$ which is the trivial solution $y = 0$. However, this point is not stable, since any perturbation which takes y off this point will cause this variable to move eventually to y^{*}.

The nonlinear differential equation in $y\,(t)$, equation (II.2) can be rewritten as follows

$$\frac{d\left[\ln y\,(t)\right]}{dt} = \alpha\left[s\exp\left(\frac{\alpha-1}{\alpha}\ln y\,(t)\right) - (n+g+\delta)\right] \tag{II.3}$$

Now, we can take a first order Taylor's approximation of $\ln y\,(t)$ around the steady state. First, we need to calculate the first derivative of equation (II.3) evaluated at the steady state (SS)

$$\left.\frac{d\left[\frac{d(\ln y(t))}{dt}\right]}{d\ln y\,(t)}\right|_{SS} = -\left.(1-\alpha)\,sy\,(t)^{\frac{\alpha-1}{\alpha}}\right|_{SS} = -\,(1-\alpha)\,(n+g+\delta)$$

Therefore, the linearized equation is[95]

$$\begin{aligned}\frac{d\left[\ln y\,(t)\right]}{dt} &= -\,(1-\alpha)\,(n+g+\delta)\left[\ln y\,(t) - \ln y^{*}\right] \\ &= -\,(1-\alpha)\,(n+g+\delta)\ln y\,(t) + \alpha\,(n+g+\delta)\left[\ln s - \ln\,(n+g+\delta)\right]\end{aligned} \tag{II.4}$$

Now, we need to write this equation in terms of per capita output, $\tilde{y}\,(t)$. For this purpose, we make use of the following identity

$$\tilde{y}\,(t) = A\,(t)\,y\,(t) \tag{II.5}$$

Taking logarithms and differentiating with respect to t, we get

$$\frac{d\ln\tilde{y}\,(t)}{dt} = \frac{d\ln A\,(t)}{dt} + \frac{d\ln y\,(t)}{dt} \tag{II.6}$$

[95]Note that the first term of Taylor's approximation is lost since the function evaluated at the steady state is equal to zero.

We know that technology, $A(t)$, grows at a rate g each time period. Consequently,

$$\ln A(t) = \ln A(0) + gt \tag{II.7}$$

Differentiating equation (II.7) and using equations (II.4), (II.5) and (II.6), the following linear differential equation in $\tilde{y}(t)$ is obtained

$$\begin{aligned}
\frac{d \ln \tilde{y}(t)}{dt} &= g + (1-\alpha)(n+g+\delta)\ln A_0 + \\
&+\alpha(n+g+\delta)(\ln s - \ln(n+g+\delta)) + \\
&+(1-\alpha)(n+g+\delta)gt - (1-\alpha)(n+g+\delta)\ln \tilde{y}(t)
\end{aligned}$$

II.ii. Derivation of the exact discrete Solow growth model.

Let us first assume the following general linear stochastic differential equation in a certain variable x

$$dx(t) = [a(t)x(t) + b(t)]\,dt + c(t)\,dW(t) \tag{II.8}$$

The explicit solution of this general equation is well known (see for example Øksendal (1998)) and equal to

$$x_t = \Upsilon_{t,t_0}\left[x_{t_0} + \int_{t_0}^{t} \Upsilon_{s,t_0}^{-1} b(s)\,ds + \int_{t_0}^{t} \Upsilon_{s,t_0}^{-1} c(s)\,dW(s)\right] \tag{II.9}$$

where Υ_{t,t_0} is the fundamental solution

$$\Upsilon_{t,t_0} = \int_{t_0}^{t} a(s)\,ds$$

We are interested in finding the solution to equation (3.2), that is, the following linear stochastic differential equation:

$$d\left(\ln \tilde{y}_i(t)\right) = \left[\mu_i + (1-\lambda_i) g_i + (1-\lambda_i) g_i t - (1-\lambda_i) \ln \tilde{y}_i(t)\right] dt + \sigma_i dW_i(t) \tag{II.10}$$

This equation is a special case the general form in equation (II.8). It is somewhat simpler to solve since neither $a(t)$ or $c(t)$ are actually functions of time. Therefore, using the general solution depicted in equation (II.9), the solution of equation (II.10) can be written as

$$\begin{aligned} \ln \tilde{y}_{it} &= \Upsilon_{t,t-1}\left[\ln \tilde{y}_{it-1} + \int_{t-1}^{t} \Upsilon_{s,t-1}^{-1}\left(\mu_i + (1-\lambda_i) g_i + (1-\lambda_i) g_i s\right) ds + \right. \\ &\left. + \int_{t-1}^{t} \Upsilon_{s,t-1}^{-1} \sigma_i dW_i(s)\right] \end{aligned} \tag{II.11}$$

where $\Upsilon_{t,t-1}$ is the fundamental solution

$$\begin{aligned} \Upsilon_{t,t-1} &= \exp\left[\int_{t-1}^{t} -(1-\lambda_i)\, ds\right] = \exp\left[-(1-\lambda_i)(t-(t-1))\right] = \\ &= \exp\left[-(1-\lambda_i)\right] \end{aligned}$$

Consequently $\Upsilon_{s,t-1}^{-1} = \exp\left[-(1-\lambda_i)(s-t+1)\right]$. As a first step, we need to evaluate the following integral

$$I_1 = \int_{t-1}^{t} \left[\mu_i + (1-\lambda_i) g_i + (1-\lambda_i) g_i s\right] \exp\left[(1-\lambda_i)(s-t+1)\right] ds$$

which can be written as the sum of two integrals

$$I_1 = \left[\mu_i + (1-\lambda_i) g_i\right] I_2 + (1-\lambda_i) g_i I_3 \tag{II.12}$$

where

$$I_2 = \int_{t-1}^{t} \exp\left[(1-\lambda_i)(s-t+1)\right]$$

and

$$I_3 = \int_{t-1}^{t} s \exp\left[(1-\lambda_i)(s-t+1)\right]$$

The solution to I_2 is straightforward

$$I_2 = \frac{\mu_i + (1-\lambda_i) g_i}{(1-\lambda_i)} \left[\exp\left[(1-\lambda_i)\right] - 1\right] \tag{II.13}$$

Integrating by parts, the solution to I_3 is found

$$I_3 = \frac{\exp\left[(1-\lambda_i)\right]}{(1-\lambda_i)} t - \frac{t-1}{(1-\lambda_i)} - \frac{\exp\left[(1-\lambda_i)\right] - 1}{(1-\lambda_i)^2} \tag{II.14}$$

Substituting equations (II.13) and (II.14) into equation (II.12), the following expression for I_1 is found

$$I_1 = \frac{\mu_i - \lambda_i g_i}{(1-\lambda_i)} \left(\exp\left[(1-\lambda_i)\right] - 1\right) + \frac{g_i}{(1-\lambda_i)} + g_i \left(\exp\left[(1-\lambda_i)\right] - 1\right) t \tag{II.15}$$

Thus, using the fundamental solution and equation (II.15), the solution in equation (II.11) is the following

$$\begin{aligned} \ln \tilde{y}_{it} &= \left\{1 - \exp\left[-(1-\lambda_i)\right]\right\} \left(\frac{\mu_i - \lambda_i g_i}{1-\lambda_i}\right) + g_i \exp\left[-(1-\lambda_i)\right] + \\ &\quad + g_i \left\{1 - \exp\left[-(1-\lambda_i)\right]\right\} t + \exp\left[-(1-\lambda_i)\right] \ln \tilde{y}_{it-1} + \\ &\quad + \exp\left[-(1-\lambda_i)\right] \int_{t-1}^{t} \sigma_i \exp\left[(1-\lambda_i)(s-t+1)\right] dW(s) \end{aligned} \tag{II.16}$$

This last equation can be estimated as long as we find out what the distribution of the error term is, that is, we need to find the distribution of

$$\xi_t = \exp\left[-\left(1-\lambda_i\right)t\right]\sigma_i \int_{t-1}^{t} \exp\left[\left(1-\lambda_i\right)s\right] dW\left(s\right)$$

Since the integrand, $\exp\left[\left(1-\lambda_i\right)s\right]$, is non-random, the integral

$$\int_{t-1}^{t} \exp\left[\left(1-\lambda_i\right)s\right] dW\left(s\right)$$

is normally distributed and its mean is equal to zero (see for example Nielsen (1999)). As a direct consequence, the random variable ξ_t is also normally distributed with mean zero. The variance of ξ_t can, therefore, be calculated using the formula for the variance of a stochastic integral with a non-random integrand. In this particular case, the variance of ξ_t is given by

$$\begin{aligned} Var\left(\xi_t\right) &= \exp\left[-2\left(1-\lambda_i\right)t\right]\sigma_i^2 \int_{t-1}^{t} \exp\left[2\left(1-\lambda_i\right)s\right] ds \\ &= \frac{\sigma_i^2}{2\left(1-\lambda_i\right)}\left\{1-\exp\left[-2\left(1-\lambda_i\right)\right]\right\} \end{aligned}$$

Recall from the text in chapter 3 that $\epsilon\tilde{\ }N\left(0,1\right)$. Equation (II.16) can then be written as

$$\begin{aligned} \ln\tilde{y}_{it} &= \left\{1-\exp\left[-\left(1-\lambda_i\right)\right]\right\}\left(\frac{\mu_i-\lambda_i g_i}{1-\lambda_i}\right)+g_i\exp\left[-\left(1-\lambda_i\right)\right]+ \\ &\quad +g_i\left\{1-\exp\left[-\left(1-\lambda_i\right)\right]\right\}t+\exp\left[-\left(1-\lambda_i\right)\right]\ln\tilde{y}_{it-1}+ \\ &\quad +\frac{\sigma_i}{\left[2\left(1-\lambda_i\right)\right]^{\frac{1}{2}}}\left\{1-\exp\left[-2\left(1-\lambda_i\right)\right]\right\}^{\frac{1}{2}}\epsilon_{it} \end{aligned}$$

which is the exact discretized version of the linear stochastic differential equation (II.10).

II.iii. Linearization of the Ramsey-Cass-Koopmans growth model.

This section outlines the steps of the linearization of the system on nonlinear differential equations obtained from the derivation of the RCK model. The steps are equivalent to those of the Solow growth model, the only difference is that now we are dealing with a system of equations. For simplicity of exposition, the system is written below again

$$\frac{dy}{dt} = \alpha y\left[y^{\frac{\alpha-1}{\alpha}}s-\Gamma\right] \tag{II.17}$$

$$\frac{ds}{dt} = (1-s)\left[\alpha y^{\frac{\alpha-1}{\alpha}}\left(s-\frac{1}{\theta}\right)-\alpha\Gamma+\frac{\Phi}{\theta}\right] \tag{II.18}$$

where

$$\Gamma = n+g+\delta$$
$$\Phi = \rho+\theta g+\delta$$

The steady state values of y and s are:

$$y^* = \left(\frac{\alpha}{\Phi}\right)^{\frac{\alpha}{1-\alpha}} \tag{II.19}$$

$$s^* = \frac{\alpha\Gamma}{\Phi} \tag{II.20}$$

The system of nonlinear equations (II.17) and (II.18) can be written as

$$\frac{d\ln y}{dt} = \alpha\left[\exp\left(\frac{\alpha-1}{\alpha}\ln y\right)\exp\left(\ln s\right)-\Gamma\right]$$
$$\frac{d\ln s}{dt} = \left[\exp\left(-\ln s\right)-1\right]\left\{\alpha\exp\left(\frac{\alpha-1}{\alpha}\ln y\right)\left[\exp\left(\ln s\right)-\frac{1}{\theta}\right]-\alpha\Gamma+\frac{\Phi}{\theta}\right\}$$

To calculate the first order Taylor's approximation of the system around the steady state, we need to calculate all four first order partial derivatives of the system evaluated at the steady state (SS). These are shown below

$$\begin{aligned}
\frac{\partial\left(\frac{d\ln y}{dt}\right)}{\partial\ln y} &= -(1-\alpha)\,y^{\frac{\alpha-1}{\alpha}}s\Big]_{SS} = -(1-\alpha)\,\Gamma \\
\frac{\partial\left(\frac{d\ln y}{dt}\right)}{\partial\ln s} &= \alpha y^{\frac{\alpha-1}{\alpha}}s\Big]_{SS} = \alpha\Gamma \\
\frac{\partial\left(\frac{d\ln s}{dt}\right)}{\partial\ln y} &= -(1-\alpha)\left(\frac{1}{s}-1\right)y^{\frac{\alpha-1}{\alpha}}\left(s-\frac{1}{\theta}\right)\Big]_{SS} \\
&= -\frac{(1-\alpha)\,(\Phi-\alpha\Gamma)\,(\alpha\theta\Gamma-\Phi)}{\alpha^2\theta\Gamma} \\
\frac{\partial\left(\frac{d\ln s}{dt}\right)}{\partial\ln s} &= -\frac{1}{s}\left[\alpha y^{\frac{\alpha-1}{\alpha}}\left(s-\frac{1}{\theta}\right)-\alpha\Gamma+\frac{\Phi}{\theta}\right]+\alpha\left(\frac{1}{s}-1\right)y^{\frac{\alpha-1}{\alpha}}s\Big]_{SS} \\
&= \Phi-\alpha\Gamma
\end{aligned}$$

Thus, the linearized system is the following

$$\begin{aligned}
\frac{d\ln y}{dt} &\simeq -(1-\alpha)\,\Gamma\,[\ln y-\ln y^*]+\alpha\Gamma\,[\ln s-\ln s^*] \\
\frac{d\ln s}{dt} &\simeq -\frac{(1-\alpha)\,(\Phi-\alpha\Gamma)\,(\alpha\theta\Gamma-\Phi)}{\alpha^2\theta\Gamma}\,[\ln y-\ln y^*]+(\Phi-\alpha\Gamma)\,[\ln s-\ln s^*]
\end{aligned}$$

Substituting y^* and s^* for their corresponding values (see equations (II.19) and (II.20)), the system reduces to

$$\begin{aligned}
\frac{d\ln y}{dt} &= -\alpha\Gamma\ln\Gamma-(1-\alpha)\,\Gamma\ln y+\alpha\Gamma\ln s \\
\frac{d\ln s}{dt} &= -(\Phi-\alpha\Gamma)\left[\frac{\Phi}{\alpha\theta\Gamma}\ln\left(\frac{\alpha}{\Phi}\right)+\ln\Gamma\right]-\frac{(1-\alpha)\,(\Phi-\alpha\Gamma)\,(\alpha\theta\Gamma-\Phi)}{\alpha^2\theta\Gamma}\ln y+ \\
&\quad+(\Phi-\alpha\Gamma)\ln s
\end{aligned}$$

This linear system of differential equations needs to be written in terms of per capita output, $\tilde{y}$. Following the same steps as in the Solow model, that is, using equations

(II.5), (II.6) and(II.7), the following system of linear differential equations is obtained

$$\begin{aligned}
\frac{d\ln\tilde{y}}{dt} &= g+\Gamma\left[(1-\alpha)\ln A(0)-\alpha\ln\Gamma\right]+(1-\alpha)\Gamma gt-(1-\alpha)\Gamma\ln\tilde{y}+\alpha\Gamma\ln s \\
\frac{d\ln s}{dt} &= -(\Phi-\alpha\Gamma)\left[\frac{\Phi}{\alpha\theta\Gamma}\ln\left(\frac{\alpha}{\Phi}\right)+\ln\Gamma-\frac{(1-\alpha)}{\alpha}\left(1-\frac{\Phi}{\alpha\theta\Gamma}\right)\ln A(0)\right]+ \\
&\quad +\frac{(1-\alpha)(\Phi-\alpha\Gamma)(\alpha\theta\Gamma-\Phi)}{\alpha^2\theta\Gamma}gt-\frac{(1-\alpha)(\Phi-\alpha\Gamma)(\alpha\theta\Gamma-\Phi)}{\alpha^2\theta\Gamma}\ln\tilde{y}+ \\
&\quad +(\Phi-\alpha\Gamma)\ln s
\end{aligned}$$

The system can be written in matrix form as

$$dX=\left[PX+Q+Rt\right]dt$$

where $X'=\begin{pmatrix}\ln\tilde{y} & \ln s\end{pmatrix}$ and P, Q and R are (2×2), (2×1) and (2×1) matrices of coefficients respectively with individual elements, p_{ij}, q_{ij} and r_{ij}

$$\begin{aligned}
p_{11} &= -(1-\alpha)\Gamma \\
p_{12} &= \alpha\Gamma \\
p_{21} &= -\frac{(1-\alpha)(\Phi-\alpha\Gamma)(\alpha\theta\Gamma-\Phi)}{\alpha^2\theta\Gamma} \\
p_{22} &= \Phi-\alpha\Gamma \\
q_{11} &= g+\Gamma\left[(1-\alpha)\ln A(0)-\alpha\ln\Gamma\right] \\
q_{21} &= -(\Phi-\alpha\Gamma)\left[\frac{\Phi}{\alpha\theta\Gamma}\ln\left(\frac{\alpha}{\Phi}\right)+\ln\Gamma-\frac{(1-\alpha)}{\alpha}\left(1-\frac{\Phi}{\alpha\theta\Gamma}\right)\ln A(0)\right] \\
r_{11} &= (1-\alpha)\Gamma g \\
r_{21} &= \frac{(1-\alpha)(\Phi-\alpha\Gamma)(\alpha\theta\Gamma-\Phi)}{\alpha^2\theta\Gamma}g
\end{aligned}$$

II.iv. Derivation of the exact discrete Ramsey-Cass-Koopmans model.

This section outlines the solution of the linear system of stochastic equations obtained in the RCK model. The steps are equivalent to those of the Solow growth model, however, now we are dealing with matrices. In matrix form the system of stochastic differential equations can be written in the following form

$$dX = [PX + Q + Rt]\, dt + V dW \tag{II.21}$$

The explicit solution of this system is therefore the following

$$X_t = \Upsilon_{t,t-1}\left[X_{t-1} + \int_{t-1}^{t} \Upsilon_{s,t-1}^{-1}(Q + Rs)\, ds + \sum_{l=1}^{2}\int_{t-1}^{t} \Upsilon_{s,t-1}^{-1} V^l dW_1(s)\right] \tag{II.22}$$

where V^l is the lth column of matrix V and $\Upsilon_{t,t-1}$ is the fundamental matrix satisfying both, $\Upsilon_{t-1,t-1} = I$ and the homogeneous matrix stochastic differential equation

$$d\Upsilon_{t-1,t-1} = P\Upsilon_{t,t-1}dt + \sum_{l=1}^{2} V^l \Upsilon_{t,t-1} dW_1(t)$$

In this case, the fundamental matrix is

$$\Upsilon_{t,t-1} = \exp\left[P\left(t - (t-1)\right)\right] = \exp(P)$$

where

$$\exp(P) = \sum_{j=0}^{\infty} \frac{1}{j!} P^j$$

and as a consequence

$$\Upsilon_{s,t-1}^{-1} = \exp\left[-P\left(s-(t-1)\right)\right]$$

As a first step towards the solution in equation (II.22), we need to evaluate the following integral

$$I_1 = \int_{t-1}^{t} \exp\left[-P\left(s-(t-1)\right)\right]\left(Q+Rs\right)ds$$

This integral can be written as the sum of two integrals, I_2 and I_3, as

$$I_1 = I_2 Q + I_3 R \tag{II.23}$$

where

$$I_2 = \int_{t-1}^{t} \exp\left[-P\left(s-(t-1)\right)\right]ds$$

and

$$I_3 = \int_{t-1}^{t} \exp\left[-P\left(s-(t-1)\right)\right]sds$$

The solution of I_2 is straightforward

$$I_2 = \left[I - \exp\left(-P\right)\right]P^{-1} \tag{II.24}$$

The solution to I_3 is found by integrating by parts

$$I_3 = \left[I - \exp\left(-P\right)\right]P^{-1}t - P^{-1} + \left[I - \exp\left(-P\right)\right]\left(P^{-1}\right)^2 \tag{II.25}$$

Substituting equations (II.24) and (II.25) into equation (II.23), we find the following expression for I_1

$$I_1 = [I - \exp(-P)] P^{-1} \left(Q + P^{-1}R + Rt\right) - P^{-1}R \qquad \text{(II.26)}$$

Using the fundamental solution and substituting equation (II.26) into the solution in equation (II.22), the following equation is found

$$X_t = \exp(P) X_{t-1} + (\exp(P) - I) P^{-1} \left(Q + P^{-1}R + Rt\right) - \exp(P) P^{-1}R + \xi_t$$

where

$$\xi_t = \sum_{l=1}^{2} \left[\int_{t-1}^{t} \exp(P(t-s)) V_l dW_l(s)\right]$$

The error term, ξ_t, is a composite of two integrals, with both integrands being non-random. Therefore, the components of ξ_t are normally distributed with mean zero (see for example Nielsen (1999)) and as a result, ξ_t is also normally distributed with zero mean. The covariance matrix of this error term can then be calculated according to the formula for the covariance of a stochastic integral with a non-random integrand

$$E[\xi_t \xi_t'] = \int_{t-1}^{t} \exp(P(t-s)) VV' \exp(P'(t-s)) \, ds$$

We can also change the variable of integration from s to u where $u = t - s$, to obtain the following expression for the covariance matrix

$$E[\xi_t \xi_t'] = \int_{0}^{1} \exp(Pu) VV' \exp(P'u) \, du$$

or, equivalently, as it is written in chapter 3

$$E\left[\xi_t \xi_t'\right] = \int_0^1 \exp\left(Ps\right) VV' \exp\left(P's\right) ds$$

APPENDIX III Appendix to chapter 4.

This appendix gives detailed explanations of two algorithms used in this thesis for the optimization of the likelihood function of the switching regressions model; the Expectation-Maximization algorithm of Hartley (1977,1978) and the Simulated Annealing algorithm (Corana *et al* (1987)). These two algorithms are compared in chapter 5 in terms of their performance.

III.i. The Expectation-Maximization algorithm.

A description of this algorithm is provided in van Norden and Vigfusson (1996). However, it is worth providing here a brief explanation since its performance compared to other algorithms is studied in chapter 5. There are also GAUSS procedures available from the Bank of Canada which implement this algorithm. The Expectation-Maximization (EM) algorithm is an iterative procedure used to solve the first order conditions of the likelihood function which uses Ordinary Least Squares (OLS) and Weighted Least Squares (WLS). The general idea is to replace the unobservables, that is y_t^*, by their conditional expectations in the model depicted in equations (4.1) and (4.2). As before, the probability density function of Δy_t is defined by

$$g(\Delta y_t) = \Phi\left(-\beta_{3,0} - \beta_{3,1} z_t\right) f_{1,t} + \Phi\left(\beta_{3,0} + \beta_{3,1} z_t\right) f_{2,t}$$

where again $f_{r,t}(\Delta y_t)$ for $r = 1, 2$ is defined in equation (4.4). Thus, from equation (4.5), the log-likelihood function is the following

$$L = \sum_{t=1}^{T} \ln g(\Delta y_t)$$

Differentiating with respect to $\beta_{v,s}$ for $r = 1, 2$ and $s = 1, \ldots, k_r$ the following first order conditions for optimisation are found

$$\frac{\partial L}{\partial \beta_{r,0}} = \sum_{t=1}^{T} w_{rt}(\triangle y_t)\left(\frac{\Delta y_t - \beta_{r,0} - \sum_{j=1}^{k_r} \beta_{r,,j}\Delta y_{t-j}}{\sigma_r}\right) = 0 \qquad \text{(III.1)}$$

$$\frac{\partial L}{\partial \beta_{r,s}} = \sum_{t=1}^{T} w_{rt}(\triangle y_t)\left(\frac{\Delta y_t - \beta_{r,0} - \sum_{j=1}^{k_r} \beta_{r,,j}\Delta y_{t-j}}{\sigma_r}\right) = 0 \qquad \text{(III.2)}$$

where

$$\begin{array}{l} w_{1t}(\triangle y_t) = \Phi\left(-\beta_{3,0} - \beta_{3,1} z_t\right) \frac{f_{1t}}{g(\triangle y_t)} \\ w_{2t}(\triangle y_t) = \Phi\left(\beta_{3,0} + \beta_{3,1} z_t\right) \frac{f_{2t}}{g(\triangle y_t)} \end{array} \qquad \text{(III.3)}$$

These will be later used as the calculated weights for the WLS regression. Differentiating with respect to $\beta_{3,0}$ and β_{31}, the next two first order conditions are found

$$\frac{\partial L}{\partial \beta_{3,0}} = \sum_{t=1}^{T} \frac{\phi\left(-\beta_{3,0} - \beta_{3,1} z_t\right) f_{1t} - \phi\left(-\beta_{3,0} - \beta_{3,1} z_t\right) f_{2t}}{g(\triangle y_t)} = 0 \qquad \text{(III.4)}$$

$$\frac{\partial L}{\partial \beta_{3,1}} = \sum_{t=1}^{T} \frac{\phi\left(-\beta_{3,0} - \beta_{3,1} z_t\right) f_{1t} - \phi\left(-\beta_{3,0} - \beta_{3,1} z_t\right) f_{2t}}{g(\triangle y_t)} z_t = 0 \qquad \text{(III.5)}$$

Since the conditional expectation of y_t^* is

$$\begin{aligned} \xi_{3t} &= E\left[y_t^* | \triangle y_t\right] \\ &= \beta_{30} + \beta_{31} z_t - w_{1t}(\triangle y_t) \frac{\phi\left(-\beta_{30} - \beta_{31} z_t\right)}{\Phi\left(-\beta_{30} - \beta_{31} z_t\right)} \\ &\quad + w_{2t}(\triangle y_t) \frac{\phi\left(-\beta_{30} - \beta_{31} z_t\right)}{\Phi\left(\beta_{30} + \beta_{31} z_t\right)} \end{aligned} \qquad \text{(III.6)}$$

equations (III.4) and (III.5) can be written as

$$\frac{\partial L}{\partial \beta_{3,0}} = \sum_{t=1}^{T} (\xi_{3t} - \beta_{30} - \beta_{31} z_t) = 0 \qquad \text{(III.7)}$$

$$\frac{\partial L}{\partial \beta_{3,1}} = \sum_{t=1}^{T} (\xi_{3t} - \beta_{30} - \beta_{31} z_t)\, z_t = 0 \qquad \text{(III.8)}$$

Also, the first derivatives with respect to σ_r are

$$\frac{\partial L}{\partial \sigma_r} = \sum_{t=1}^{T} w_{rt} (\Delta y_t) \left[\sigma_r^2 - \left(\Delta y_t - \beta_{r,0} - \sum_{j=1}^{k_r} \beta_{r,,j} \Delta y_{t-j} \right)^2 \right] = 0 \qquad \text{(III.9)}$$

for $r = 1, 2$.

The algorithm then has 2 sequential stages in each iteration. First, given the values of the parameters to be estimated, the conditional expectation of the unobservable variable is formed. This is the E step. Then, given this conditional expectation, the likelihood function is maximized, resulting in a new set of parameter values which will be used in the next iteration. This is the M step. Thus, the steps of the algorithm are as follows:

<u>*Step 1*</u>. For some initial values of the parameters to be estimated in the model, calculate the matrices of weights for the WLS regression (see *Step 2*) defined as $\mathbf{W}_r = diag\{w_{r1}, ..., w_{rT}\}$ for $r = 1, 2$ using equation (III.3). For the same values of the parameters, calculate the conditional expectations vector $\boldsymbol{\xi}_3 = (\ \xi_{31},\ \ldots,\ \xi_{3T}\ )'$ using equation (III.6).

<u>*Step 2.*</u> Next, from equations (III.1) and (III.2), the estimates of the parameters $\boldsymbol{\beta}_r$ for $r = 1, 2$ can be calculated using WLS as follows

$$\boldsymbol{\beta}_r = \left[\mathbf{X}_r'\mathbf{W}_r\mathbf{X}_r\right]^{-1}\left[\mathbf{X}_r'\mathbf{W}_r\mathbf{y}\right]$$

where

$$\boldsymbol{\beta}_r = \begin{pmatrix} \beta_{r0} \\ \beta_{r1} \\ \vdots \\ \beta_{rk_1^r} \end{pmatrix} \quad \mathbf{X}_r = \begin{pmatrix} 1 & \triangle y_0 & \cdots & \triangle y_{1-k_1^r} \\ 1 & \triangle y_1 & \cdots & \triangle y_{2-k_1^r} \\ \vdots & \vdots & & \vdots \\ 1 & \triangle y_{T-1} & \cdots & \triangle y_{T-k_1^r} \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} \triangle y_1 \\ \vdots \\ \triangle y_T \end{pmatrix} \tag{III.10}$$

and the estimates of the parameters in the switching equation

$$\boldsymbol{\beta}_3 = \begin{pmatrix} \beta_{30} \\ \beta_{31} \end{pmatrix} \tag{III.11}$$

are obtained using an OLS regression of the conditional expectation vector of y^* on $\mathbf{Z}$

$$\boldsymbol{\beta}_3 = \left[\mathbf{Z}'\mathbf{Z}\right]^{-1}\mathbf{Z}'\boldsymbol{\xi}_3$$

where $\mathbf{Z} = \begin{pmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_T \end{pmatrix}$

<u>*Step 3.*</u> Given the estimated parameters in *Step 2* and the matrices of weights obtained in *Step 1*, the estimated variances of each regime are calculated as

$$\sigma_r^2 = \frac{1}{\sum w_r\left(\triangle y_t\right)}\left(\mathbf{y} - \mathbf{X}_r\boldsymbol{\beta}_r\right)'\mathbf{W}_r\left(\mathbf{y} - \mathbf{X}_r\boldsymbol{\beta}_r\right)$$

for $r = 1, 2$.

Steps 1 to 3 are repeated iteratively using the new parameter estimates as the new initial values until the selected convergence criteria is reached. However, even if the EM algorithm always converges to a local maximum of the likelihood function as Hartley states, there still remains the problem of multimodality of the likelihood function.

III.ii. The Simulated Annealing algorithm.

The discussion of the algorithm in this section relates to maximization problems, whereas the thermodynamics example provided in chapter 4 relates to minimisation problems.

Step 1. In this step, several parameter values need to be initialized before starting the algorithm. First of all, initial values of the parameters to be estimated need to be supplied to the algorithm. These are stored in a column vector, Θ_0. This is also needed whenever a function is optimised numerically. The advantage when using the SA algorithm is that the choice of these initial values is not crucial, whereas this is not the case when using traditional algorithms. In addition, the following parameters also need to be initialized for the simulated annealing to work:

1.vm^0: this is an initial step vector. It is a column vector with rows equal to the number of parameters to be estimated. The algorithm looks for potential new optima in a circle centered on the current parameter value of radius equal to the corresponding element of this vector step vector. The initial step vector is not very important, because after the first iteration it will be modified according to the temperature (see below). In the thermodynamic example, this vector corresponds to the maximum movement of the molecules at a given temperature. If the vector is big, which happens when the temperature is high, the molecules can move around a lot. If the vector is small, which occurs at low temperatures, the molecules move in a small area.

2.T^0: the initial temperature. The initial temperature is very important because if it is too small the algorithm may never escape the local maxima. This is because it influences the step vector over which the points are selected via its influence on the number of downhill moves that are permitted. Therefore, the temperature has to be high enough for the step vector to be big enough to cover all the possible points in the domain of the parameters.

3.ϵ: a criterion for termination of the algorithm.

4.N_ϵ: a number of successive temperature reductions after which the algorithm tests for termination using ϵ.

5.N_S: a number of cycles before the step vector is adjusted and a varying criterion c for the step vector.

6.N_T: a number of loops before the temperature is cut down using a reduction coefficient r_T which also needs to be initialized.

7.lb and ub: two column vectors with the lower and upper bounds of the parameters respectively.

Once this parameters have been initialized, the likelihood function is computed for the initial parameter values $L_0 = L(\Theta_0)$. This is equivalent to the current level of energy in the system. The current optimum values of the parameters and the likelihood function are initialized to the current values, $\Theta_{opt} = \Theta_0$ and $L_{opt} = L_0$. The last point accepted for evaluation is also initialized $\Theta_l = \Theta_0$, since this is just the first evaluation of the function.

Step 2. The algorithm now starts trying new points given the allowable movement in the molecules, vm. A new point, Θ_{trial}, is generated by performing a random move for the jth element of the parameter vector $\Theta_{trial} = \Theta_l + \eta v m_j e_j$ where η is a pseudorandom number in the range $[-1, 1]$, e_j is a dummy column vector with a 1 in the jth row and vm_j is the jth element of the step vector vm.

Step 3. If the jth element of the new generated parameter vector lies outside the domain given by the lower and upper bounds of the parameter vector, lb and ub, then the algorithm generates a random point between the bounds for trial.

Step 4. In this step, the algorithm checks whether the likelihood function is higher at this point. In the thermodynamics example this is equivalent to check whether the energy of the system has decreased. Thus, the algorithm evaluates the likelihood function at the new generated point. If the value of the function increases then accept the new point, $\Theta_l = \Theta_{trial}$. If the value of the function is greater than any other point so far, store it as a new optimum $\Theta_{opt} = \Theta_{trial}$. Conversely if the value of the function decreases accept the point using the Metropolis criteria. This is a very important part of the algorithm and what makes it different from traditional methods of optimisation. Even if the value of the function decreases, there is still a probability of accepting this point for further evaluations. This is the reason why the algorithm can escape from local optima and find the global maximum somewhere else. The Metropolis criteria uses the Boltzmann probability distribution and accepts the point with probability π

$$\pi = e^{\left(\frac{L(\Theta_{trial}) - L(\Theta_l)}{T}\right)}$$

That is, the smaller the extent of the downhill move and the higher T, the more probable the downhill move is to be permitted. To implement this criteria, a pseudorandom number is generated in the range $[0, 1]$. If the generated number is less than π, then the point is accepted, otherwise, the point is rejected.

Step 5. Steps 2 to 4 are carried out sequentially each time allowing only the jth element of the parameter vector to be optimised to change. Every one of the loops uses the last accepted point $\mathbf{\Theta}_l$ as the starting point for the loop.

Step 6. Repeat steps 2 to 5 N_S times to give the algorithm a chance to form a good view of the function.

Step 7. Once the algorithm has a rough idea of how the function looks like, the step length vector, vm, can be adjusted. This is accomplished by modifying the step length vector so that approximately half of the total number of evaluations are accepted. The reason behind this, is that too many accepted function evaluations with respect to the number of rejections implies that the function is being examined with steps that are too small, whereas too many rejections mean that the points chosen for trial are being generated too far from the current point.

Let $ratio_j$ be equal to the number of accepted evaluations when the jth element of the parameter vector is changed divided by N_S. Update each jth element of the step length vector vm as follows

$$\begin{array}{lll} \text{if } ratio_j > 0.6 \text{ then} & vm_j = vm_j\left(1 + c_j\frac{ratio_j - 0.6}{0.4}\right) \\ \text{if } ratio_j < 0.4 \text{ then} & vm_j = \frac{vm_i}{1 + c_j\frac{0.4 - ratio_j}{0.4}} \\ \text{otherwise} & vm_j = vm_j \end{array}$$

Therefore the parameter c_j initialized in *Step 1* is used to control the variation of jth element of the step vector.

Step 8. Repeat steps 2 to 7 N_T times. After this, the temperature can be reduced (see below).

Step 9. Check whether the termination criteria is met using ϵ and N_ϵ. That is, if for the last N_ϵ temperatures, the final function values do not differ from the corresponding value at the current temperature by more than ϵ and, also, the difference between the final function value at the current temperature and the current optimal function is less than ϵ then the program halts. If the termination criteria is not met then reduce the temperature according to $T = r_T N_T$ and go back to step 2 again.

Some considerations about the parameters in the simulated annealing algorithm.

The initial temperature is very important because if it is too small the algorithm may never escape the local maxima. This is because it influences the step vector over which the points are selected by influencing the number of downhill moves that are allowed. Therefore, the temperature has to be high enough for the step vector to be big enough to cover all the possible points in the domain of the parameters.

Corana *et al.* (1987) suggested the following parameter values to control the simulated annealing algorithm

$$\begin{aligned}
&N_S = 20 \\
&N_T = \max(100, 5n) \\
&c_i = 2 \qquad \text{for } i = 1, \ldots, n \\
&N_\epsilon = 4 \\
&r_T = 0.85
\end{aligned}$$

where n is the number of parameters to be estimated. However depending on the function and the number of parameters that need to be estimated this algorithm may take too long to run. Goffe *et al.* (1994) suggested ways of choosing the appropriate values of N_T and r_T to reduce execution time. They also suggested putting lower and upper limits when possible for the estimated parameters which allows a lower starting temperature since the domain decreases.

The GAUSS code for Simulated Annealing used in this thesis has been written by E. G. Tsionas and used in Goffe *et al.* (1994)

APPENDIX IV Appendix to chapter 6.

The analysis in chapter 6 uses data for the G7 countries from the first quarter of 1970 to the last quarter of 1994. Data on both output and technology is required. Output data comes from the **International Financial Statistics (IFS)** published by the International Monetary Fund (see chapter 5, section 5.2 for more details about these data). Total Factor Productivity (TFP) is used as a measure of technology. TFP measures the component of output which remains after taking into account labour and capital inputs. To compute this measure several variables and some assumptions are required. First of all, a functional form needs to be assumed for the production function. Following the bulk of the literature in this area and also following the theoretical models of chapter 2, a Cobb-Douglas production function is assumed (see chapter 2, equation (2.1)). Typically, in a Cobb-Douglas production function, the share of labour in output, $1-\alpha$, is assumed to be time invariant. However, it is well known that the share of labour in output changes across time and consequently, in the empirical work of chapter 6, the share of labour is allowed to change over time. Taking logarithms on both sides of the production function, the following measure of TFP (a_t) for each time period t is obtained:

$$a_t = \frac{1}{1-\alpha_t} y_t - l_t - \frac{\alpha_t}{1-\alpha_t} k_t \qquad \text{(IV.1)}$$

where lower case letters denote the logarithm of the variables. Thus, data on output, the capital stock, labour input and a measure of the share of capital in income are needed to be able to compute this measure of TFP. Data on gross fixed capital formation comes also from the **IFS** excluding the USA for which the data comes from the **OECD Main Economic indicators**. These data are seasonally adjusted and transformed into constant prices by using the GDP deflator. The times series on the capital stock are constructed using the perpetual inventory formula

$$K_{i,t} = (1 - \delta_i)^t K_{i,0} + \sum_{J=0}^{t-1} (1 - \delta_i)^j I_{i,t-j}$$

where $K_{i,t}$ and $I_{i,t}$ are the capital stock and the gross investment in country i at time t respectively and δ_i is the depreciation rate in country i. For estimation purposes, the depreciation rate is set to be the same for all the countries and equal to 0.04.[96] For the initial values of $K_{i,0}$ the simple backasting steady state values of the averages of gross investment over the first 12 observations of the sample (that is the first 3 years of data) divided by the rates of depreciation is used. Since in the empirical estimation only the last 92 observations are used, there are enough data points at the beginning of the sample (more than 5 years of quarterly data) to make sure that the calculated capital stock has stabilized by the beginning of the sample used in estimation.

Labour input is calculated as employment times an index of the number of weekly hours worked in manufacturing. Quarterly data for total employment as an index is obtained from the **OECD Main Economic Indicators**, this is multiplied by the total labour force in the indexed year from the **OECD Annual Labour Force Statistics**. The number of weekly hours worked is from the **OECD Main Economic Indicators** and the gaps missing are extrapolated from annual data from the **Yearbook of Labour Statistics**. All these data are not seasonally adjusted so a multiplicative model is used to adjust it.[97]

[96] A depreciation rate of 0.02 was additionally used to construct the capital series. However, the results of the switching regressions models of Chapter 6 were robust to the value of depreciation used to construct the series of capital.

[97] The multiplicative model is a decomposition method. The underlying assumption is that the data series is made up of the multiplication of several components, namely, a trend, a cycle, a seasonal and a random component. Once the seasonal component is identified, the series is divided by the seasonal

Following Hall and Jones (1999) the productivity measure incorporates a correction for natural resources used as inputs. For this purpose, the value added in the mining and quarrying industry (ISIC 2) is subtracted from the measure of output. This rests on the assumption that capital and labour inputs in mining do not add any extra value to the measure of output and, therefore, all the value added of the mining industry is attributed to natural resource inputs. This correction is needed because the United Kingdom discovered oil in the North Sea during the sample. Without this correction in would appear that the UK suddenly became more productive. For this purpose an index of production in the mining and quarrying industry (**UN Statistical Yearbook**) is multiplied by the value added in the mining and quarrying industry in 1985 (**UN National Accounts Statistics: Main Aggregates and detailed tables**). The index of production, however, is annual, therefore the figures have to be extrapolated to quarterly figures.

The share of capital, α, is calculated as one minus the compensation of employees paid by resident producers over GDP at current prices. Both GDP at current prices and the compensation of employees come from the **OECD National Accounts: Main Aggregates**. The data is annual, therefore it needs to be extrapolated to quarterly figures. To smooth the fluctuations of the data for alpha, the actual series, however, is calculated as a moving average process for 9 quarters. That is, for each quarter, the value of alpha is the average of the actual value of alpha in that particular quarter and four quarters behind and four quarters ahead.

To be able to compare all the series, the data on output and capital in constant prices has to be converted to common currency units (US dollars in this case) using an

---

effect to obtain the seasonally adjusted series (see for example Makridakis, Wheelwright and McGee (1983)).

appropriate exchange rate. Here we use the Purchasing Power Parity (PPP) exchange rate in 1985 from the **OECD Industrial Structure Statistics** (1994).

BIBLIOGRAPHY

Abramovitz, M. (1986) "Catching Up, Forging Ahead, and Falling Behind", Journal of Economic History, Vol. 46, No. 2, 385-406.

Aghion, P. and P. Howitt (1992) " A model of growth through creative destruction" Econometrica, Vol. 60, 323-351.

Balke, N. S. and M. A. Wynne (1996) "Are deep recessions followed by strong recoveries? Results for the G-7 countries", Applied Economics, Vol. 28, 889-897.

Barro, R. J. (1991), "Economic Growth in a Cross-Section of Countries", Quarterly Journal of Economics, Vol. 106, 407-444.

Barro, R. J. (1997), "Economic Growth and Convergence" in Determinants of Economic Growth: A Cross-Country Empirical Study, MIT Press,Cambridge.

Barro, R. J., N. G. Mankiw and X. Sala-i-Martin (1995), "Capital mobility in neoclassical models of growth", American Economic Review, Vol. 85, No. 5, 103-115.

Barro, R. J. and X. Sala-i-Martin (1992), "Convergence", Journal of Political Economy, Vol. 100, No. 2, 223-251.

Barro, R. J. and X. Sala-i-Martin (1995), Economic Growth, McGraw Hill, New York.

Barro, R. J. and X. Sala-i-Martin (1997), "Technological Diffussion, Convergence and Growth", Journal of Economic Growth, Vol. 2, No. 1, 1-26.

Baumol, W. J. (1986), "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show". American Economic Review, Vol. 76, No. 5, 1072-1085.

Baumol, W. J. and E. N. Wolff (1988), " Productivity Growth, Convergence, and Welfare: Reply", American Economic Review, Vol. 78, No. 5, 1155-1159.

Bean, C. R. (1991) "The external constraint in the UK" in External constraints on macroeconomic policy: the European experience, Ed. G. Alogoskoufis, L. Papademos and R. Portes, 193-215.

Beaudry, P. and G. Koop (1993), "Do recessions permanently change output?", Journal of Monetary Economics, Vol. 31, 149-163.

Bergstrom, A. R. (1990), "Recent advances in econometrics: Continuous Time Econometric Modelling", Oxford University Press.

Bernard, A.B., (2001), "Trends and transitions in the long run growth of nations" Working Paper, http://mba.tuck.dartmouth.edu/pages/faculty/andrew.bernard /longrunv22.pdf.

Bernard, A.B. and S. Durlauf (1995), "Convergence in International Output", Journal of Applied Econometrics, Vol.10, 97-108.

Bernard, A.B. and S. Durlauf (1996), "Interpreting tests of convergence hypothesis", Journal of Econometrics, Vol. 71, 161-173.

Bernard, A.B. and C. I. Jones (1996a), "Productivity and Convergence Across U.S. States and Industries", Empirical Economics, Vol. 21, No.1, 113-135.

Bernard, A.B. and C. I. Jones (1996b), "Technology and Convergence", The Economic

Journal, Vol. 106, 1037-1044.

Bernard, A.B. and C. I. Jones (1996c), "Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries", The American Economic Review, Vol. 86, No. 5, 1216-1238.

Bernard, A.B. and C. I. Jones (1996d) "Productivity across industries and countries: time series theory and evidence", Review of Economics and Statistics, Vol. 78, No. 1, 135-146.

Binder, M. and M.H. Pesaran (1999), "Stochastic Growth models and their Econometric Implications", Journal of Economic growth, Vol. 4 No. 2, 139-183.

Cameron, G., J. Proudman and S. Redding (1998) "Productivity convergence and international openness", Bank of England, Working Paper Series No. 77.

Cass, D. (1965), "Optimum growth in an aggregative model of capital accumulation", Review of Economic Studies, Vol.32 (Jul), 233-240.

Coakley, J., F. Kulasi, and R. Smith (1996), "Current account solvency and the Feldstein-Horioka puzzle", The Economic Journal, Vol 106, 620-627.

Cochrane, J. H. (1988) "How Big Is the Random Walk in GNP?", Journal of Political Economy, Vol. 96, No. 5, 893-920.

Cogley, T. (1990) "International Evidence on the Size of the Random Walk in Output", Journal of Political Economy, Vol. 98, No. 3, 501-518.

Corana, A., M. Marchesi, C. Martini, and S. Ridella (1987) "Minimizing Multimodal Functions of Continuous Variables with the "Simulated Annealing" Algorithm", ACM Transactions on Mathematical Software, Vol. 13, No. 3, 262-280.

de la Fuente, A. (1997), "The empirics of growth and convergence: a selective review", Journal of Economic Dynamics and Control, Vol. 21, 23-73.

De Long, J. B. (1988), "Productivity Growth, Convergence, and Welfare: Comment", American Economic Review, Vol. 78, No. 5, 1138-1154.

Demery, D. and N.W. Duck (1992) "Are economic fluctuations really persistent?. A reinterpretation of some international evidence", The Economic Journal, Vol. 102 (Sep), 1094-1101.

Diamond, P. (1965), "National debt in a Neoclassical growth model", American Economic Review, Vol. 55, 1126-1150.

Dollar, D. and E. N. Wolff (1994), "Capital intensity and TFP convergence in industry in manufacturing, 1963-1985" in Convergence of Productivity, Cross National Studies and Historical Evidence, W. J. Baumol, R. R. Nelson and E. N. Wolff (eds), Oxford University Press, Oxford, 197-224.

Domar, E. V. (1946), "Capital expansion, rate of growth and employment", Econometrica, Vol. 14, 137-147.

Dowrick, S. (1992), "Thecnological catch up and diverging incomes: patterns of economic growth 1960-88", The Economics Journal, Vol. 102, 600-610.

Dowrick, S. and D. T. Nguyen (1989), "OECD Comparative economic growth 1950-85: catch-up and convergence", American Economic Review, Vol.79, No.5, 1010-1030.

Durlauf, S.N. (1996), "Controversy: On the convergence and divergence of growth rates", The Economic Journal, Vol. 106 (Jul), 1016-1018.

Durlauf, S.N., J.F. Helliwell and B. Raj (1996), "Long-run Economic Growth", Empirical Economics, Vol. 21, 1-10.

Durlauf, S.N. and P.A. Johnson (1995), "Multiple regimes and cross?country growth behaviour", Journal of Applied Econometrics, Vol. 10, 365-384.

Durlauf, S. N. and D. T. Quah (1999) " The new empirics of economic growth", in Handbook of Macroeconomics, J. B. Taylor and M. Woodford (eds), Vol 1A, North-Holland.

Eaton, J. and S. S. Kortum (1995), "Trade in ideas: patenting and productivity in the OECD", National Bureau of Economic Research (Cambridge MA) Working Paper No. 5049.

Evans, P. (1998), "Using panel data to evaluate growth theories", International Ecnomic Review, Vol. 39, Issue 2, 295-306.

Evans, P. and G. Karras (1996a), "Convergence revisited", Journal of Monetary Economics, Vol. 37, No. 2 (Apr), 249-265.

Evans, P. and G. Karras (1996b), "Do economies converge?. Evidence from a panel of US States", Review of Economics and Statistics, Vol.78, No. 3, 384-388.

Féve, P. and Y. le Pen (2000) "On modelling convergence clubs", Applied Economic Letters, Vol 7, 311-314.

Frenkel, J. A. and A. Razin (1996), Fiscal policies and growth in the world economy, MIT Press.

Friedman, M. (1992), "Do old fallacies ever die?", Journal of Economic Literature, Vol. 30, No. 4 (Dec), 2129-2132.

Galor, O. (1996), "Convergence? Inferences from theoretical models", The Economic Journal, Vol. 106 (Jul), 1056-1069.

Goffe, W. L., G. D. Ferrier and J. Rogers (1994), "Global optimisation of statistical functions with simulated annealing", Journal of Econometrics, Vol. 60, 65-99.

Goldfeld, S. M., H. H. Kelejian and R.E. Quandt (1971) "Least squares and maximum likelihood estimation of switching regressions" Econometrics Research Program, Research Memorandum No. 130, Princeton University.

Goldfeld, S.M. and R.E. Quandt (1973a), "The estimation of structural shifts by switching regressions", Annals of Economic and Social Measurement, Vol. 2, No. 4, 475-485.

Goldfeld, S.M. and R.E. Quandt (1973b), "A Markov model for switching regressions", Journal of Econometrics, Vol. 1, 3-16.

Gouriéroux, C. and A. Monfort (1996), "Simulation-Based Econometric Methods", Oxford University Press.

Gourieroux, C., A. Monfort and A. Trognon (1984) "Pseudo Maximum Likelihood Methods: Theory", Econometrica, Vol. 52, No. 3, 681-700.

Granger , C. W. J. (1998) "Overview of nonlinear time series specifications in economics", first draft, University of California.

Granger, C. W. J. and T. Terasvirta (1992) "Experiments in modelling nonlinear relationships between time series", in M. Casdagli and S Eubank (eds.), Nonlinear Modelling and Forecasting, Proceedings of a workshop on nonlinear modelling and forecasting held September, 1990, in Santa Fe, New Mexico, (Redwood City, Calif., Addison-Wesley), 189-197.

Granger, C. W. J. and T. Terasvirta (1993) Modelling Nonlinear Economic Relationships, Oxford University Press.

Granger, C. W. J., T. Terasvirta and H. Anderson (1993), "Modelling non-linearity over the business cycle", in J. H> Stock and M. W. Watson (eds) New Research in Business Cycles, Indicators and Forecasting, Chicago University Press, 311-325.

Gregory, A. W. and M. R. Veall (1985), "Formulating Wald tests of nonlinear restrictions", Econometrica, Vol. 53, 1465-1468.

Gregory, A. W. and M. R. Veall (1987), "Formulating Wald tests of the restrictions implied by the rational expectations hypothesis", Journal of Applied Econometrics, Vol. 2, 61-68.

Grossman, G. M. and E. Helpman (1991a), Innovation and Growth in the Global Economy, Cambridge MA, MIT Press.

Grossman, G. M. and E. Helpman (1991b), "Trade, knowledge spillovers, and growth", European Economic Review, Vol.35, 517-526.

Hall, R. E. and C. I. Jones (1997), "Levels of economic activity across countries", American Economic Review, Vol. 87, Issue 2, 173-177.

Hall, R. E. and C. I. Jones (1999), "Why do some countries produce so much more output per worker than others?", The Quarterly Journal of Economics, Vol. 114,83-116.

Hamilton, J. D. (1989) "A new approach to the economic analysis of nonstationaty time series and the business cycle", Econometrica, Vol. 57, No. 2, 357-383.

Hamilton, J. D. (1990) "Analysis of time series subject to changes in regime", Journal of Econometrics, Vol. 45, 39-70.

Hamilton, J. D. (1994) Time Series Analysis, chapter 22, Princeton University Press.

Hamilton, J. D. (1996) "Specification testing in Markov-switching time-series models" Journal of Econometrics, Vol. 70, 127-157.

Hansen, B. E. (1992) "The likelihood ratio test under nonstandard conditions: testing the Markov switching model of GNP", Journal of Applied Econometrics, Vol. 7, S61-S82.

Hansen, B. E. (1994) "Erratum: the likelihood ratio test under nonstandard conditions: testing the Markov switching model of GNP", Journal of Applied Econometrics, Vol. 11, 195-198.

Harrod, R. (1948), Towards a dynamic economics, London and New York: Macmillan.

Hartley, M. J. (1977) "On the estimation of a general switching regression model via maximum likelihood methods" Department of Economics Discussion Paper No. 415, State University of New York at Buffalo.

Heston, A. (1994) "National Accounts. A brief review of some problems in using national accounts data in level of output comparisons and growth studies" Journal of Development Economics, Vol. 44, 29-52.

Heston, A. and R. Summers (1991), "The Penn World Table (mark 5): an expanded set of international comparisons, 1950-1988", The Quarterly Journal of Economics, Vol. 106, No. 2, 327-368.

Heston, A. and R. Summers (1996), "International Price and Quantity comparisons: Potential and Pitfalls", American Economic Review, Vol. 86, 20-24.

Holland, A. and A. Scott (1998), "The determinants of UK business cycles", The Economic Journal, Vol. 108, 1067-1092.

Im, K. S., M. H. Pesaran and Y. Shin (1995), "Testing for unit roots in Dynamic Heterogeneous Panels", DAE Working Papers Amalgamated Series No. 9526.

Islam, N. (1995), "Growth Empirics: A Panel Data Approach", Quarterly Journal of Economics, Vol. 110, No. 4, 1127-1170.

Johansen, S. (1988) "Statistical Analysis of cointegration vectors", Journal of Economic Dynamics and Control, Vol. 12, 231-254

Jones, C. I. (1997) "Convergence Revisited" Journal of Economic Growth, Vol. 2, No. 2, 131-153.

Kiefer, N. M. (1978) "Discrete parameter variation: efficient estimation of a switching regression model", Econometrica, Vol. 46, No. 2, 427-434.

Kim C. J. (1994), "Dynamic linear models with markov-switching", Journal of Econometrics, Vol. 60, 1-22.

King, R. G. and S. T. Rebelo, (1993), "Transitional dynamics and economic growth in the neoclassical model", American Economic Review, Vol. 83, Issue 4, 908-931.

Klenow, P. J. and A. Rodriguez-Clare (1997a) "The Neoclassical revival in growth economics: Has it gone too far?", NBER Macroeconomics Annual, 73-102.

Klenow, P. J. and A. Rodriguez-Clare (1997b) "Economic growth: A review essay" Journal of Monetary Economics, Vol. 40 597-617.

Kloeden, P. E. and E. Platen, (1992), (1995 corrected printing), "Numerical Solution of Stochastic Differential Equations", Springer.

Kocherlakota, N. R. and K. M. Yi (1996) "A simple time series test of endogenous vs. exogenous growth models: an application to the United States" The review of Economics and Statistics, Vol. 78 No. 1, 126-134.

Koop, G., M.H. Pesaran and S.M. Potter (1996), "Impulse response analysis in nonlinear multivariate models", Journal of econometrics, Vol. 74, 119-147.

Koop, G., and S.M. Potter (2000), "Nonlinearity, structural breaks or outliers in economic time series?" In William Barnet *et al* (ed.) Nonlinear econometric modelling, Cambridge, Cambridge University Press.

Koop, G., and S.M. Potter (2001), "Are apparent findings of nonlinearity due to structural instability in economic time series?" Econometrics Journal, Vol. 4, 35-55.

Koopmans, T. C. (1965), "On the concept of optimal economic growth", in The Econometric Approach to Development Planning, Amsterdam, North Holland, 1965.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin (1992), "Testing the null hypothesis of stationarity against the alternative of a unit root", Journal of Econometrics, Vol 54, 159-178.

Lee K. (1997) "Cross-country Interdependencies in Growth Dynamics: A Model of Output Growth in the G7 Economies, 1960-94", (mimeo) University of Leicester, November.

Lee, K, M. H. Pesaran and R. Smith (1995) "Growth and Convergence: A multi-country empirical Analysis of the Solow Growth model",Working Paper, University of Cambridge.

Lee, K, M. H. Pesaran and R. Smith (1997) "Growth and convergence in a multi-country empirical stochastic Solow model", Journal of Applied Econometrics, Vol.12, No. 4, 357-392.

Lee, K, M. H. Pesaran and R. Smith (1998) "Growth empirics: a panel data approach - a comment", Quarterly Journal of Economics, Vol. 113, 319-323.

Lee, M., R. Longmire, L. Matyas and M. Harris (1998) "Growth convergence: some panel data evidence" Applied Economics, Vol. 30, 907-912.

Leung, C. K. Y. and D. T. Quah (1996) "Convergence, endogenous growth and productivity disturbances" Journal of Monetary Economics, Vol. 38 535-547.

Levine, R. and D. Renelt (1992) "A sensitivity analysis of cross-country growth regressions" American Economic Review, Vol. 82, Issue 4, 942-963.

Lucas, R. E. (1988), "On the mechanics of development planning", Journal of Monetary Economics, Vol. 22 No. 1, 3-42.

Lucas, R. E. (1993), "Making a miracle", Econometrica, Vol. 61, Issue 2, 251-272.

Maddala, G. S. (1983) Limited-dependent and qualitative variables in econometrics, Cambridge University Press.

Maddala, G. S. and I. M. Kim, (1998), "Unit Roots, Cointegration, and Structural Change", Cambridge University Press.

Maddala, G. S. and E. D. Nelson (1975) "Switching regression models with exogenous and endogenous switching", Proceedings from the American Statistical Association, 423-426.

Maddala, g. S. and S. Wu, (2000), "Cross-country growth regressions: problems of heterogeneity, stability and interpretation", Applied Economics, Vol. 32, p635-642.

Maddison, A. (1982), Phases of capitalist development, Oxford University Press, Oxford.

Maddison, A (1987), "Growth and slowdown in advanced capitalist economies: techniques of quantitative assessment", Journal of Economic Literature, Vol. 25, 649-698.

Maddison, A (1991), Dynamic forces in capitalist development, Oxford University Press, Oxford.

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

Maudos, J., J. M. Pastor, and L. Serrano (2000), "Convergence in OECD countries: technical change, efficiency and productivity", Applied Economics, Vol. 32, 757-765.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller (1953) "Equation of state calculations by fast computing machines" Journal of Chemistry and Physics, Vol. 21, 1087-1090.

Miller, S. M. (1996) "A note on cross-country growth regressions", Applied Economics, Vol. 28, 1019-1026.

Mills, T.C. (1995a), "Business cycle asymmetries and non-linearities in UK macroeconomic time series", Ricerche Economiche, Vol. 49, 97-124.

Mills, T.C. (1995b), "Are there asymmetries or non-linearities in UK output?", Applied Economics, Vol. 27, 1211-1217.

Nerlove, M. (1996), "Growth rate convergence, fact or artifact?", Working Paper, University of Maryland.

Nerlove, M. (1999), "Properties of alternative estimators of dynamic panel models: an empirical analysis of cross-country data for the study of economic growth", in Analysis of panels and limited dependent variable models, C. Hsiao, K. Lahiri, L-F. Lee and M. H. Pesaran (eds), Cambridge University Press, 136-170.

Nerlove, M. (2000), "Growth rate convergence, fact or artifact?. An essay on panel data econometrics", in Panel data econometrics: future directions, J. Krishnakumar and E. Ronchetti (eds), North-Holland, 3-34.

Nielsen, Lars Tyge (1999) "Pricing and Hedging of Derivative Securities", Oxford University Press.

Obstfeld, M. and K. Rogoff (1996), Foundations of International Macroeconomics, The MIT Press, Cambridge.

Øksendal, Bernt (1998), "Stochastic Differential Equations: An Introduction with Applications", 5th edition, Springer.

Parente, S. L. and E. C. Prescott (1994), "Barriers to technology adoption and development", Journal of Political Economy, Vol. 102, Issue 2, 298-321.

Peel, D. A. and A. E. H. Speight (1998), "Threshold nonlinearities in output: some international evidence", Applied Economics, Vol. 30, 323-333.

Peña, D. (1994) "Discussion. Second-generation Time-series Models: A Comment on

'Some Advances in Nonlinearlinear and Adaptive Modelling in Time-series'", Journal of Forecasting, Vol. 12, 133-140.

Pesaran, M.H. and S. M. Potter (1997), "A floor and ceiling model of US output", Journal of Economic Dynamics and Control, Vol. 21, 661-695.

Phillips, P. C. B. and J. Y. Park (1988), "On the formulation of Wald tests of nonlinear restrictions", Econometrica, Vol. 56, 1065-1083.

Potter, S. M. (1995), "A nonlinear approach to US GNP", Journal of Applied Econometrics, Vol. 10, 109-125.

Proudman, J. and S. Redding (eds.) (1998), "Openness and growth", Proceedings of the Bank of England academic conference on the relationship between openness and growth in the United Kingdom, September 15th 1997

Quah, D.T. (1993a) "Empirical cross-section dynamics in economic growth", European Economic Review, Vol. 37 (April), 426-434.

Quah, D.T. (1993b) "Galton's fallacy and tests of the convergence hypothesis", Scandinavian Journal of Economics, vol. 95, No. 4 (Dec), 427?443.

Quah, D.T. (1996a) "Twin Peaks: Growth and convergence in models of distribution dynamics", The Economic Journal, Vol. 106 (Jul), 1045-1055.

Quah, D.T. (1996b) "Empirics for Economic Growth and Convergence", European Economic Review, Vol. 40, 1353-1375.

Quah, D.T. (1996c) "Convergence empirics across countries with (some) capital mobility", Journal of Economic Growth, Vol. 1, 95-124.

Quandt R. E. (1972) "A New Approach to Estimating Switching regressions", Journal of the American Statistical Association, Vol 67, No. 338, 306-310.

Quandt R. E. and J. B. Ramsey (1978) "Estimating Mixtures of Normal Distributions and Switching Regressions" and comments by M. J. Hartley, D. W. Hosmer, N. M. Kiefer, D. A. Binder, E. B. Fowlkes, P. Bryant, B. R. Clarke and C. R. Heathcote, N. L. Johnson, Journal of the American Statistical Association Vol. 73, 730-752.

Ramsey, F.(1928), "A mathematical theory of saving", The Economic Journal, Vol. 38 (Dec), 543-559.

Rebelo, S. (1991), "Long run policy analysis and long run growth", Journal of Political Economy, Vol. 99, No. 3, 500-521.

Rodriguez, F. and J. D. Sachs (1999), "Why do resource abundant economies grow more slowly?", Journal of Economic Growth, Vol. 4, issue 3, 277-303.

Romer, P. M. (1986), "Increasing Returns and Long-Run Growth", Journal of Political Economy, Vol. 94, No. 5, 1002-1037.

Romer, P.M. (1987), "Growth based on increasing returns due to specialization", American Economic Review, Vol. 77, No. 2, 56-62.

Romer, P. M. (1990), "Endogenous technological change", Journal of Political Economy, Vol. 98, No. 5, Part II, S71-S102.

Romer, P. M. (1993), "Idea gaps and object gaps in economic development", Journal of Monetary Economics, Vol. 32, 542-573.

Romer, P. M. (1994), "The origins of Endogenous Growth", Journal of Economic Perspectives, Vol. 8, No. 1, 3-22.

Romer, D. (1996), Advanced Macroeconomics, McGraw Hill, New York.

Sachs , J. (1981) "The Current Account and Macroeconomic Adjustment in the 1970's", Brookings Papers on Economic Activity, Vol 1, 201-268.

Sachs, J. D. and A. M. Warner (1997), "Fundamental sources of long run growth" American Economic review, Vol. 87, Issue 2, 184-188.

Sala-i-Martin, X. (1990), "Lecture notes on Economic Growth: Introduction to the literature and neo-classical models", Yale University Center Discussion Paper, No. 621 (Dec).

Sala-i-Martin, X (1994), "Cross-sectional Regressions and the Empirics of Economic Growth", European Economic Review, Vol. 38 (Apr), 739-747.

Sala-i-Martin, X (1996a), "The classical approach to convergence analysis", The Economic Journal, Vol. 106 (Jul), 1019-1036.

Sala-i-Martin, X (1996b), "Regional cohesion: Evidence and theories of regional growth and convergence", European Economic Review, Vol. 40, 1325-1352.

Sala-i-Martin, X (1996c), "I just ran four million regressions" Mimeo, Columbia University, December.

Sala-i-Martin, X (1997), "I just ran two million regressions", American Economic Review, Vol. 87,Issue 2, 178-183.

Samuelson, P. A. (1958), "An exact consumption-loan model of interest with or without the social contrivance of money", Journal of Political Economy, Vol. 66, 467-482.

Sawyer, K. R.,(1993), "Continuous Time Financial Models: Statistical Applications of Stochastic Processes" in Handbook of Statistics 11, Econometrics, Maddala, G. S., C. R. Rao and H. D. Vinod, North-Holland.

Sichel D. E. (1989) "Are Business Cycles Asymmetric? A Correction", Journal of Political Economy, Vol. 97, No. 5, 1255-1260.

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

Solow, R. M. (2000), Growth theory: an exposition, (2nd edition), Oxford University Press, Oxford.

Summers, R. and A. Heston (1991), "The Penn World Table (Mark 5): An Expanded Set on International Comparisons, 1950?1988", Quarterly Journal of Economics, Vol. 106, No. 2, 327-368

Taylor, A. M. (1999), "Sources of convergence in the late nineteenth century", European Economic Review, Vol. 43, 1621-1645.

Temple, J. R. W. (1998), "Robustnesstests of the augmented Solow model", Journal of

Applied Econometrics, Vol. 13, 361-375.

Terasvirta, T. (1994), "Specification, estimation and evaluation of Smooth Transition Autoregressive models", Journal of the American Statistical Association, Vol. 84, 208-218.

Terasvirta, T. and H. M. Anderson (1992) "Characterizing nonlinearities in business cycles using smooth transition autoregressive models", Journal of Applied Econometrics, Vol. 7, S119-S136.

Tiao, G. C. and R. S. Tsay (1994) "Some Advances in Non.linear and Adaptive Modelling in Time-series", Journal of Forecasting, Vol. 12, 109-131.

Tong, H. (1993), Non-linear Time Series: A Dynamical System Approach, Oxford University Press.

Tsionas, E. G. (2000), "Real convergence in Europe - How robust are econometric inferences?"

Van de Klundert, T. and S. Smulders (1998) "Capital mobility and catching up in a two-country, two-sector model of endogenous growth" CentER Discussion Paper, Macroeconomics, No. 9813.

Van Norden S. and R. Vigfusson (1996) "Regime-Switching models: A Guide to the Bank of Canada Gauss Procedures", Working Paper 96-3, Bank of Canada.

White, H. (1987) "Specification Testing in dynamic Models" in Advances in Econometrics, Fifth World Congress, Vol. II, Truman F. Bewkey ed., Cambridge, Cambridge Press.

White, H. (1994) "Estimation, Inference and Specification Analysis", Econometric Society Monographs, Cambridge University Press.

Wolff, E. (1991) "Capital formation and Productivity convergence over the long term", The American Economic Review, Vol. 81, No. 3, 565-579.

Wolff, E. (1996), "The Productivity Slowdown: The Culprit at Last? Follow-Up on Hulten and Wolff", The American Economic Review, Vol. 85, No. 5, 1239-1252.