LONG RUN DETERMINANTS OF ECONOMIC GROWTH IN A CROSS SECTION OF COUNTRIES: THEORY AND EVIDENCE

by

Ali Enis Uysal

Ph.D. Dissertation in Economics

London School of Economics and Political Science

To my father

Acknowledgements

This research has been facilitated by the attention and suggestions of my supervisor Christopher Pissarides. I am also indebted to Daron Acemoglu for his attention, encouragement, and comments. I wish to thank Robert Barro, Peter Boone, William Buiter, Peter Clark, Ed Green, Peter Orszag, Edward Prescott, and Hilary Steedman for their comments. I have greatly benefited from the hospitality and the resources of the Centre for Economic Performance at LSE where I was employed for three years. I wish to thank Richard Layard for this opportunity. Parts of this thesis have been presented to the Money and Macroeconomic Workshop, the Seminar for Research Students in Economics, and the International Economic Performance Seminars at LSE, and Money- Macro Seminar at the Cambridge University. I wish to thank participants at those seminars for their comments.

Abstract

This thesis studies long run economic growth in a cross section of countries. Its main objective is to constitute the necessary empirical and theoretical means for explaining the disparities in growth rates across countries. It consists of four non-coherent chapters. The first chapter is an empirical study of post-war economic growth in a wide range of countries. It uses the data provided by Summers and Heston (Penn World Table) and examines the empirical determinants of growth by using advanced panel-data techniques. Chapter 2 is a theoretical model of technology acquisition in a world where innovation is a costly process. It stresses the importance of innovative activity on long run economic growth, and shows how countries may develop at different rates even when they share a common technological frontier. Chapter 3 is another empirical work where the attention is focused on the economic performance of six European countries during 20th Century. We find that World-War-2 has been a major influence in economic activity and left permanent effects on relative incomes. The last chapter of the thesis contains a theoretical econometrics work. It provides consistent criteria for simultaneous selection of autoregressive order with cointegrating rank. A Monte Carlo experiment stimulates the performance of these criteria in small samples.

Table of Contents

List of Figures

List of Tables

Preface

Recently there has been a surge of interest in economic growth both from theoretical and empirical perspectives. Starting with Romer (1986) and Lucas (1988), many macroeconomists have started formulating ideas for growth generating mechanisms, something that was left out at an intuitive level or completely hand waved in previous literature. One of the reasons for this interest in endogenous growth theory was the enrichment of available data sets that made cross country comparisons possible. Most notably, the Penn World Table, a panel data set that is put together by Summers and Heston and contains PPP corrected post-war aggregate data for a wide range of countries, became available at the beginning of 80s. It is followed with an enormous number of empirical research that complemented the new developments in theoretical growth literature.

This thesis is also inspired by these new exiting developments in the growth literature. Its main objective is to contribute to our understanding of cross country differentials in per capita income growth rates. It consists of four non-coherent but complementary chapters. Among them Chapter 1 and Chapter 3 contain some empirical work. Chapter 1 is an analysis of post-war economic growth in a wide range of countries. We find strong evidence on that long run growth systematically varies with country specific factors, and conclude a long term divergence of incomes across the countries. This supports the newly developed ideas in the theoretical literature where economic growth is endogenously generated and cannot be taken as granted. Chapter 3 is a time series analysis of incomes across six European countries during the 20th Century. The main conclusion of this chapter is the long run stability of incomes among these countries. This supports the view that economic growth is a natural consequence of a world wide technological progress. All countries, perhaps with some time lag, adopt the new progress into the production, and maintain its relative level in the long run (which is of course conditional on relative factor endowments being unchanged). This seems somewhat in contrast to the conclusions of Chapter 1 where long run growth was found to vary with country specific elements that remained unchanged in time.

One theoretical explanation to the paradox is given in Chapter 2. It contains a theoretical growth model that can explain the empirical findings of both chapters along with the previous empirical research. In this model, countries are divided into clusters that are determined with factor endowments. Economic growth is a result of a world wide technological progress, and it varies across clusters. However, countries that are in the same cluster grow at the same rate. Hence, one would expect economic growth to vary systematically with factor endowments in a wide range of countries that are belong to different clusters. Moreover, we should observe a stable long run relationship between the member countries of the same cluster. Thus, the findings of both empirical chapters are compatible with the theoretical model provided in this section.

Besides their empirical findings, Chapter 1 and Chapter 3 also contribute to the literature with their original econometrics approaches. Both chapters include some theoretical econometrics discussions and some related theorems (mainly included in their Technical Appendices). This econometrics side of the thesis is complemented with a theoretical econometrics work in the last chapter. We propose consistent criteria for the simultaneous selection of autoregressive order with cointegrating rank. Such criteria are useful to avoid asymptotic Type-I errors if the main purpose of analysis is not inference.

The econometric calculations, simulations, and the data analysis are carried out with GAUSS programming language on a PC environment or on a Sun/UNIX terminal. The excellent IT services at the Centre for Economic Performance, where I was employed through out my Ph.D., have greatly facilitated these empirical work. I like to take this opportunity to thank Richard Layard and Adam Lubanski for their hospitality and helpfulness. The computer programs that are used in this thesis are not included along with it, but soft copies can be obtained from my email address upon request (uysal@lse.ac.uk).

CHAPTER 1

Endogeneity in Long Run Growth Performance

1.1 Introduction

Between 1960 and 1989, the per capita income in a typical country has grown by a factor of two[1]. Cross sectional mean and standard deviation of the average growth rates have been 1.88% and 1.79% respectively. If countries were ranked with respect to their growth performance, then in the top 10% of the table per capita incomes have more than tripled during the 29-year period. Yet, in the bottom 10%, incomes were less in 1989 than they were in 1960. The best performing country in our sample of 107 countries has been South Korea with an average growth rate of 6.66% (this growth rate is consistent with doubling income every 10.5 year, and the per-capita income in that country has multiplied by a factor of 6.74 during the period). By contrast, the poorest performance has been in Chad with a negative growth rate of 2.1%, where the per-capita income has decreased to 54% of its 1960 level. Furthermore, these shocking disparities have also been strongly persistent in some cases. Japan, for example, grew faster than USA for 22 of the 29 years.

There have been two alternative proposals to explain this heterogeneity in growth rates: According to the traditional view, long term growth is an exogenous[2] process, and should therefore occur at a

---

[1] The data is taken from PWT (mark 5.6) which contains 152 countries all together and 115 countries with comparable income data from 1960 to 1989. In this paper, the number of countries is further reduced to 107, since some other variables are not available for the full set. See also Summers and Heston (1991).

[2] *Exogenous*, in this paper, is not interpreted as *manna from heaven*, but as a phenomenon that evolves outside the model. Just as competitive prices are exogenous to a firm but endogenously determined at the market level, economic growth may be exogenous for a given country, but endogenously determined for the whole world. This is more likely to be the

homogeneous rate around the world (see Solow (1956)). Heterogeneity is a result of a process of temporary deviations from and converging back to the steady state. On the contrary, in the newer approach, economic growth is endogenously determined within each country. According to this paradigm, economies can exhibit different growth rates indefinitely because of the differences in preferences, micro structure, government policy, international transactions, size (population and/or human capital), and similar country specific factors (see Romer (1986), Lucas (1988), Young (1991) among others).

This paper is an attempt to discuss and quantify the relative role of endogeneity in long run economic performance. "Long run", here, refers to "permanent" disparities in growth performance and not *long lasting temporary deviations* due to some changes in factor endowment. This is contrary to the focus of most recent empirical work that *a priori* treat the cross sectional disparities as long lasting temporary deviations. Starting with the influential work by Kormendi and Meguire (1985)[3], many authors studied the impact of changes in several socio-economic variables in a "cross country growth regression" framework. According to this setting, the growth rates in per capita incomes that are averaged over a reasonably long sample period (30 years in most cases) are regressed on several explanatory variables of interest that also includes the initial values of log per capita incomes. In almost all cases, a negative coefficient for the initial level of incomes is obtained and interpreted as evidence of a *conditional convergence* across the countries[4]. What is more important is that other coefficients are interpreted

---

case if there is a high degree of technological diffusion across the countries, and all countries are small compare to the world.

[3] The earlier studies of similar type include Feder (1982), and Landau(1983). There has been a rapid increase in the literature during 90's. See Levine and Renelt (1992), Fischer (1993), Barro and Lee (1993), King and Levine (1993), Barro (1994), and Sachs and Warner (1995).

[4] There have been other approaches to test if there is a convergence across the countries in terms of per capita income levels. See Bernard and Durlauf (1991), and Quah (1993b).

as the effects of the corresponding variables on *long run* economic growth (owing to the relatively long length of the sample period). This paper differs from this framework in two fundamental ways. First, the primary objective is to estimate the effect, if any, of the *level of an explanatory variable* on the long run growth, but not the effects due to a change in its level (notice that previous literature is focused on the latter by completely ignoring the possibility of the former). Such effects, if found in the data, would be consistent with the newly developed endogenous growth literature and the ideas therein, and suggest that countries can exhibit differentials in their long run growth performance as a result of their factor endowments. Second, we exclude one of the central explanatory variables of previous empirical work based on the theoretical analysis; namely, the initial values of the per capita incomes. It is shown in Technical Appendix-A that similar implications for the cross-country growth regressions can be generated with alternative approaches to economic growth if the initial values of incomes are included as an explanatory variable. Therefore coefficients lose their interpretability unless a priori restrictions on the true growth generating mechanism are enforced.

To make the empirical analysis more transparent, we first derive the cross country growth regression implications for alternative theories of economic growth. The similarity between the predictions of endogenous and exogenous growth models about the regressions that include initial levels of incomes is also shown in this section. The predictions, on the other hand, differ for the regressions that do not include initial levels of incomes as an explanatory variable. The empirical section concentrates on this identifying implication and studies the data for systematic variations of long run growth rates with country specific factors in a panel-data-random-coefficients set-up. Attention is focused on five explanatory variables; namely, schooling, physical capital investment rate, the share of government expenditures in GDP, openness to international trade, and the population growth rate as they were implied by the recent endogenous growth literature. The results show that the investment rate in physical capital formation is the major determinant of long run growth. The other factors are important to the extent that they influence the physical capital investment rate. Furthermore, as a natural consequence of our econometric framework, *business cycles* are also identified and decomposed into world-wide and idiosyncratic components. We find that countries that are highly endowed with physical and human capital are less affected by world-wide productivity shocks. This is interpreted as evidence to the idea that bigger economies can allocate risk into a more diverse variety of industries and, as a result, are more stable (see Acemoglu and Zilibotti (1995)). In addition, the data suggest that there is an inverse relationship between the impact of world-wide productivity shocks on the economy and the share of government expenditures. This could indicate some rigidities in the government sector and that a higher degree of government involvement in economy acts as a buffer against world-wide productivity shocks (see Acemoglu and Zilibotti (1995)). There is a positive relationship between openness to international trade and the extent to which world-wide shocks influence home economy. This is not very surprising since one would expect open economies to be more exposed to productivity shocks elsewhere. In addition, the results also have some implications for the convergence hypothesis literature. Contrary to the main result generally found in cross sectional data analysis (see Barro (1991) and Barro and Salla-i Martin (1992a) among many others), we conclude a slow divergence across the countries in terms of their per capita incomes. In particular, the gap between the poor and the rich nations has been widening in every year since

1960 (except in 1981). Underlying it is the big gap in physical capital investment rates between poor and rich nations.

The second part of the empirical section studies the time series aspects of the innovations to the average growth rates. It is shown that idiosyncratic shocks to incomes are quite large in magnitude (about 5% of the GDP) and much of them are permanent. This is also contrary to the previous findings of the empirical growth literature that suggests the disparities in growth rates are a result of long lasting temporary deviations (i.e., idiosyncratic shocks are temporary, though persistent).

The rest of this chapter is organised as follows; Section 1.2 discusses the dynamics of incomes for alternative theories of economic growth. Section 1.3 includes the econometric considerations and reports the quantitative results. Section 1.4 concludes the chapter.

1.2 Dynamics of Incomes

Exogenous Growth Models

To begin with, it will be convenient to review an augmented Solow growth model[5]. Consider an economy that uses physical capital, K(t), human capital, H(t), and labour, L(t), to produce a single output, Y(t);

$$Y(t) = K(t)^{\alpha} H(t)^{\beta} [A(t)L(t)]^{1-\alpha-\beta} \qquad (2.1)$$

Here, the aggregate technology exhibits constant returns to scale. A(t) represents the state of technology, and is assumed to grow at an exogenous rate "b". Labour grows exogenously at rate "n," and capital inputs are accumulated by saving a constant portion of income (Solow assumption):

[5] The main features of this model are borrowed from Mankiw, Romer, and Weil (1992).

$$\frac{d\,\tilde{k}(t)}{dt} = s_K \cdot \tilde{y}(t) - (n+g+\delta) \cdot \tilde{k}(t) \quad (2.2)$$

$$\frac{d\,\tilde{h}(t)}{dt} = s_H \cdot \tilde{y}(t) - (n+g+\delta) \cdot \tilde{h}(t) \quad (2.3)$$

At the steady state equilibrium of such an economy, per capita incomes grow at the exogenous rate "b". Furthermore, cross-country differences in preferences, micro structure and other similar factors can make a difference only at the level and not at the rate of growth. Formally, for a country starting with initial steady state per capita income $e^{y(0)}$, the logarithm of the per capita income, $y_i(t) = \ln Y_i(t) - \ln L_i(t)$, evolves by

$$y_i(t) = bt + y_i(0) \quad (2.4)$$

where "b" is the same for all countries, and relative steady state levels, $y_i(0)$, may differ in the cross section based on the country specific factors (i.e., n, s_K, and s_H in this model).

Endogenous Growth Models:

In endogenous growth models, countries can undergo sustained economic growth by accumulating reproducible factors only. Unlike the exogenous growth models, economies do not converge to a steady state even in the absence of an external technological progress. One common and easy way of incorporating this view into the model studied above is to assume a per capita technology that is constant returns to scale:

$$y(t) = A \cdot k(t)^{\alpha} \cdot h(t)^{1-\alpha} \quad (2.5)$$

Equation (2.5) is virtually identical to (2.1) except that β=1-α and the technology parameter "A" is assumed to be time invariant. Just as in the exogenous growth models, the capital inputs (h and k) are assumed to be accumulated in the same way by saving a constant portion of the current income;

$$\frac{d\,k(t)}{dt} = s_K \cdot y(t) - (n+\delta) \cdot k(t) \qquad (2.6)$$

$$\frac{d\,h(t)}{dt} = s_H \cdot y(t) - (n+\delta) \cdot h(t) \qquad (2.7)$$

If these savings rates are large enough, this economy grows at the endogenous rate "b" given by

$$b = A s_K^{\alpha} \cdot s_H^{1-\alpha} - \delta - n \qquad (2.8)$$

Similar to the exogenous growth models, in the absence of stochastic disturbances, per capita incomes in this economy would be given by

$$y_i(t) = b_i t + y_i(0) \qquad (2.9)$$

where "i" is the cross sectional index. The difference here is that cross-country differences in factor endowments can make a difference not only at the relative level of per capita incomes, but also at its rate of growth. In particular, the growth rate of incomes in this over simplified model is a function of the savings rates and the population growth rates (i.e., n, s_K, and s_H).

So far, the growth rates are assumed to be constant in time for both exogenous and endogenous growth models. Yet, one may wish to relax this assumption by allowing a stochastic rate of growth instead. Considering also that economies may be subject to various temporary shocks, η, the observed series of log-per-capita incomes, $\hat{y}$, should be specified as a unit root process:

$$\hat{y}_i(t) = \hat{y}_i(0) + b_i \cdot t + \sum_{\tau=1}^{t} \varepsilon_i(\tau) + \eta_i(t) \qquad (2.10)$$

$$E(\eta_i(t)) = 0 \qquad E(\varepsilon_i(t))=0$$

where $\varepsilon_i(t) = b_i(t) - b_i$[6]. This implied non-stationarity of incomes, and how to separate the permanent shocks, ε, from those which are temporary, η, have recently been studied by many econometricians. (see Beveridge and Nelson

[6] Notice that in the presence of stochastic shocks the solutions to the models are not correct. Nevertheless, on the basis of these models one would expect (2.10) as an adequate representation of real life data when sufficient dynamics is built on the disturbances.

(1981), Campbell and Mankiw (1987), Cochrane (1988), Blanchard and Quah (1989), and Quah (1992) among others). Notice that the formulation in (2.10) is quite general and highlights two significant differences of exogenous growth models from their endogenous analogues. Firstly, in exogenous growth models the long run average growth rates are identical across all countries (i.e., b_i=b for all i), and country specific factors can make a difference only on the level of relative incomes. Secondly, stochastic component of incomes between any two economies can deviate only temporarily. This is because permanent shocks in exogenous growth models are as a result of a world wide technological progress. Therefore, permanent components should be driven by the same unit root process in all countries (i.e., $\sum_{\tau=1}^{t}\varepsilon_i(t)-\sum_{\tau=1}^{t}\varepsilon_j(t)$ should be a stationary process). These two important predictions can be put together to claim what is known as the "convergence hypothesis" of exogenous growth models. The issue is recently studied by many authors including Barro and Salla-i Martin (1992), Quah (1993a), Bernard and Durlauf (1995), and Uysal (1996). Although, this paper is not directly concerned with testing the convergence hypothesis, our empirical results have also some ramifications on this current debate.

1.3 Quantitative Analysis

The objective of this section is to quantify the theoretical discussion of Section 1.2, and estimate the relative contribution of country specific factors to long run economic performance. In other words, we study the cross sectional variation of per capita growth rates with several country specific factors. As a direct consequence of our econometric approach, we also decompose the growth rates into idiosyncratic and world-wide components, and study their time series aspects.

First of all, however, it is appropriate to present some related preliminary observations. As clarified in Section 1.2, if long run growth were completely exogenously determined, then the disparities in time averages of growth rates should have vanished as the length of the sampling period increased. Figure-1.1 plots the cross sectional mean and the standard deviation of average growth rates from the base year 1960 to the date. The cross sectional dispersion of average growth rates rapidly decreases with the length of the period, but stabilises around 1.64%. This result contradicts to the predictions of exogenous growth models, and is in line with the earlier findings of several other time series analyses of incomes (see Bernard and Durlauf (1991) and/or Quah (1993a)). It is, therefore, worthwhile to study this dispersion in per capita income growth rates for systematic variation with several country specific elements.

1.3.1 Econometric Methodology

In this paper, we adopt the panel data random coefficients approach to study the effects of several explanatory variables on long run growth performance. On the basis of the arguments developed in the previous sections,

$$g_i(t) \quad = \quad x_i\,\beta(t) \,+\, \lambda(t) \,+\, \mu_i \,+\, v_i(t) \qquad (3.1)$$

is a natural formulation to study the year by year changes in the logarithm of per capita incomes, $g_i(t)$. Here, the expected values of the growth rates are determined separately by the country specific factors (x_i) within each country. The effects of the variables on the growth rate, as well as the common intercept (i.e. the time effects, λ), are permitted to vary in time. According to the exogenous growth models, for example, all the entries in $\beta(t)$ is expected to average zero in longer horizons. Individual effects, μ, are included to capture the remaining cross sectional variation in growth rates, and assumed to be independently identically distributed across the countries. Similarly, the idiosyncratic disturbances are also assumed to be independently identically distributed in the cross section with an arbitrary time structure:

$$\mu_i \sim \mathrm{IID}(0,\sigma_\mu^2)$$
$$v_i \sim \mathrm{IID}(0,\Omega)$$

where μ_i is the individual effect, and v_i is the column vector of idiosyncratic shocks to country "i" during the period. Similar panel data set-ups are considered in the previous empirical literature (most notably see Kiefer (1980), Chamberlain (1982) and (1984) for theoretical discussion and MaCurdy (1982) for an empirical application of a special restricted form on earnings).

One can re-formulate this model in terms of a simultaneous equations system (SES) where each equation corresponds to a single period;

$$\begin{bmatrix} g_i(1) \\ g_i(2) \\ \vdots \\ g_i(T) \end{bmatrix} = \begin{bmatrix} \lambda(1) \\ \lambda(2) \\ \vdots \\ \lambda(T) \end{bmatrix} + \begin{bmatrix} b'(1) \\ b'(2) \\ \vdots \\ b'(T) \end{bmatrix} \cdot x_i' + \begin{bmatrix} \mu_i + \varepsilon_i(1) \\ \mu_i + \varepsilon_i(2) \\ \vdots \\ \mu_i + \varepsilon_i(T) \end{bmatrix} \tag{3.2}$$

Equivalently, the observations can be stacked in a matrix form so that each column represents a single cross sectional unit;

$$G = B \cdot x + E \tag{3.3}$$

where G and x are matrixes of observed growth rates and the explanatory variables respectively (a constant is also included in x to account for time effects). E is the matrix of disturbances where each row corresponds to a single period and each column corresponds to a single cross sectional unit. B is the matrix of coefficients in which rows correspond to a particular period and columns to a particular explanatory variable (first column being the vector of time effects).

Without a priori restrictions, the ordinary least squares (OLS) estimate for B is consistent and asymptotically efficient. Furthermore, one can apply the Wald statistics to test the null hypothesis that the effects of explanatory variables are identical from one year to another in the usual way. If the null hypothesis is not rejected, then more efficient and consistent estimates of the restricted form coefficients can be obtained by applying the generalised least squares (GLS) method.

It is possible in this set-up that the explanatory variables are correlated with the changes in incomes in some years but do not effect the average performance in the long run. In other words, some variables could influence the changes in incomes sometimes positively and sometimes negatively in such a way that positive and negative effects cancel out in the long run. Therefore, one might be more interested in the 'mean effect' of an explanatory variable on the growth rate rather than its effect for every year during the sample period. If the null hypothesis that the coefficients are time invariant cannot be rejected, then one can proceed with panel data GLS regressions as described by MaCurdy (1982), or Kiefer (1980). Similarly, the minimum distance (MD) estimator, as described in Chamberlain (1982), can also be used to obtain the consistent and efficient estimates of the mean effects, and it turns out that they are all numerically identical in this case. If the null hypothesis is rejected, however, there are no consistent estimators for the mean effects in the cross sectional dimension. In other words, the consistent estimates are available only as the length of the sampling period goes to infinity. Nevertheless, one can construct unbiased estimators for the mean effects that are also consistent estimators for the averages of the realised effects during the sample period. We produce three such estimators: First is a simple average of the coefficient estimates for B that are obtained with OLS from the SES set-up. It is equivalent to the standard approach of previous empirical literature where average growth rates are regressed on explanatory variables. Second is a weighted average of these coefficients that minimises certain moments of the sample. It is also numerically equivalent to the pseudo-GLS estimate that assumes time invariant coefficients, and to the Chamberlain's MD estimator. Third is a more general GLS that is obtained through a *random coefficients* formulation. That is, it takes the additional cross sectional heterogeneity in the variance-covariance structure of the error terms into account before applying the GLS method. It also makes a difference how one formulates the variance covariance matrix of the innovations to the random coefficients. It is assumed in this paper that the innovations to the coefficients are serially uncorrelated. This is clearly a restrictive assumption, and the efficiency of the estimates can be improved with more elaborate techniques that allow intertemporal correlation (e.g., VAR models). All these three estimators are unbiased and consistent as the time dimension increases (the last one being asymptotically more efficient). We report all three estimates since the short sample properties are not known. The procedures are formally described in Technical Appendix-B.

Another important issue is to study the dynamic properties of innovations to the changes of log per capita incomes. Among those, measures of persistence in Ω, its decomposition into permanent and transitory components, and the remaining cross sectional variation, σ_{μ}^2, are of particular interest. More specifically, σ_{μ}^2 would tell us how much of the between group variation in growth rates is not explained with our set of explanatory variables and therefore the potential benefits from additional explanatory variables. Ω is the variance covariance matrix of the idiosyncratic shocks. Its decomposition into permanent and transitory components, therefore, is important to assess how much, for example, seemingly-a-like- countries can drift apart in the long run. Furthermore, the estimates of persistence in the idiosyncratic shocks can help us to understand the extent to which short run policies can influence the economy in the longer run. These findings can also be useful in locating our results with respect to the related empirical literature. The theoretical concerns related to these issues will be discussed in more detail later in the chapter.

1.3.2 Data

In this paper five variables are considered as the candidate determinants of long run endogenous growth; saving rate in physical capital, the savings rate in human capital, the share of government expenditures in the GDP, openness to international trade (the share of exports in the GDP), and the population growth rate. Among those, the savings rate in human capital and the savings rate in physical capital are building blocks of any endogenous growth theory (see Romer(1986), Lucas (1988) among many others). The effects of the share of government expenditures on the endogenous growth performance are studied by Barro and Salla-i Martin (1992b). They show that there is an optimum share of government expenditures that maximises the growth rate of per capita incomes. The potential effects of openness to international trade on long run performance are also widely discussed in the endogenous growth literature. Recently Lucas (1993) highlights the theoretical and empirical observations on the rapid increase in exports of some fast growing East Asian economies. Stokey (1991) and Young (1991) stress how international trade can influence the endogenous growth rate through learning by doing mechanisms.

With this specification, we assume that these five are the only country specific variables that could directly affect the long run growth rate. The theoretical literature, on the other hand, also includes potential effects of several other variables. It is assumed here that other environmental variables such as political stability, degree of financial development, income distribution, monetary stability, and perhaps others not yet considered in the literature, have no direct impact on long run economic growth, but perhaps via one of these five variables. The influence of the variables that are left out and orthogonal to the explanatory variables of this empirical exercise are summarised within the individual effects (i.e., μ_i).

Since the actual investment rate in human capital accumulation is not observable, we use the same proxy, SCHOOL, as was suggested by Mankiw-Romer-Weil (1992). It is approximately the percentage of the working-age population that is enrolled in secondary school[7]. It is clearly a rough approximation for the actual savings rate in human capital. It ignores the quality inputs (number student per teacher, number of computers and other equipment per student etc.) as well as the other forms of human capital investment; such as primary and higher education, vocational courses, and on the job training. Nevertheless, the variable is shown to capture satisfactorily the relative human capital endowment across the countries in Mankiw Romer and Weil (1992). There is also an ambiguity on choosing an appropriate measure of openness to international trade. In previous literature the share of exports, the share of imports, and total share of international trade are all considered as proxies for openness. Sachs and Warner (1995) uses a set of criteria, and dummy those countries that satisfy these criteria as open. In this paper we use share of exports as our measure of openness to international trade. It should be noted, however, this measure is far from being perfect, and more elaborate measures that take relative sizes of countries into account are needed. The other three explanatory variables and

[7] More precisely, SCHOOL is the product of the fraction of eligible population (aged 12 to 17) enrolled in the secondary school with the fraction of the working-age population that is of school age (aged 15 to 19). The data is obtained from the UNESCO yearbook. See Mankiw Romer and Weil (1992) for more details.

the growth rates are taken from the Penn World Table 5.6[8] (also known as PWT 5.6 or the Summers & Heston data set). Since the time means of these variables are not observable, we use their time averages during the sample period. For data sets that are short in time dimension, this approach can produce unreliable results. But because the number of years covered in PWT 5.6 is relatively large, we hope such biases would not make significant effects on our conclusions (see Data Appendix).

1.3.3 Empirical results

Effects of Factors Endowment:

Table-1.1 includes the summary results for OLS regressions of the simultaneous equations model in (3.3) for alternative specifications. More precisely, six settings are considered. For each of these settings, the table reports the range of coefficients for each explanatory variable, the Wald test results for the null hypothesis that the coefficients are time invariant, and the number of years in which they have been statistically significant. In addition, the full range of coefficients are illustrated in Figure-1.2.

According to these OLS results, the effects of the explanatory variables changed significantly from one year to another. The Wald tests reject the null hypothesis that the effects are time invariant in most cases. The traditional exogenous growth models would interpret these time varying correlation coefficients as evidence for the co-movements of the business cycle disturbances across similarly endowed countries. That is, world-wide productivity shocks are probably one of the several sources of the business cycle fluctuations, and the short-run impacts of these shocks may differ across the countries depending on their factor endowments. One would

[8] The Penn World Table provides cross sectionally comparable yearly post-war data (from 1950 to 1991) for a large set of countries (152 all together). See Summers and Heston (1991) for more detail.

expect, for example, the oil shocks in 1970's had different consequences between industrialised and third-world countries. The same explanations also go through with the endogenous growth models[9]. The two approaches differ only on the accumulation of these asymmetric (endowment based) disturbances in the long run. According to endogenous growth framework, the differences can accumulate to be significant, whereas, in exogenous growth models such endowment based asymmetries have no long run consequences. It is therefore worth checking the cumulative contribution of these endowment based shocks during the 29-year sample period. Tables 1.2 and 1.3 present the shares of endogenous growth for six individual cases; namely US, UK, Japan, *All*, *Rich* and *Poor* (also see Figure-1.3 for the historical decomposition of US growth). Here, *All*, *Rich* and *Poor* are the averages for all countries, for richer countries (that are richer than the average in 1960), and for poorer countries respectively. The shares are calculated as the ratio of square cumulative effect of a variable to the summation of square cumulative effects of all the variables (including the remaining shocks). Formally,

$$\text{share of } x_i \text{ in overall growth} = 100 \cdot \frac{\left(x_i \cdot \sum_t \beta_i(t)\right)^2}{\sum_k \left(x_k \cdot \sum_t \beta_k(t)\right)^2 + \left(\sum_t e(t)\right)^2}.$$

According to Table-1.3, country specific elements account for most of the long run growth. Contrary to the predictions of the exogenous growth models, the time effects are not significant at the end of 29 year period. They

9 This result is more difficult to justify with many endogenous growth models in the literature. This is because most endogenous growth models link the effects of explanatory variables to economic growth via production technology parameters that are probably not as much volatile over time. On the other hand, the endogenous growth models that are in the sprit of Romer (1990) are more flexible with time varying effects of country specific factors. This is because endogenous growth in such models are generated with endogenous technology adaptation. Next chapter provides one such endogenous growth model where the time-varying coefficients are interpreted as *endowment based technological progress.*

account for only about 2% of the overall growth performance. The physical capital investment rate is the most important determinant of long run growth in developed countries. It alone explains about 75% of growth for *Rich* and about 50% for *All*. Also, the marginal contribution of exports and government expenditures are not very big for *Rich*, though they are substantial for *Poor*. From the endogenous growth point of view, the effects of the schooling and population growth rate are probably the most surprising. Neither of the two variables have the predicted strong marginal impact on the growth performance. In addition, the overall contribution of the idiosyncratic shocks is considerably small, but in some cases at important levels. There will be more discussion about the characteristics of these idiosyncratic shocks later in this section.

It is also evident in Table-1.3 that growth experiences have been quite different for *Rich* and *Poor*. Table-1.4 is a year by year comparison of these differences (also see Figures 1.4 and 1.5 for graphical illustration). According to these results, *Rich* has grown faster than *Poor* for 28 of the 29 years. Moreover, the differences have been statistically significant in 13 of the 28 years. A formal test for the joint null hypothesis that there have been no differences in the growth experiences of *Poor* and *Rich* is rejected at all reasonable significance levels. It is, therefore, only fair to conclude that the gap between *Rich* and *Poor* has been widening statistically significantly during the post-war years. This also confirms with the earlier findings by Quah (1993b) who suggests a divergence into clubs.

Another parallel result that relates to the endowment based differences in growth experience is obtained from the cross product matrix of the coefficients (see Table-1.5). After controlling for the remaining three factors, it is evident from this matrix that countries with high schooling and physical capital investment rates experience smaller cycles than the others.

This is because, schooling and investment rates have often opposite sign with the world wide productivity shocks (i.e., time effects), and therefore likely to be bigger during recessions or vice versa. It supports the idea that highly endowed countries are able to diversify the risk into a larger variety of industries, and as a result, they are less effected from the world wide productivity shocks (see Acemoglu and Zilibotti (1995) for a theoretical discussion). The share of government expenditures is also negatively correlated with the time effects. It can be interpreted as a support for the idea that governments can play a buffer role in the economy. It is, however, not clear whether some rigidities in the government sector is the reason for this buffer role, or the effectiveness of short-run policies. The only variable that is positively correlated with the world cycles is openness to international trade. It reflects that open economies are more tied to the others, and therefore more exposed to productivity shocks elsewhere. In brief, these five explanatory variables endogenise the extent to which the world cycles influence the home economy.

One important econometric problem for the above findings arises from the possible endogeneity of the explanatory variables. The direction of causality of the correlation between growth and physical capital investment, for example, is ambiguous in the theoretical growth literature. This is because, one could expect that a high rate of physical investment generally accompanies rapid growth that has been generated as a result of other factors. This would be the case if capital is substantially mobile across countries, and high rate of growth promotes the rate of return on physical capital investment. Then, the share of investment in GDP is also endogenously determined with the growth rate being one of its determinants. To test for the existence of such reverse causality in the data, the same exercise is repeated for the last 19 years of the sample and by using the averages of the initial 10 years as the instrumental variables (IV). The results did not change significantly, and the formal Durbin-Wu-Hausman (DWH) specification tests could not reject the null hypothesis that the estimated coefficients are the same as those estimated by OLS for this 19 year period. The results of the IV estimation are summarised in Table-1.6.

The estimates for the *mean effects* of the explanatory variables are reported in Table-1.7. The first column is a simple time average of the coefficients for the corresponding variable. It is numerically equivalent to the usual OLS regression of average growth rates on the explanatory variables. The second column is the MD estimator that uses the unrestricted simultaneous equations system OLS coefficients to estimate the restricted form mean effects as described by Chamberlain (1982). It is equivalent to the GLS estimation that assumes time invariant coefficients, random individual effects and idiosyncratic shocks with arbitrary time structure. Finally, the third column includes the estimated mean effects under the random coefficients specification (see Technical Appendix-B).

It should be noted first that all these three estimators are unbiased for the mean effects but not consistent in the cross sectional dimension. This is due to our earlier finding that the coefficients vary in time and that the formal tests reject the restriction for time invariant coefficients (see Table-1.1). On the other hand, all three estimators are consistent for a weighted average of the realisation of random coefficients during the 29 year period. In other words, the same parameters estimate two different concepts simultaneously: mean effects, and a weighted average of what happened during the sample period. Consequently, two t-statistics are reported. The first one is the usual t-statistics for the average of the realised coefficients that is asymptotically (i.e., as the size of cross sectional units increases) normal under the null hypothesis. It provides a consistent test statistic against the alternative. The second one is only a quasi t-statistic that is constructed from the ratio of estimated mean effects to their standard deviations. Since the estimators are not consistent, they do not provide a consistent test for the null hypothesis that mean effects are zero. In other words, as the number of cross sectional units goes to infinity, these t-tests still remain bounded even if the null is incorrect. In addition, it is normally distributed only if the time series observations are large or the coefficients are random draws from a multivariate normal distribution. As a result, it requires additional distributional assumptions to convert these t-tests into probability measures under the null hypothesis. Hence, one should be more sceptical on interpreting these quasi t-statistics on hypothesis testing about the mean effects of the explanatory variables.

It is evident from the first column of Table-1.7 that the averages of the realised coefficients for *Investment* and *Exports* have been positive and significantly different from zero. However, there is not enough evidence to support the hypothesis that the coefficients for *School, Government* and *Population Growth* have also had an effect on average growth performance during the period. This result is consistent with endogenous growth models with high degree of capital mobility. In these models, the factor endowments provoke income growth that is also accompanied by high investment rate in physical capital. This is enforced by capital mobility to equalise the real interest rate across the countries. In return, other factors do not have any marginal contribution to economic growth over and beyond that is explained by physical capital investment rate. The significant marginal contribution of exports on growth could be due to the ideas that are developed by Grossman and Helpman. It could indicate that openness to international trade stimulates innovative activity through increased competition.

If the parameters in Table-1.7 are interpreted as estimators for mean effects (i.e., what usually happens as oppose to what has happened), then the third column provides the efficient estimates. Again, we find only exports and physical capital investment rates as significant determinants of long run growth with quite high quasi t-ratios. Given that the sample period is quite long, the probability of obtaining these high quasi t-ratios is very low for reasonable distributional assumptions of the random coefficients. It is not clear from the table, to what extent the other three variables influence the long run growth. There is not enough evidence to reject the null hypothesis that the mean effects for these variables are zero. On the other hand, because the quasi t-tests are not consistent against the alternative, a low value does not necessarily favour the null hypothesis against the alternative. As a result, there is not any conclusive evidence in the data to measure the mean effects of schooling, government expenditures and population growth rates on long run growth performance.

Idiosyncratic Shocks:

The analysis of macro economic disturbances is another important concern of the recent empirical literature. Table-1.8 reports the covariogram associated with the residuals of the simultaneous equations system in (3.1). It includes the mean, maximum, minimum and a χ^2-test for the validity of the restriction that these autocovariances are time invariant (the mean autocovariances can be estimated by the optimal minimum distance procedure as described by Chamberlain (1982)[10]). The formal tests reject the null hypothesis that the autocovariances of the same order has been the same during the sample period for most cases. This indicates a non-stationarity of the idiosyncratic shocks in the sense that the second moments change through time[11].

---

[10] Because the time series dimension is large relative to the cross sectional observations, a non-singular covariance matrix for the covariance matrix of the residuals cannot be estimated. We use the Moore-Penrose inverse of the estimated covariance matrix to perform the Optimum Minimum Distance procedure.

[11] Significant fluctuations in autocovariance structure of random variables are referred as non-stationarity in the previous empirical work (see MaCurdy (1982)). The reader, however,

Regardless of whether the stationarity assumption fails or not, one might be interested in various decomposition of the idiosyncratic shocks. In this paper we estimate 8 different structural decomposition. In each case, the parameters of interest are estimated both using MD and PML procedures and the results are reported in Table-1.9. In all cases, the idiosyncratic shocks are assumed to consist of three main components; namely permanent disturbances, temporary disturbances, and random individual effects. In addition, both temporary and permanent components are modelled as either as white noise or first order autoregressive processes (the temporary component is named as 'measurement errors' if white noise is assumed). These are clearly very special structural forms for more general reduced form ARMA processes. In fact, one can model the idiosyncratic shocks as an ARMA process of arbitrary order and decompose it into its three components with arbitrary relative magnitudes (see Quah (1992)). Nevertheless, we hope that the restrictive structures are close approximations to the reality, and hence, informative on the magnitudes of the relative components and their measure of persistence.

According to Table-1.9, the idiosyncratic shocks are mainly composed of permanent disturbances and individual random effects. The temporary component disappears whenever it is modelled as an autoregressive process (i.e., the estimated autocorrelation coefficient is equal to one), or the relative size is small when restricted to be white noise. In fact, the minimum distance procedure always results with no temporary component except when we restrict the idiosyncratic shocks to white noise measurement errors and random individual effects. This is consistent with our earlier findings and supports the endogenous growth models as oppose to the traditional

should notice that the non-stationarity here does not imply an ever-increasing variance of the idiosyncratic shocks as it would for unit root processes. In fact, the unconditional time series expectation of the autocovariances for idiosyncratic shocks are constant. It is the realisation of these autocovariances what varied during the sample period.

exogenous growth theory. Moreover, the data rejects the convergence hypothesis once again since countries cannot converge if they are subject to idiosyncratic disturbances that have permanent effects.

The variance of the random individual effects can be used to obtain a quasi-R^2 as a measure of performance of the cross country growth regressions. Formally, we use the ratio of square estimated mean growth rate to its summation with the variance of random individual effects.:

$$\text{quasi}-R^2 = \frac{(x \cdot \beta)^2}{(x \cdot \beta)^2 + \sigma^2}$$

The results are reported in Table-1.10. Accordingly, the five explanatory variables explain most of the disparities in average growth rates across the countries. Introduction of additional explanatory variables may change the estimates for contribution of the five explanatory variables to economic growth, but cannot significantly improve the overall performance. This also indicates that an important part of the omitted variables, if any, influence long run growth through one of these five explanatory variables. It is also consistent with the endogenous growth models where country specific factors affect the long run growth also by increasing the physical capital investment rate at the same time.

1.4 Concluding Remarks

This chapter is an attempt to quantify the relative importance of country specific factors on long run economic performance. First, alternative growth models are studied to highlight the identifying implications. We argue that endogenous and exogenous growth models differ in one main implication: In endogenous growth models endowment based shocks can average significantly different than zero even in relatively long horizons.

The main empirical results of the chapter can be summarised as follows: First, we show that yearly growth performances are correlated with country specific factors, and that those correlation coefficients change from one year to another. This is consistent with both exogenous and endogenous growth models and indicates a co-movement of business cycle disturbances across similarly endowed countries. Moreover, the evidence indicates that countries with bigger factor endowments are less affected by the world wide productivity shocks. This is consistent with the idea that such countries are in average better in diversifying the risk into a greater variety of industries. In addition, a smaller government sector and a higher degree of openness to international trade indicate that world wide productivity shocks have a bigger impact on such countries. This could be because open economies are closely tied to the rest of the world and that governments follow counter cyclical fiscal policies.

Then, the mean effects of country specific factors are estimated. We find evidence to support the hypothesis that mean growth rates differ systematically across countries with the five explanatory variables of this empirical exercise; schooling rate as a proxy for human capital accumulation, physical capital investment rate, share of government consumption in the GDP, openness to international trade, and population growth rate. Among them, both physical capital investment rate and openness to international trade are statistically significant, though, only physical capital investment rate is at a quantitatively important level. This is consistent with the endogenous growth models where the country specific factors effect the physical capital investment rate and long run per capita growth simultaneously and proportionally.

Next, the idiosyncratic disturbances are studied. The residuals of the growth regression are decomposed into three main components with eight alternative structural specifications: Temporary shocks, Permanent shocks, and random individual effects. We find that the idiosyncratic disturbances consist of mainly large permanent innovations, and that a typical shock ranges from 4 to 6 percent of the GDP. This indicates that short run policies in economic performance can be quite important even for quite long horizons. The size of random individual effects is small, however, indicating that the five explanatory variables are capturing most of the cross sectional variation of average growth rates.

Figure - 1.1

The Mean and Standard Deviation of Average Growth Rates

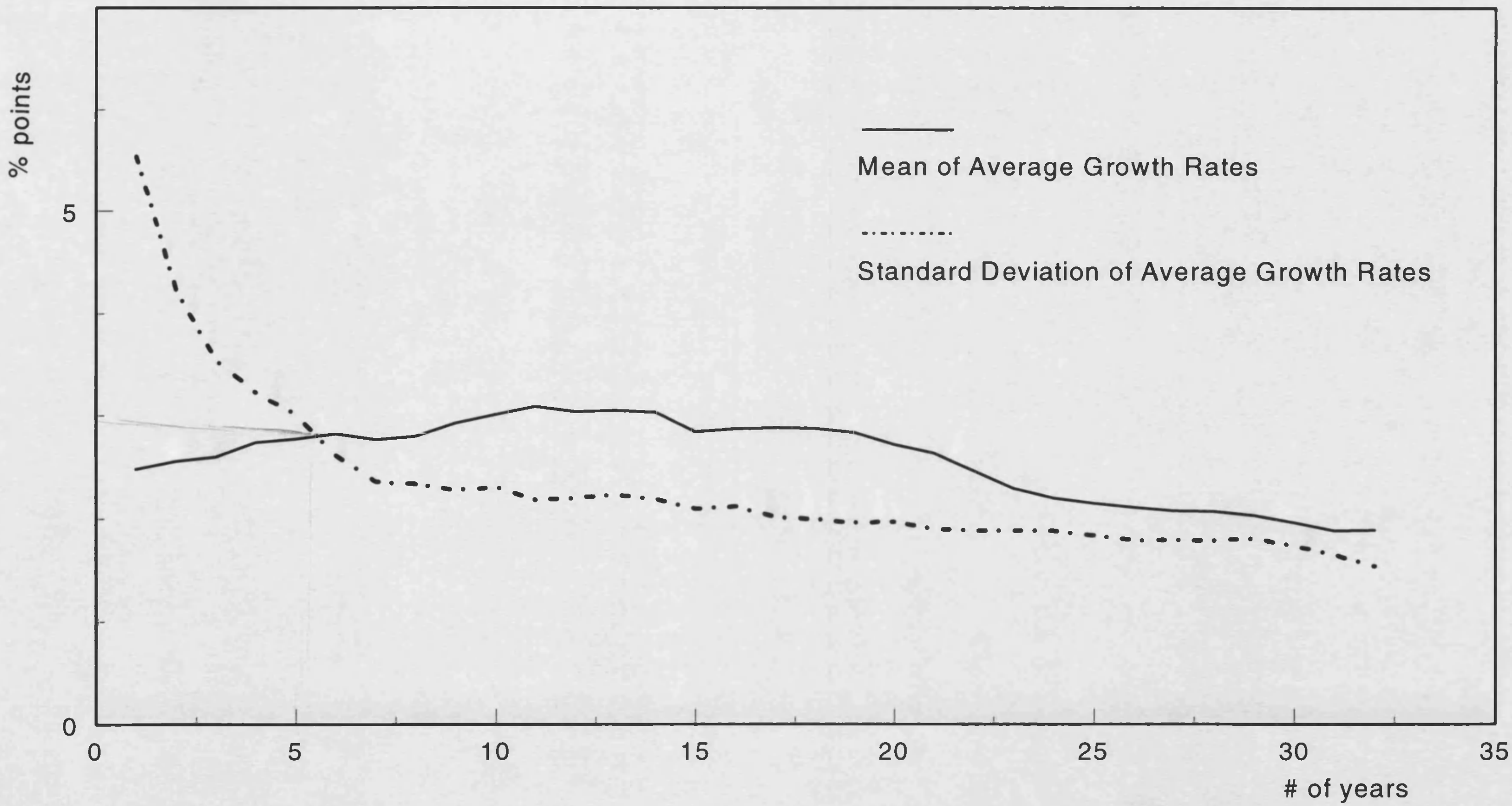


Figure - 1.2

Historical Movements in Model Coefficients

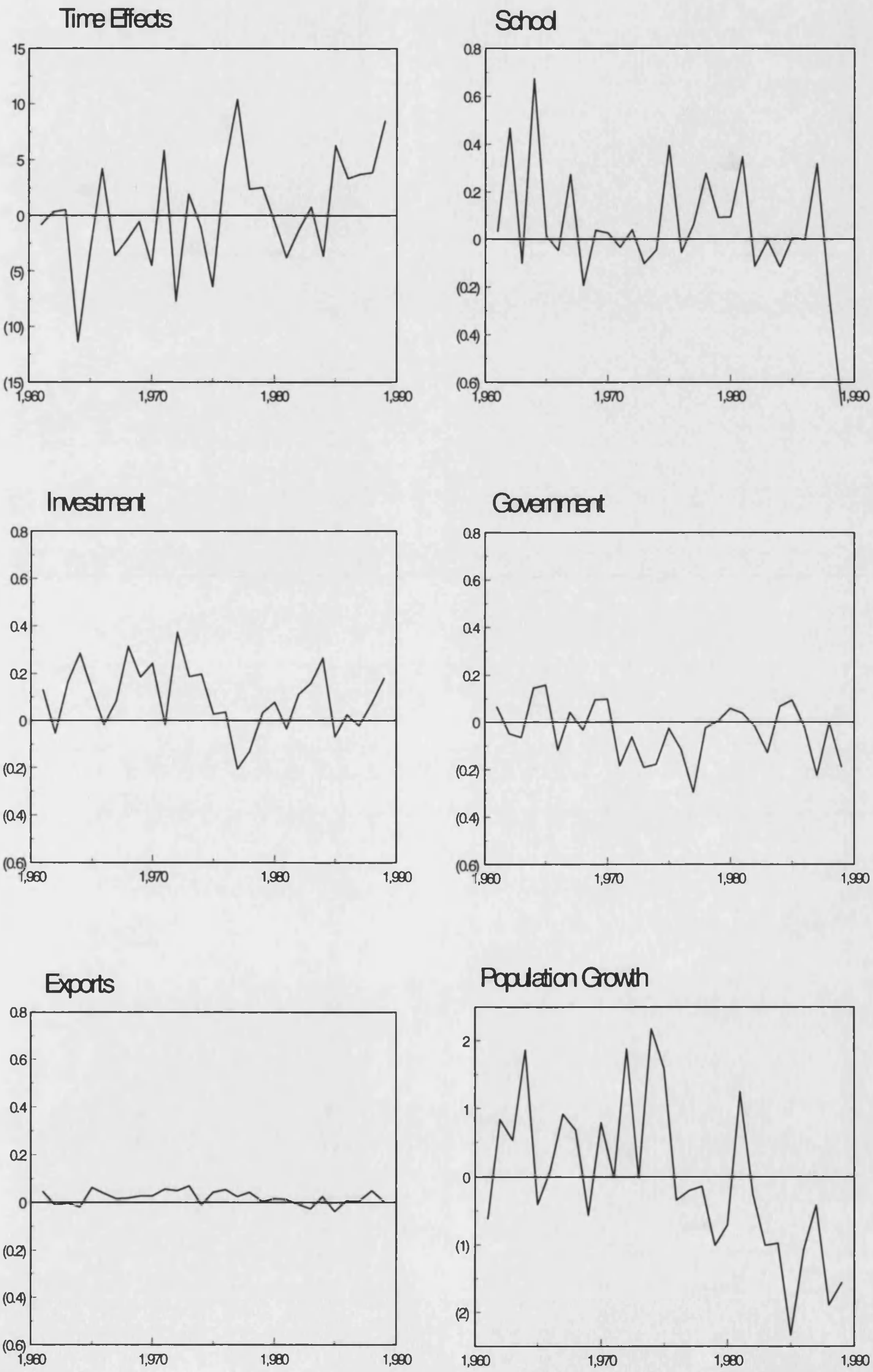


Figure - 1.3
Historical Decomposition of Growth in US

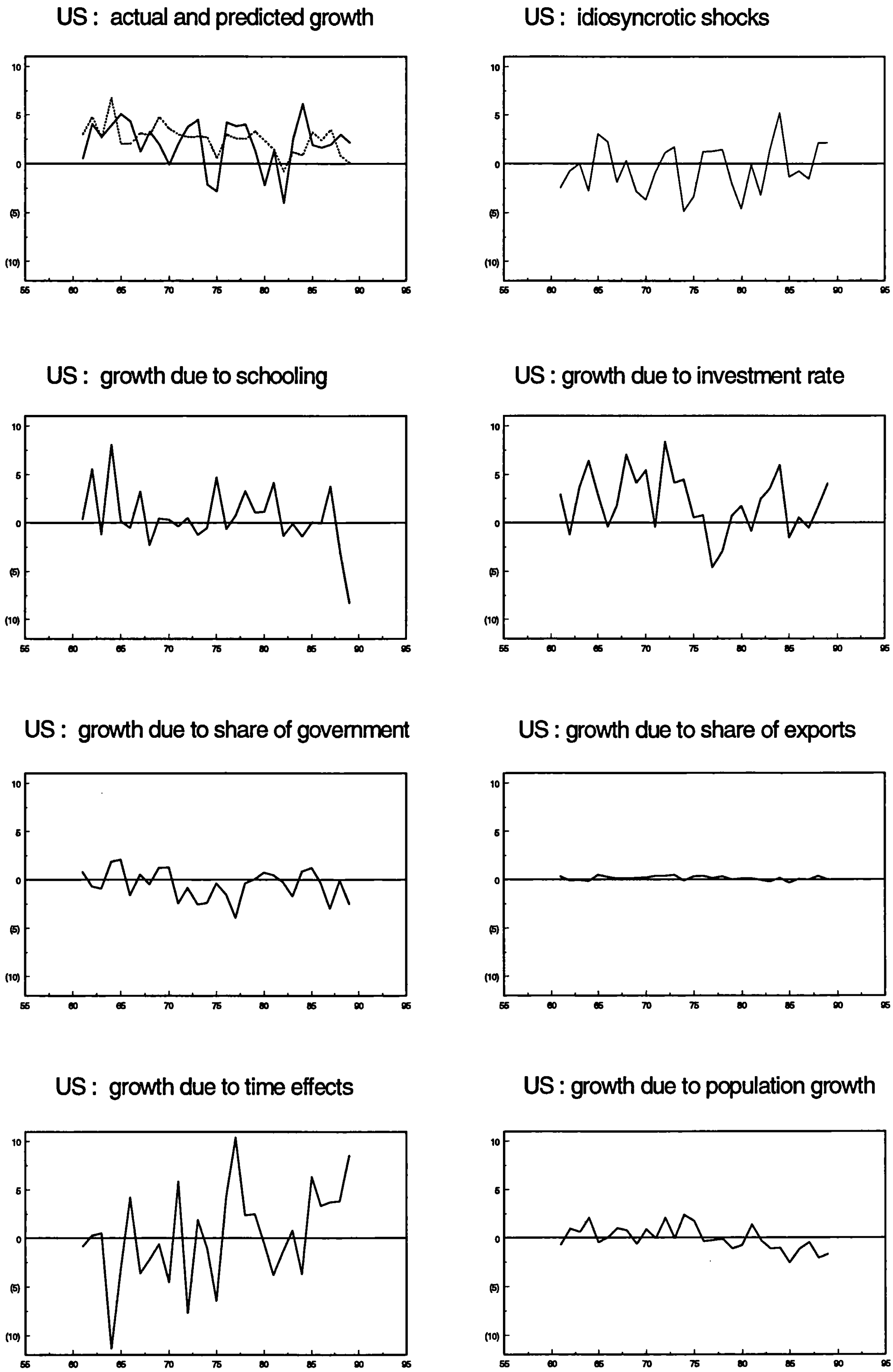


Figure - 1.4

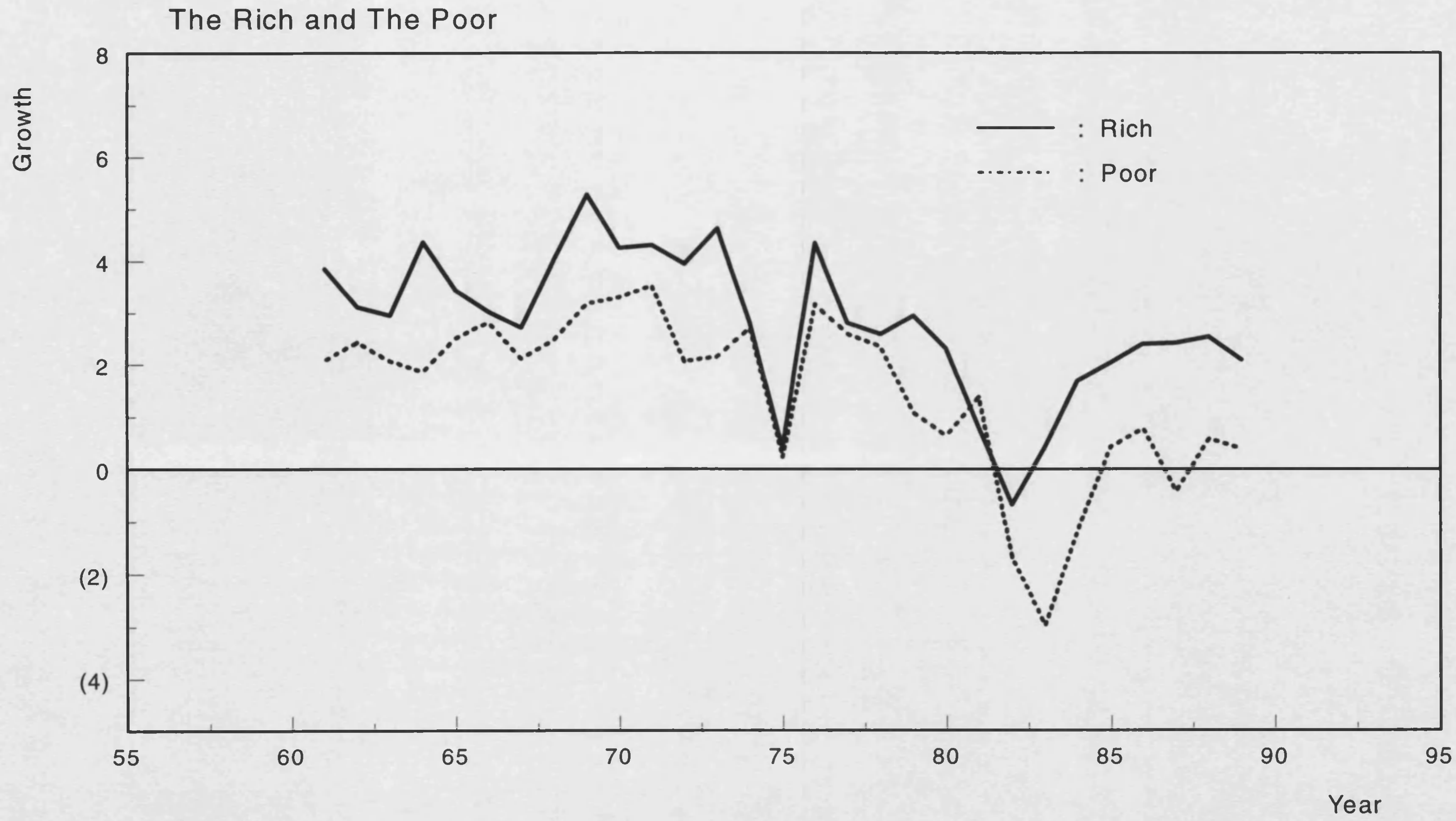

The Rich and The Poor
Growth
8
6
4
2
0
(2)
(4)
: Rich
: Poor
55
60
65
70
75
80
85
90
95
Year

Figure - 1.5

The Difference Between Rich and Poor

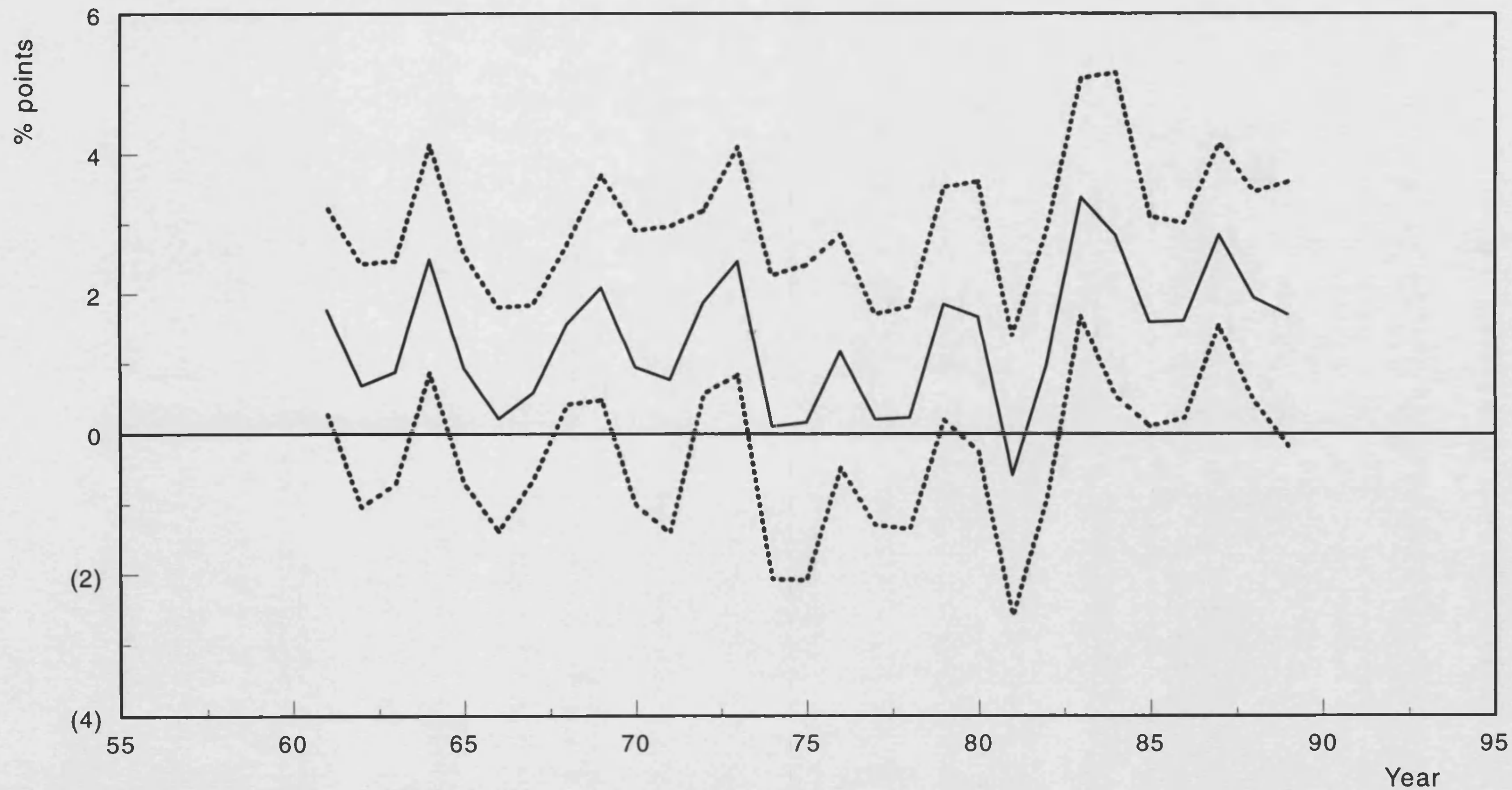


Table - 1.1

The OLS Results for Alternative Settings

		Time Effects	School	Investment	Government	Exports	Pop
SES-1	Range of Coef.	[-11.37, 8.50]	[-0.70, 0.67]	[-0.2, 0.37]	[-0.30, 0.16]	[-0.04, 0.07]	[-2.34, 2.17]
	Wald Test	REJECT	REJECT	REJECT	REJECT	ACCEPT	REJECT
	Years	5	2	4	2	3	5
SES-2	Range of Coef.	[-7.43, 10.99]	[-0.49, 1.01]		[-0.25, 0.13]	[-0.04, 0.08]	[-2.25, 1.93]
	Wald Test	REJECT	REJECT		REJECT	REJECT	REJECT
	Years	3	4		4	5	6
SES-3	Range of Coef.	[-6.40, 10.12]	[-0.52, 0.46]	[-0.18, 0.37]	[-0.29, 0.17]		
	Wald Test	REJECT	ACCEPT	REJECT	REJECT		
	Years	3	2	9	2		
SES-4	Range of Coef.	[-3.69, 7.91]	[-0.20, 0.78]		[-0.27, 0.12]		
	Wald Test	REJECT	ACCEPT		REJECT		
	Years	8	6		5		
SES-5	Range of Coef.	[-5.99, 3.30]	[-0.47, 0.42]	[-0.10, 0.37]			
	Wald Test	REJECT	ACCEPT	REJECT			
	Years	6	1	7			
SES-6	Range of Coef.	[-4.66, 3.02]	[-0.02, 0.70]				
	Wald Test	REJECT	ACCEPT				
	Years	6	7				

Table - 1.2

The Endowments

	School	Invest.	Govern	Exports	Pop. Growth
US	11.86	22.28	13.28	7.60	1.10
UK	8.89	17.69	18.34	24.71	0.29
Japan	10.90	32.45	8.69	11.97	0.93
All					
Rich	8.64	22.28	13.64	31.38	1.23
Poor	4.46	12.74	19.22	30.84	2.42

Table - 1.3

The Shares of Factor Endowments in Cumulative Growth (%);

	Time Effects	School	Invest.	Govern	Exports	Pop. Growth	Idiosyn Shocks
US	1.53	6.34	83.11	4.78	0.40	0.03	4.21
UK	2.29	5.31	78.18	13.60	6.34	0.00	0.62
Japan	0.66	2.30	75.72	0.88	0.43	0.00	20.44
All	1.56	2.03	52.45	9.68	9.61	0.11	34.16
Rich	1.25	2.95	73.67	4.69	7.41	0.04	17.40
Poor	1.73	1.49	39.95	12.63	10.90	0.16	44.05

Table = 1.4

Year by Year Growth Experience

Year	Rich	Poor	Difference	t-ratio
1961	3.85	2.08	1.78	2.37
1962	3.12	2.42	0.70	0.78
1963	2.96	2.07	0.89	1.09
1964	4.37	1.86	2.50	3.01
1965	3.45	2.51	0.94	1.12
1966	3.04	2.82	0.22	0.27
1967	2.72	2.12	0.59	0.93
1968	4.04	2.47	1.57	2.67
1969	5.28	3.19	2.09	2.55
1970	4.26	3.31	0.95	0.96
1971	4.31	3.53	0.78	0.69
1972	3.96	2.07	1.89	2.83
1973	4.63	2.16	2.47	2.97
1974	2.82	2.72	0.11	0.10
1975	0.40	0.23	0.17	0.15
1976	4.35	3.17	1.18	1.39
1977	2.82	2.60	0.21	0.28
1978	2.59	2.36	0.24	0.29
1979	2.95	1.09	1.87	2.19
1980	2.33	0.65	1.68	1.70
1981	0.79	1.38	-0.58	-0.57
1982	-0.68	-1.68	1.00	1.01
1983	0.41	-2.98	3.39	3.91
1984	1.70	-1.16	2.86	2.42
1985	2.04	0.43	1.62	2.11
1986	2.41	0.79	1.63	2.28
1987	2.44	-0.42	2.86	4.30
1988	2.55	0.58	1.97	2.58
1989	2.11	0.39	1.72	1.78

Table - 1.5

The Cross Product Matrix of the Coefficients

	Time Effects	School	Invest.	Govern.	Exports	Pop. Growth
Time Effects	656.55					
School	-16.88	1.86				
Invest.	-10.68	-0.12	0.78			
Govern.	-9.96	0.19	0.03	0.42		
Exports	0.21	-0.01	0.05	-0.03	0.03	
Pop. Growth	-101.67	4.08	1.18	-0.05	0.02	36.11

Table - 1.6

The First Stage Coefficients in 2SLS

	School	Investment	Govern.	Export	Pop. Gr.
Constant	0.268 (0.55)	3.754 (2.91)	3.449 (2.46)	5.236 (2.12)	0.772 (4.26)
School	1.104 (24.08)	0.056 (0.46)	-0.034 (-0.26)	-0.125 (-0.54)	-0.112 (-6.54)
Investment	0.001 (0.04)	0.845 (19.40)	-0.049 (-1.04)	-0.032 (-0.38)	0.003 (0.48)
Govern.	-0.0.16 (-1.04)	-0.109 (-2.70)	0.887 (20.22)	0.017 (0.22)	0.004 (0.68)
Export	0.005 (0.91)	0.017 (1.16)	0.018 (1.16)	1.075 (39.09)	0.001 (0.70)
Population Growth	0.434 (3.65)	0.220 (0.69)	0.612 (1.78)	-1.103 (-1.82)	0.756 (16.99)
Summary of Second Stage Coefficients in 2SLS Dependent Variable: Annual growth					
Range of Coeffs.	[-0.71, 0.68]	[-0.25, 0.34]	[-0.30, 0.14]	[-0.04, 0.07]	[-2.95, 2.73]
Wald Test	Accept	Accept	Accept	Accept	Reject

Table - 1.7

Mean effects

	Simple Average of the OLS	Minimum Distance	Random Coefficients
Time effects	0.283 [0.37] (0.24)	0.191 [0.32] (0.14)	0.148 [0.25] (0.25)
School	0.049 [0.86] (0.66)	0.053 [1.22] (0.68)	0.053 [1.21] (1.21)
Investment	0.094 [3.86] (2.41)	0.102 [5.48] (2.24)	0.101 [5.38] (5.36)
Government	-0.038 [-1.67] (-1.19)	-0.021 [-1.20] (-0.59)	-0.019 [-1.12] (-1.11)
Exports	0.019 [2.76] (2.06)	0.014 [2.58] (1.38)	0.014 [2.60] (2.60)
Pop. Growth	-0.034 [-0.19] (-0.12)	-0.082 [-0.61] (-0.26)	-0.072 [-0.54] (-0.53)

Table - 1.8

Covariogram of the Idiosyncratic Shocks:

Lag	0	1	2	3	4	5	6	7
Mean	29.17	3.31	1.01	-3.63	1.89	-0.60	1.23	-0.00
Max.	65.26	18.31	12.03	5.77	4.85	11.99	16.10	9.29
Min.	16.10	-19.62	-10.99	-32.24	-7.13	-7.27	-7.53	-5.47
χ^2-test	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject

Lag	8	9	10	11	12	13	14	15
Mean	-0.74	2.47	-0.51	2.96	-1.25	-1.84	-2.83	4.22
Max.	8.35	10.18	11.53	8.35	3.76	5.63	6.45	8.54
Min.	-7.27	-4.40	-12.66	-8.66	-6.42	-6.71	-6.08	-8.19
χ^2-test	Reject	Accept	Reject	Accept	Accept	Accept	Reject	Reject

Lag	16	17	18	19	20	21	22	23
Mean	0.76	0.47	0.20	-0.86	-1.16	0.49	-0.89	-0.55
Max.	5.38	4.99	4.53	6.75	4.28	6.89	2.17	2.31
Min.	-6.16	-9.95	-7.15	-3.68	-4.73	-6.77	-5.95	-7.52
χ^2-test	Accept	Reject	Reject	Reject	Accept	Reject	Accept	Accept

Lag	24	25	26	27	28
Mean	-0.87	-3.64	3.55	0.86	0.96
Max.	3.18	1.92	10.34	3.31	0.96
Min.	-2.39	-6.27	-0.33	-1.47	0.96
χ^2-test	Accept	Accept	Accept	Accept	-

Table - 1.9

The Structural Decomposition of the Idiosyncratic Shocks
(Using Minimum Distance procedure)

			1	2	3	4	5	6	7	8
Perm.	W.N.	σ^2	16.01	-	-	-	16.01	-	-	-
	AR(1)	σ^2	-	15.97	-	-	-	15.97	-	15.97
		ρ	-	0.19	-	-	-	0.19	-	0.19
Temp.	AR(1)	σ^2	-	-	-	16.01	-	-	16.01	0.00
		ρ	-	-	-	1.00	-	-	1.00	---
Ind. Effects		σ^2	0.69	0.45	1.51	0.69	0.69	0.45	0.69	0.45
Measu. Error		σ^2	-	-	6.65	-	0.00	0.00	0.00	0.00

The Structural Decomposition of the Idiosyncratic Shocks
(Using Pseudo-Maximum Likelihood procedure)

			1	2	3	4	5	6	7	8
Perm.	W.N.	σ^2	35.48	-	-	-	35.47	-	-	-
	AR(1)	σ^2	-	35.60	-	-	-	28.09	-	21.55
		ρ	-	0.05	-	-	-	0.20	-	0.29
Temp.	AR(1)	σ^2	-	-	-	35.48	-	-	35.47	6.60
		ρ	-	-	-	1.00	-	-	1.00	0.62
Ind. Effects		σ^2	0.53	0.41	1.45	0.53	0.53	0.31	0.53	0.41
Measu. Error		σ^2	-	-	9.12	-	0.00	3.80	-0.00	2.93

Table - 1.10

Quasi - R^2

	Simple Average of the OLS	Minimum Distance	Random Coefficients
1	0.878	0.885	0.884
2	0.917	0.922	0.921
3	0.766	0.779	0.777
4	0.878	0.885	0.884
5	0.878	0.885	0.884
6	0.917	0.922	0.921
7	0.878	0.885	0.884
8	0.917	0.922	0.921

1.5 Technical Appendix - A

Cross Country Growth Regressions:

The popular approach in the analysis of cross country differentials in post war growth experiences is to regress average growth rates on several explanatory variables which also include initial values of per capita incomes. A negative coefficient for the initial values is interpreted as convergence conditional on the other explanatory variables. Furthermore, this is taken as an evidence in favour of exogenous growth models and argued to conflict with the ideas in endogenous growth theory. However, this conjecture is not correct, and a negative coefficient for the initial incomes is also expected in endogenous growth models. To see this notice that the regression coefficient of initial incomes is given by

$$\hat{\rho} \;=\; \frac{\mathrm{cov}\left(g_i, y_{i,0} \mid x_i\right)}{\mathrm{var}\left(g_i \mid x_i\right)} \tag{A.1}$$

where

$$g_i \;=\; \frac{1}{T}\left(y_{i,T} - y_{i,0}\right), \tag{A.2}$$

x_i are the set of explanatory variables, and $y_{i,T}$ and $y_{i,0}$ are the logarithms of the final and the initial values of per capita incomes. By substituting the identity in (A.2) into (A.1), one obtains the necessary and sufficient condition for obtaining a negative coefficient[12]:

$$\hat{\rho} < 0 \quad \text{iff} \quad \mathrm{cov}\left(y_{i,T}, y_{i,0} \mid x_i\right) < \mathrm{var}\left(y_{i,0} \mid x_i\right) \tag{A.3}$$

Hence, a negative coefficient for the initial incomes is obtained whether or not growth is endogenously determined (if one includes sufficient explanatory variables to explain the cross sectional variation of incomes). The difference that comes from the endogeneity of growth is that the correlation

---

12 The negative coefficient in a cross country growth regression is called as (conditional) β-convergence. It is clear from (A.3) that a decreasing cross sectional variance of incomes, what is known as (conditional) σ-convergence, is sufficient for (conditional) β-convergence.

between explanatory variables and log per capita incomes would increase over time.

1.6 Technical Appendix - B

Estimating Mean Effects

Consider the panel data set-up

$$g_i(t) \;=\; x_i\,\beta(t) \;+\; \lambda(t) \;+\; \mu_i \;+\; \nu_i(t) \tag{B.1}$$

We assume that the individual effects, μ, and the idiosyncratic disturbances are independently identically distributed in the cross section;

$$\mu_i \;\sim\; \mathrm{IID}(0,\sigma_\mu^2)$$

$$\nu_i \;\sim\; \mathrm{IID}(0,\Omega)$$

The *mean effect* of a variable on the long run growth rate is defined as the unconditional expected value of its coefficients in (B.1). To estimate these mean effects, we first re-formulate the system in its simultaneous equations set-up (SES) where each equation corresponds to a single period;

$$\begin{bmatrix} g_i(1) \\ g_i(2) \\ \vdots \\ g_i(T) \end{bmatrix} = \begin{bmatrix} \lambda(1) \\ \lambda(2) \\ \vdots \\ \lambda(T) \end{bmatrix} + \begin{bmatrix} \beta'(1) \\ \beta'(2) \\ \vdots \\ \beta'(T) \end{bmatrix} \cdot x_i + \begin{bmatrix} \mu_i + \varepsilon_i(1) \\ \mu_i + \varepsilon_i(2) \\ \vdots \\ \mu_i + \varepsilon_i(T) \end{bmatrix} \tag{B.2}$$

Here T is the number of periods. This SES can also be represented in matrix form;

$$G = B \cdot x + E \tag{B.3}$$

where G and x are matrixes of observed growth rates and the explanatory variables respectively (a constant is also included in x to account for time effects). E is the matrix of disturbances where each row corresponds to a single period and each column corresponds to a single cross sectional unit. B is the matrix of coefficients in which rows correspond to a particular period and columns to a particular explanatory variable (first column being the vector of time effects). Notice that this coefficient matrix B can be efficiently estimated by using OLS;

$$\hat{B} = Gx'(xx')^{-1} \tag{B.4}$$

We consider five different methods to estimate the mean effects. Two of those are numerically equivalent to the others. Next section reviews all these five methods and establishes the equivalence results.

(1) The OLS on Average Growth Rates:

This is the standard approach taken by the previous empirical literature. To estimate the mean effects we simply calculate the average growth rates during the sample period and regress them on the explanatory variables:

$$\hat{\bar{\beta}}_1 = (xx')^{-1} x\bar{g}' \tag{B.5}$$

where

$$\bar{g} = \frac{\iota_T' G}{T} \tag{B.6}$$

and ι_T is a Tx1 vector of ones.

Let β be the unconditional mean of $\beta(t)$. Then, by substituting (B.2) and (B.6) into (B.5), one can obtain

$$\hat{\bar{\beta}}_1 = \beta + \frac{\Gamma'\iota_T}{T} + (xx')^{-1}\frac{xE'\iota_T}{T} \tag{B.7}$$

where $\Gamma = B - \iota_T\beta'$ is the deviation of realised coefficients from their unconditional mean. Clearly $\hat{\bar{\beta}}_1$ is an unbiased estimator both for mean effects, β, and the average of the realised coefficients during the sample period, $\bar{\beta} = \frac{B'\iota_T}{T} = \bar{\beta}$. Under the usual regularity assumptions, however, it is not consistent for the mean effects as the cross sectional observations increase unboundedly for a fixed T.

The covariance matrix of the estimates is given as follows;

$$V\left(\hat{\bar{\beta}}_I\right) = E\left(\left(\hat{\bar{\beta}}_I - \beta\right)\left(\hat{\bar{\beta}}_I - \beta\right)'\right)$$

$$= \frac{\Lambda}{T} + \left(x\,x'\right)^{-1} \frac{\iota_T' \, \Sigma \iota_T}{T^2} \tag{B.8}$$

where,

$$\Lambda = E\left((\beta - \beta(t))(\beta - \beta(t))'\right), \tag{B.9}$$

$$\Sigma = \sigma_\mu^2 \, \iota_T \, \iota'_T + \Omega. \tag{B.10}$$

In this paper, although correlation between shocks to the individual coefficients are allowed, we assume they are serially uncorrelated. Despite being restrictive, it considerably simplifies the algebra and the calculations.

The covariance matrix in (B.8) has two components. The first one is due to the variation of coefficients in time and is the source of inconsistency for mean effects in the cross sectional dimension. The second one is the usual covariance matrix and vanishes asymptotically as the number of cross sectional observations increases. We, therefore, calculate two different t-ratios: one by using the covariance matrix in (B.8), and another by using only the second component of the same covariance matrix. The first one is only a quasi t-ratio for the null hypothesis that mean effects are zero, and its asymptotic distribution is dependent on the actual distribution of the innovations. The second t-ratio has the standard asymptotic normal distribution under the null that the sample average of the coefficients is zero. Despite being less interesting, it is this t-ratio which has been reported in previous empirical literature with cross country growth regressions.

(2) The Simple Average of SES OLS coefficients:

The simplest way of obtaining a set of estimators for the mean effects is simply to take the average of the estimated SES coefficients:

$$\hat{\bar{\beta}}_2 = \frac{\hat{B}'\iota_T}{T} \tag{B.11}$$

The following theorem states that the estimators in (B.11) and (B.5) are numerically identical.

Theorem 1: Consider the model given by (B.1). Then $\hat{\bar{\beta}}_1 = \hat{\bar{\beta}}_2$.

Proof:

$$\hat{\bar{\beta}}_1 = (x\,x')^{-1}\,x\bar{g}' \quad = (x\,x')^{-1}\,\frac{x\,G'\iota_T}{T}$$

$$\hat{\bar{\beta}}_1 = \frac{\hat{B}'\iota_T}{T} \quad = \hat{\bar{\beta}}_2$$

(3) Chamberlain's Minimum Distance Estimator:

One easy way of gaining efficiency in estimating mean effects is to use a weighted average of the estimated SES coefficients rather than the simple average. This can be done by using the Chamberlain's minimum distance estimator. In particular, we estimate mean effects such that

$$\hat{\bar{\beta}}_3 = \underset{\beta}{\arg\min}\ \mathrm{vec}\left(\hat{B} - \beta'\otimes\iota_T\right)'\ \hat{\Omega}_B^{-1}\ \mathrm{vec}\left(\hat{B} - \beta'\otimes\iota_T\right) \tag{B.12}$$

where $\hat{B}$ and $\hat{\Omega}_B$ are the OLS estimates of the SES coefficients and their covariance matrix. The explicit solution to the minimisation problem in (B.12) is given by

$$\hat{\bar{\beta}}_3 = \left(\iota_T'\,\hat{\Sigma}^{-1}\,\iota_T\right)^{-1}\hat{B}'\hat{\Sigma}^{-1}\,\iota_T \tag{B.13}$$

where $\hat{\Sigma}$ is the OLS estimate of the residual covariance matrix in (B.10). Clearly (B.13) is a weighted average of the estimated coefficients in SES set-up. One can expend this expression to obtain

$$\hat{\hat{\beta}}_3 = \beta + \left(\iota_T' \hat{\Sigma}^{-1} \iota_T\right)^{-1} \hat{\Gamma}' \hat{\Sigma}^{-1} \iota_T + \left(\iota_T' \hat{\Sigma}^{-1} \iota_T\right)^{-1} \left(x x'\right)^{-1} x E' \hat{\Sigma}^{-1} \iota_T \qquad \text{(B.14)}$$

It is clear from (B.14) that minimum distance estimator provides an unbiased estimate both for the mean effects and a weighted sample average of the realised coefficients. Its asymptotic covariance matrix is given by

$$\begin{aligned} V\left(\hat{\hat{\beta}}_3\right) &= E\left(\left(\hat{\hat{\beta}}_3 - \beta\right)\left(\hat{\hat{\beta}}_3 - \beta\right)'\right) \\ &= \frac{\iota_T' \Sigma^{-2} \iota_T}{\left(\iota_T' \Sigma^{-1} \iota_T\right)^2} \Lambda + \left(\iota_T' \Sigma^{-1} \iota_T\right)^{-1} \left(x x'\right)^{-1} \qquad \text{(B.15)} \end{aligned}$$

Just as the covariance matrix in (B.8), (B.15) has two components. The first one is due to the variation of coefficients in time and is the source of the inconsistency. The second one is due to the estimation of coefficients from the cross sectional observations and vanishes asymptotically.

(4) Quasi-GLS Estimator:

If one assumes time invariant coefficients, the SES set-up in (B.3) can be re-formulated by using the "vec" operator;

$$g = \left(x' \otimes \iota_T\right) \beta + \nu \qquad \text{(B.16)}$$

where ν is vec(E), and independently identically distributed with covariance matrix $\left(I_N \otimes \Sigma\right)$. One can efficiently estimate β from (B.16) with GLS and by using the OLS estimate of Σ from the SES in (B.3). The explicit formulae for the estimate is given by

$$\hat{\hat{\beta}}_4 = \left(\left(x x'\right)^{-1} \otimes \left(\iota_T' \hat{\Sigma}^{-1} \iota_T\right)^{-1}\right) \left(x \otimes \iota_T' \hat{\Sigma}^{-1}\right) g \qquad \text{(B.17)}$$

Notice that when coefficients vary in time this is only a quasi-GLS and that the estimates are not efficient. This is because, the covariance matrix of the residuals is then mis-specified. The following theorem states that it is numerically equivalent to the minimum distance estimator in (B.13).

Theorem 2: Consider the model given by (B.1). Then $\hat{\bar{\beta}}_3 = \hat{\bar{\beta}}_4$.

Proof:

$$\begin{aligned}
\hat{\bar{\beta}}_4 &= \left((x\,x')^{-1} \otimes \left(\iota_T' \hat{\Sigma}^{-1} \iota_T\right)^{-1}\right)\left(x \otimes \iota_T' \hat{\Sigma}^{-1}\right) \mathrm{vec}(G) \\
&= (x\,x')^{-1} \left(\iota_T' \hat{\Sigma}^{-1} \iota_T\right)^{-1} \mathrm{vec}\left(\iota_T' \hat{\Sigma}^{-1} G\, x'\right) \\
&= \left(\iota_T' \hat{\Sigma}^{-1} \iota_T\right)^{-1} (x\,x')^{-1}\, x\, G' \hat{\Sigma}^{-1} \iota_T \;=\; \hat{\bar{\beta}}_3
\end{aligned}$$

(5) Random Coefficients Estimator:

If the coefficients in (B.3) vary in time, then the correct covariance matrix of the disturbances in (B.16) is given by

$$E(vv') = \left(I_N \otimes \Sigma\right) + \left(x' \Lambda\, x \otimes I_T\right) \tag{B.18}$$

In this case, the efficient estimate of β can be obtained by using GLS and a consistent estimate of the correct covariance matrix in (B.18):

$$\hat{\bar{\beta}}_5 = \left(\left(x \otimes \iota_T'\right) \hat{\Omega}_V^{-1} \left(x' \otimes \iota_T\right)\right)^{-1} \left(x \otimes \iota_T'\right) \hat{\Omega}_V^{-1}\, g \tag{B.19}$$

where $\hat{\Omega}_V$ is an estimate of the covariance matrix in (B.18) that is calculated by using the results of OLS regression of the SES. Under certain regularity conditions, the covariance matrix of the estimated mean effects is given by

$$\begin{aligned}
V\left(\hat{\bar{\beta}}_5\right) &= E\left(\left(\hat{\bar{\beta}}_5 - \beta\right)\left(\hat{\bar{\beta}}_5 - \beta\right)'\right) \\
&= \left(\left(x \otimes \iota_T'\right) \Omega_V^{-1} \left(x' \otimes \iota_T\right)\right)^{-1}
\end{aligned} \tag{B.20}$$

1.7 Data Appendix

List of Countries in Empirical Exercise:

1. ALGERIA
2. ANGOLA
3. BENIN
4. BOTSWANA
5. BURKINA FASO
6. BURUNDI
7. CAMEROON
8. CENTRAL AFR.R.
9. CHAD
10. CONGO
11. EGYPT
12. GABON
13. GAMBIA
14. GHANA
15. GUINEA
16. IVORY COAST
17. KENYA
18. LESOTHO
19. MADAGASCAR
20. MALAWI
21. MALI
22. MAURITANIA
23. MAURITIUS
24. MOROCCO
25. MOZAMBIQUE
26. NIGER

27. NIGERIA
28. RWANDA
29. SENEGAL
30. SOMALIA
31. SOUTH AFRICA
32. SWAZILAND
33. TOGO
34. TUNISIA
35. UGANDA
36. ZAIRE
37. ZAMBIA
38. ZIMBABWE
39. BARBADOS
40. CANADA
41. COSTA RICA
42. DOMINICAN REP.
43. EL SALVADOR
44. GUATEMALA
45. HAITI
46. HONDURAS
47. JAMAICA
48. MEXICO
49. NICARAGUA
50. PANAMA
51. TRINIDAD&TOBAGO
52. U.S.A.
53. ARGENTINA
54. BOLIVIA
55. BRAZIL

56. CHILE
57. COLOMBIA
58. ECUADOR
59. GUYANA
60. PARAGUAY
61. PERU
62. SURINAME
63. URUGUAY
64. VENEZUELA
65. BANGLADESH
66. HONG KONG
67. INDIA
68. INDONESIA
69. IRAN
70. ISRAEL
71. JAPAN
72. JORDAN
73. KOREA, REP.
74. MALAYSIA
75. MYANMAR
76. PAKISTAN
77. PHILIPPINES
78. SAUDI ARABIA
79. SINGAPORE
80. SRI LANKA
81. SYRIA
82. THAILAND
83. AUSTRIA
84. BELGIUM

85. CYPRUS
86. DENMARK
87. FINLAND
88. FRANCE
89. GERMANY, WEST
90. GREECE
91. ICELAND
92. IRELAND
93. ITALY
94. LUXEMBOURG
95. MALTA
96. NETHERLANDS
97. NORWAY
98. PORTUGAL
99. SPAIN
100. SWEDEN
101. SWITZERLAND
102. TURKEY
103. U.K.
104. AUSTRALIA
105. FIJI
106. NEW ZEALAND
107. PAPUA N.GUINEA

1.8 References

Acemoglu, D., `Market Imperfections, Investment in Skills and Growth', 1993, CEPR Conference Paper.

Acemoglu, D., and Zilibotti, F., `Was Prometheus Unbound By Chance? Risk, Diversification and Growth', 1995, Unpublished Draft Copy.

Barro, R. J., `Economic Growth in a Cross Section of Countries', 1991, Quarterly Journal of Economics, Vol. 106, pp.407-43.

Barro, R. J., and Sala-i Martin, X., `Convergence across States and Regions', 1991, Brookings Papers on Economic Activity, pp.107-182.

-------, `Convergence', 1992a, Journal of Political Economy, Vol. 100, pp.223-51.

-------, `Public Finance in Models of Economic Growth', Review of Economic Studies, 1992b, Vol. 59, pp.645-61.

Barro, R. J., and Lee, J.W., `Losers and Winners in Economic Growth', 1993, National Bureau of Economic Research, Working Paper, No.4341.

Baumol, W.J., `Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show', 1986, American Economic Review, Vol.76, pp.1073-85.

Baumol, W.J., and Wolff, E.N., `Productivity Growth, Convergence, and Welfare: Reply', 1988, American Economic Review, Vol. 78, pp.1155-59

Bernard, A. B., and Durlauf S. N., `Convergence of International Output Movements', 1991, National Bureau of Economic Research, Working Paper No. 3717.

Beveridge, S., and Nelson, C. R., `A new Approach to Decomposition of Economic Time series into Permanent and Transitory Components with Particular Attention to Measurement of the `Business Cycle", 1981, Journal of Monetary Economics, Vol. 7, pp.151-54.

Blanchard, O. J., and Quah, D., `The Dynamic Effects of Demand and Supply Disturbances', 1989, American Economic Review, Vol.79, pp.655-73.

Ben-David, D., `Equalizing Exchange: A Study of the Effects of Trade Liberalization', 1991, National Bureau of Economic Research, Working Paper No. 3706.

Campbell, J. Y., and Mankiw N. G., `Are Output Fluctuations Transitory', 1987, Quarterly Journal of Economics, Vol. 102, pp.857-80.

Canning, D., `The Effect of Transportation Network on Economic Growth', 1994, CEP Seminar Paper.

Cass, D., `Optimum Growth in an Aggregative Model of Capital Accumulation', 1965, Review of Economic Studies, Vol. 32, pp.233-40.

Chamberlain, G., `Multivariate Regression Models for Panel Data', 1982, Journal of Econometrics, vol.18, pp.5-46.

Chamberlain, G., `Panel Data', 1984, Handbook of Econometrics, vol. II, edited by Z. Griliches and M. Intriligator, pp.1247-318.

Cochrane, J., `How Big Is the Random Walk in GDP', 1988, Journal of Political Economy, Vol. 96, pp.893-920.

De Long, J. B., `Productivity Growth, Convergence, and Welfare: Comment', 1988, American Economic Review, Vol. 78, pp.1139-54.

Feder, G., `On Exports and Economic Growth', 1982, Journal of Development Economics, Vol.12, pp.59-73.

Friedman, M., `Do Old Fallacies Ever Die?', 1992, Journal of Economic Literature, Vol. 30, pp.2129-32.

Fischer, S., `The Role of Macroeconomic Factors in Growth', 1993, Journal of Monetary Economics, Vol.32, pp.485-512.

Kiefer, N.M., `Estimation of Fixed Effects Models for Time Series of Cross-Sections with Arbitrary Intertemporal Covariance', 1980, Journal of Econometrics, Vol.14, pp.195-202.

King, R.G., and Levine R., `Finance and Growth: Schumpeter Might Be Right', 1993, Quarterly Journal of Economics, Vol. , pp.717-737.

Lucas, R. E., `On the Mechanics of Economic Development', 1988, Journal of Monetary Economics, Vol. 22, pp.3-22.

-------, `Making Miracle', 1993, Econometrica, Vol. 61, pp.251-72.

Levine, R., and Renelt, D., `A Sensitivity Analysis of Cross Country Growth regressions', 1992, American Economic Review, Vol.82, pp.942-63.

MaCurdy, T.E., `The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis', 1982, Journal of Econometrics, Vol.18, pp.83-114.

Mankiw, N. W., Romer D., and Weil D. N., `A Contribution to the Empirics of Economic Growth', 1992, Quarterly Journal of Economics, Vol. 107, pp.407-432.

Nelson, R. R., and Phelps E. S., `Investment in Humans, Technological Diffusion, and Economic Growth', 1966, American Economic Review, Papers and Proceedings, Vol. 56, pp.69-75.

Parente, S.L., and Prescott, E. C., ` Technology Adoption and Growth', 1991, NBER Working Paper: 3733.

-------, `Barriers to Technology Adoption and Development', 1994, Journal of Political Economy, Vol. 102, pp.298-321.

Quah, D., `The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds', 1992, Econometrica, Vol. 60, No. 1, pp.107-18.

-------, `Galton's Fallacy and Tests of the Convergence Hypothesis', 1993a, Scandinavian Journal of Economics, Vol. 95, pp.427-43.

------, `Empirical cross-section dynamics in economic growth', 1993b, European Economic Review, Vol. 37, pp.426-34.

Rebelo, S., `Long-Run Policy Analysis and Long-Run Growth', 1991, Journal of Political Economy, Vol. 99, pp.500-21.

Ram, R., `Government Size and Economic Growth: A New Framework and Some Evidence From Cross Section Time Series Data', 1986, American Economic Review, Vol.76, pp.191-203.

Romer, P. M., `Increasing Returns and Long-Run Growth', 1986, Journal of Political Economy, Vol. 94, pp.1002-37.

------, `Endogenous Technological Change', 1990, Journal of Political Economy, Vol. 98, pp.S71-S102.

Solow, R. M., `A Contribution to the Theory of Economic Growth', 1956 Quarterly Journal of Economics, Vol. 70, pp.65-94.

------, `Perspectives on Growth Theory', 1994, The Journal of Economic Perspectives, Vol., pp.45-54.

Stokey, N. L., `Human Capital Product Quality and Growth', 1991, Quarterly Journal of Economics, Vol.106, pp.587-616.

Summers, R., and Heston A., `The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988', 1991, Quarterly Journal of Economics, Vol. 106, pp.327-368.

Uysal, E., `Modernisation Costs, Acquiring New Technology, And Endogenous Long Run Growth', 1996, Unpublished draft copy.

Young, A., `Learning by Doing and the Dynamic Effects of International Trade', 1991, Quarterly Journal of Economics, Vol.151, pp.369-405.

CHAPTER 2

Modernisation Costs, Acquiring New Technology, and Endogenous Long Run Growth

2.1 Introduction

It seems there is a new consensus among economists that sustained long run economic growth is a result of advances in *productive knowledge*. Every period, new inventions hit the boundary of *technological frontiers*, and countries that are able to adopt the new progression, or perhaps older ones that were not yet utilised, enjoy an improvement in their per capita productivity. However, the issues related to the mechanisms that generate new inventions (inventive activity) and enable a country to acquire them (innovative activity) still remain unclear and controversial.

Inventive activity has clearly played a central role in the history of economic development. It is still probably the main engine of world-wide technological progress. However, it appears that this historical importance has rapidly decreased for a small open economy, and continues to decrease, with the introduction of advanced transport and communication networks. All countries today are consuming goods, and producing with methods, equipment, and machinery that were first originated elsewhere in the world. The increasing volume of international trade, direct foreign investment, multi-national companies, tourism, international broadcasting, the internet, world wide railroad and highway networks are all contributing to the diffusion of new discoveries and other types of progress in productive knowledge. Today, the important economic issue is to understand the mechanisms that determine which innovations are taken up and integrated into the actual lines of production of a given economy that face a common world wide technological frontier.

Acquisition of a new invention (innovative activity), by definition, requires a change somewhere in the production process. It could be in the form of introduction of new assembly lines, or new production procedures, or simply new physical machinery and equipment. In many cases, as a result, the entrepreneur who wishes to innovate has to incur additional *modernisation costs* besides the actual physical investment. These modernisation costs might take many different forms: The loss of skills that were developed specifically for the old technology (Schumpeterian explanations of creative destruction), direct or indirect taxes to be paid, the cost of training of labour for the new elements of the production process, and the lower rate of production until the workers develop experience with the new environment are probably among the important ones. Each of these costs effectively reduces the profitability of acquiring new technology, and entrepreneurs may, therefore, choose not to innovate with every new progression in the technological frontier. In return, what determines these additional modernisation costs of innovative activity also determines the degree of technological sophistication and the relative economic welfare in modern-age economies.

This paper concentrates on the adoption of new discoveries into actual production. Following Romer (1990), the world wide technological frontier is formulated by means of an expanding variety of capital inputs. But, the model deviates from his and other R&D derived growth models (see Grossman and Helpman (1991) among many others) in a number of fundamental ways. The attention in that literature mainly focused on the costs of R&D and their effects on economic growth. Every country invents its own products, and once a new product is invented it is innovated with no additional cost. Here inventive activity is overlooked and the technology is assumed to be global. Instead, the paper focuses on post-invention cost benefit analysis of innovative activity. A more similar approach is taken by learning by doing models (see Young (1991) and Young (1994) for an example). However, these models are rather ad-hoc and do not bring an explanation to the fundamental differences across the countries. In this paper all countries have equal access to the same production technology from the same world-wide prices. In other words, if a country (or the entrepreneurship in that country to be precise) wishes to innovate a particular capital input, there are no external barriers to prevent that action. What makes a country different from the others is its preferences and micro structure that determine the cost of acquiring technology endogenously within the system. As a result, private innovation decisions differ, and countries may progress in different ways that may cause permanent disparities in growth rates of incomes. We also show how countries may stuck in no-growth equilibrium and temporary policies may help overcome this problem.

The rest of this chapter is organised as follows; the following section introduces the model, and then examines the steady state equilibrium dynamics. The possibilities for the multiple equilibria in long run growth rates are also analysed in this section. Section 2.3 provides some concluding remarks.

2.2 The Model

The model in this paper is built on two main components. The first component studies the household and business sector behaviours for a given rate of income growth and the exogenous parameters of the model.

Households decide how much to work, to save , to consume, and to invest on education in order to maximise their lifetime utility. Business sector decides how much to produce, to employ labour and physical capital, and whether or not to acquire new technologies. The second component of the model takes the acquirements of new technologies as given and studies the implied rate of economic growth. Although the progress in production technology is assumed to be available world wide and from the same prices, some countries may not acquire all the new progression due to lack of business sector incentives for innovative activity. This, in turn, may generate cross country differentials in the rate of economic growth. General equilibrium is determined to satisfy the internal consistency of two components of the model. Because both higher rate of economic growth and incentives for acquiring new technologies re-enforce each other, multiple equilibria are possible. In such cases, we show that temporary changes in policy parameters may result in permanent shifts in long run growth rate of incomes.

2.2.1 Determination of Human Capital and Innovative Activity

The economy in this model consists of many, identical households, and a competitive business sector. Households live for three periods, and are the suppliers of factors of production that include their effective labour as well as financial capital. The business sector uses the financial capital to buy physical capital inputs from a world-wide market of technology. Hence, all countries in this model potentially face the same technological frontier, though some may choose not to utilise it to the full extent. When a new technology is acquired, there is an additional modernisation cost to be paid before the new capital input can be used in productive activity. Output is then determined by the amount of effective labour, and the quantity and variety of capital inputs that are employed.

The Household Problem

This is an overlapping generations model in which each household lives for three periods; childhood, parenthood, and retirement. All economic decisions are made during parenthood, and individuals who are in their childhood or retirement periods are not involved in productive activity. Individuals in this model have utility during parenthood and retirement defined over consumption. Every period the new generation (children) arrives at the constant rate L, each of which is assigned to an individual from the previous generation (parents). Parents then decide on the educational level of their children, which is costly, and the allocation of their income for consumption during parenthood and retirement. They also give a constant portion of their income to their own parents as a gratitude for their investments in education during their childhood.

Formally, the representative household faces the following maximisation problem;

$$\max_{\{c_{t,P},c_{t,R}\}} \frac{c_{t,P}^{1-\sigma}-1}{1-\sigma}+e^{-\rho}\cdot\frac{c_{t,R}^{1-\sigma}-1}{1-\sigma} \qquad \rho>0, \tag{1}$$

subject to

$$c_{t,P}+k_t = \omega_t(h_t)\cdot\ell_t\cdot(1-\theta) \tag{2}$$

$$c_{t,R} = k_t\cdot e^{r_t}+\omega_{t+1}(h_{t+1})\cdot\ell_{t+1}\cdot\theta \tag{3}$$

$$h_{t+1} = m\cdot(1-\ell_t)^{\gamma} \tag{4}$$

where $c_{t,P}$ and $c_{t,R}$ are the consumption of generation 't' (i.e. people who are in their parenthood during the calendar date t) during parenthood and retirement respectively, h is the human capital that is developed by schooling during childhood, $\omega_t(.)$ is the wage function in period 't' that maps the human capital into the corresponding wages, ℓ is the portion of time during parenthood that is devoted for working, k is the savings during the parenthood for retirement, r is the rate of return on financial savings, and θ is the coefficient of gratefulness to the parents.

The gratuity payments are the only reason for parents to invest on the schooling of the next generation in this model. It can be interpreted as compulsory contributions to pension funds which are used in financing the state schooling system. The present formulation is not crucially important for the main results of the paper, and can be replaced with alternatives that link the welfare of two consecutive generations. Nevertheless, it simplifies the algebra and provides easy interpretation as a simple policy parameter.

Equation (2) is a simple budget constraint, and states that the consumption and savings during parenthood should add up to the net income during that period. Because life ends at the end of the third period, and parents are not directly concerned with the welfare of their children, all income is consumed during retirement without leaving any bequest to the next generation. Leisure does not have any direct utility. Hence, individuals allocate their time, either for productive activity, ℓ, or for the schooling of their children, $(1-\ell)$. As a result, children can build a higher level of human capital, and pay higher gratuity payments to their parents. Equation (4) is the relationship between the parents' effort for the schooling of their children, and the children's human capital level during parenthood.

A typical parent, takes the interest rate (r_t), the wage equations $(\omega_i(.))$, his/her own human capital endowment (h_t), and the strategic behaviour of the next generation (ℓ_{t+1}) as given, and maximise the discounted sum of utilities over his/her life-time by choosing the consumption levels (c_i) and the effort for the schooling of the next generation (1-ℓ_t). The two first order conditions (FOC) that solve this maximisation problem are as follows;

$$\begin{aligned} c_{t,P}^{-\sigma} - e^{r_t-\rho}\left[\left(\omega_t(h_t)\cdot\ell_t\cdot(1-\theta)-c_{t,P}\right)\cdot e^{r_t}+\omega_{t+1}\left(m\cdot(1-\ell_t)^{\gamma}\right)\cdot\ell_{t+1}\cdot\theta\right]^{-\sigma} &= 0 && (5)\\ \omega_t(h_t)\cdot(1-\theta)\cdot e^{r_t}-\left(\frac{\partial}{\partial h_{t+1}}\omega_{t+1}(h_{t+1})\right)\cdot m\cdot\gamma\cdot(1-\ell_t)^{\gamma-1}\cdot\ell_{t+1}\cdot\theta &= 0 && (6)\end{aligned}$$

Equation (5) is the familiar optimum policy for the intertemporal substitution of consumption between future and present (notice that the expression within the square brackets is equal to $c_{t,R}$). Similarly, Equation (6) states the optimum policy rule on allocating time endowment between schooling of children and working hours. It is contingent on the inter-generational welfare link parameter θ. This is because the return from investment on the education of next generation increases with θ thereby giving additional incentives.

Business Sector

The business sector in this model consists of many identical firms that operate in competitive output and input markets. They employ labour and capital inputs to maximise the profit:

$$\max_{\{x,\ell\}} \pi_t = Y_t - W_t - B_t \qquad (3)$$

where Y_t is output, W_t is wage bill, and B_t is rental cost of capital inputs. The production technology is given as follows;

$$Y_t = A\cdot H_t^{\beta}\cdot\sum_{(z,\tau)} x_t(z,\tau)^{\alpha}\cdot e^{\mu\cdot\tau} \qquad (4)$$

Here, A is a scaling parameter, H represents the total stock of effective labour, and x(z,τ) is the quantity of capital input (z,τ). β+α is assumed to be strictly less than one. This production technology is essentially the same as that was first studied by Romer (1990) and widely used in many papers thereafter. The two differences here are decreasing returns to scale and that capital inputs are indexed in two dimensions. The former indexes the type of the capital input (z), and the later is the date at which it was discovered (τ). The type of a capital input is set by its additional cost of innovation when the other factors are kept constant. That is, each capital input requires certain modernisation costs and some portion of that cost is exogenous to the rest of the model. Formally, the modernisation cost of capital input (z,τ) is given by

$$C(z,\tau) = \begin{cases} \sum_j \Gamma(z, h_j, G) \cdot \omega(h_j) \cdot e^{\mu \cdot (\tau - s_z)} & \text{first time innovated} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where j indexes the individuals of that country. Countries who are willing to innovate with the capital input type (z,τ) should incur the additional modernisation cost given in (8). It can be interpreted as the training cost of labour for the new environment that changes with innovative activity. The cost varies among individuals, depending on their human capital endowments. The function $\Gamma(.)$ can be interpreted as the time requirement to develop the necessary skills for the new capital input. It is assumed to be increasing in relation to the difficulty level associated with the capital input (z), to be decreasing in relation to human capital endowment of the person who is willing to learn it (h_j), and to be decreasing in the global environmental parameter (G). The overall cost is also assumed to be exponentially increasing in the degree of relative sophistication compare to the current state of technology (τ-s_z). These assumptions are not unique to this paper and also studied in other work. The effects of education on acquiring of new technology, for example, have been widely studied in the literature (see Nelson and Phelps (1966) for an example). Underlying it is the theory that education enables a person to perform or learning to perform many jobs. The global environment parameter 'G' can also be interpreted as barriers to technology adaptation. This issue is recently considered by Parente and Prescott (1994) in a more conventional growth model.

In similar set-ups, the usual practice is to constrain the innovation of more sophisticated goods contingent on the current state of technology. That is, models are so formulated that countries cannot make jumps to higher quality goods before innovating all the goods of lesser sophistication. Behind it probably is the implicit assumption that the cost of innovation would be too high that it would not be profitable to innovate very sophisticated capital inputs before gaining skills with the simpler ones first. In this model, this intuition is explicitly built in the modernisation cost function in (8). The opportunity cost of innovation is exponentially changing with the degree of separation between the quality level of the capital input (τ) and the current state of the technology (s). However, in principle agents are free to adopt new inventions regardless of whether they innovated everything up to then or not.

Because there are no externalities on the productivity of capital inputs, firms can make their innovation decisions by a simple cost benefit analysis separately for each capital input. If a capital input is innovated, then the optimum quantity is chosen in the usual way by equalising its marginal product to its marginal cost;

$$A \cdot H_t^{\beta} \cdot \alpha \cdot x_t(z,\tau)^{\alpha-1} \cdot e^{\mu \cdot \tau} = r_t \cdot P_{z,\tau} \tag{9}$$

Here $P_{z,\tau}$ is the world wide price of capital input (z,τ). To simplify the analysis, I assume it is proportional to the quality index of the capital input. More precisely, the world wide prices for the capital inputs are given as follows:

$$P_{z,\tau} = e^{\mu \cdot \tau} \tag{10}$$

The optimum quantity of the capital input (z,τ), then, can be calculated independent of its type and the calendar date it was discovered. However, it still varies with the level of human capital and the interest rate:

$$\bar{x}_t(z,\tau) = \bar{x}_t = \left(\frac{\alpha}{r_t} \cdot A \cdot H_t^{\beta}\right)^{1/1-\alpha} \qquad \text{for all } (z,\tau) \qquad (11)$$

Consequently, the total rental cost of employing capital input (z,τ) in its optimum quantity is given by

$$\begin{aligned} B_t(z,\tau) &= \bar{x}_t \cdot e^{\mu\cdot\tau} \cdot r_t \\ &= \alpha \cdot Y_t(z,\tau) \end{aligned} \qquad (12)$$

There are no externalities in the aggregation of human capital in this model. The aggregate human capital can be obtained merely by summing the human capital endowment across the individuals who are in their parenthood:

$$H_t = \sum_{j\in P} h_{j,t} \cdot \ell_{j,t} \qquad (13)$$

The competition in the labour market equalises the component of the wage for time of unit labour of individual j to its marginal product from the production with capital input (k,τ). Hence, the component of person j's wage due to the capital input (k,τ) is given by

$$\omega_t^j(z,\tau) = A \cdot x_t(z,\tau)^{\alpha} \cdot e^{\mu\cdot\tau} \cdot \beta \cdot H_t^{\beta-1} \cdot h_{j,t} \qquad (14)$$

By using (15), the total salary of person j can be obtained by summing his/her wage components over the whole set of capital inputs which are already innovated and employed at their optimum levels.

$$\omega_t^j = h_{j,t} \cdot \sum_{(z,\tau)} A^{\frac{1}{1-\alpha}} \cdot \left(\frac{\alpha}{r_t}\right)^{\frac{\alpha}{1-\alpha}} \cdot \beta \cdot e^{\mu\cdot\tau} \cdot H_t^{\frac{-1+\alpha+\beta}{1-\alpha}} \qquad (15)$$

It should be noted that wages are decreasing in relation to aggregate level of human capital for a given set of capital inputs. This is due to the decreasing returns to scale in production technology. But in equilibrium, as will be shown later in this section, the set of capital inputs is also dependent on the aggregate human capital, and its overall effect on wages is positive.

Similarly, the total wage bill for the labour services with capital input (k,τ) can be found by summing the salaries across the individuals:

$$\begin{aligned} W_t(z,\tau) &= \sum_j \omega_t^j(z,\tau)\cdot \ell_{j,t} \\ &= A\cdot\beta\cdot x_t(z,\tau)^{\alpha}\cdot e^{\mu\cdot\tau}\cdot H_t^{\beta} \\ &= \beta\cdot Y_t(z,\tau) \end{aligned} \tag{17}$$

The Steady State Equilibrium Dynamics

One of the unwelcome features of traditional growth models with infinitely lived households is that assuming financial capital mobility pre-determines the rate of growth of consumption by fixing the interest rate at its world-wide level. Consequently, levels of income and output diverge without bound unless the growth rate of output is also constrained to the same rate. This, however, is not attractive if the objective is to study the growth performance in a world with financial capital mobility. The problem disappears in the models with overlapping generations. In particular, it is possible in such models to have the consumption growth rate of a person different from that of the whole population, since individuals live only for a finite number of periods. As a result, the model builder can assume the mobility of financial capital, thereby equalising the real interest rate across the countries, with no pre-commitment for the rate of growth of the aggregate output or income. Specifically, it is possible in the present model that

$$\frac{c_{t,R}}{c_{t,P}} \neq \frac{c_{t+1,P}}{c_{t,P}}.$$

As in the infinitely lived household case, the consumption growth rate of an individual from parenthood to the retirement can be calculated from the FOC in Equation (5):

$$g_c = \ln(c_{t,P}) - \ln(c_{t,R}) = \frac{r-\rho}{\sigma}. \tag{18}$$

where r is the world wide interest rate on the financial capital. However, Equation (18) is the optimum growth rate of consumption only for an individual and does not impose any restrictions on the growth rate of aggregate variables.

Because there is no bequest to the next generation in the form of financial capital, the only state variable that links generations in this model is the human capital. In Equation (6), parents choose the optimum level of working hours (ℓ_t) based on the choice of working hours of their children (ℓ_{t+1}). Therefore, the *rational* parent has to consider the choice of working hours not only for his/her children, but also for his/her grand-children, since that will affect the choice of working hours for his/her children. In other words, the parent should solve the maximisation problem not only for himself/herself, but also all the generations after him/her.

Let $\{\ell_{t+n}\}_{n=1}^{\infty}$ be the sequence of the optimum working hours for all the grand-children of a parent of generation t. For that sequence to be an equilibrium path for working hours, it should be optimum to choose the corresponding working hour for every generation after t, given that all the other generations are going to choose as specified. Formally, it should be true that the FOC in (6) is satisfied for all the future generations as well as the generation t;

$$\omega_{t+n}\left(m\cdot\left(1-\ell_{t+n-1}\right)^{\gamma}\right)\cdot(1-\theta)\cdot e^{r} + \frac{\partial\omega_{t+n+1}\left(m\cdot\left(1-\ell_{t+n}\right)^{\gamma}\right)}{\partial\ell_{t+n}}\cdot\ell_{t+n+1}\cdot\theta \;=\; 0 \tag{19}$$

for all $n \geq 0$.

Equation (19) is a second order non-linear difference equation in the optimum choices of working hours. The current state of human capital is therefore inadequate to fully describe the equilibrium path in this paper. In fact, for every level of human capital and the choice of working hours for generation t, there exist a sequence of rational expectations[13] that would justify this choice of working hours as the optimum. More formally,

Proposition 1

Let h_t be the state of human capital for generation-t. Then for all $0<h_t<m$ and $0<\ell_t<1$, there exist a sequence of rational expectations for the choice of working hours of future generations, $\{\ell_{t+n}\}_{n=1}^{\infty}$, such that the optimality conditions given in Equation (19) are satisfied for generation-t, and all the generations afterwards, conditional on future generations maintaining the remaining part of the same sequence of rational expectations (i.e., $\{\ell_{t+n}\}_{n=m+1}^{\infty}$ for generation-(t+m), m>0).

Proof

See the Technical Appendix.

In other words, the equilibrium dynamics of human capital are dependent on the expectations in this model. It cannot be fully specified contingent only on the current state of the human capital. Moreover, for any choice of working hours there exist a set of rational expectations such that it justifies the choice as the optimum. One needs to know, therefore,

13 The rationality of expectations here intimates that for any future generation the optimal choice matches exactly with the expectations of the previous generations, conditional on that they also maintain the same expectations for the generations after them.

the state of the human capital for generation-(t-1), as well as for generation-t, to characterise the equilibrium dynamics of the human capital and the set of rational expectations that justifies it.

This kind of multiplicity of equilibria in overlapping generations models is a well-known phenomenon. There have been alternative suggestions on how to choose among the alternative solutions in similar cases. I leave this issue aside for the time being, since the main concern in this paper is the description of the steady state dynamics of incomes rather than the convergence to the steady state growth path. Hence, it suffices for the current purposes to study the possible steady state values of optimum working hours without a description how the economy converges to that point. It should first be cleared what is meant by the steady state level of working hours;

Definition: The level of working hours, ℓ, is said to be a steady state, if the sequence $\{\ell_{t+n} = \ell\}_{n=0}^{\infty}$ is a rational expectations equilibrium.

In other words, given that all the future generations will choose ℓ, and that the previous generation has chosen ℓ as well, it should also be optimum to choose ℓ for the present generation. Clearly, then, any steady state equilibrium value of the optimum working hours is also a fixed point of Equation (19), if it is an interior solution. To solve for the fixed points of (19), one should first characterise the parents' conjecture for the wage equations. It is assumed in this model that parents believe the wages change linearly with the human capital and grow at a constant rate 'g' from one generation to the next. These beliefs are consistent with the wage equation in (16) when the size of the population is infinitely large relative to an individual and the economy is on a steady state equilibrium growth path. It should be noticed, however, for finite size populations there is an aggregate externality on wages from a change in human capital of a typical parent. The effect of this externality goes to zero if individuals are infinitesimal compared to the whole population.

Formally, parents believe (almost rationally) that

$$\omega_t(h) = \omega_t \cdot h \tag{20}$$

$$\omega_{t+1}(h) = \omega_{t+1} \cdot h \quad = \omega_t \cdot e^g \cdot h \tag{21}$$

where ω_t is independent of h. Then, there are two fixed points that solve Equation (19) if $\gamma > 1$, and a single fixed point if $\gamma \leq 1$. In both cases

$$\bar{\ell} = \frac{(1-\theta)\cdot e^r}{(1-\theta)\cdot e^r + \gamma\cdot\theta\cdot e^g} \tag{22}$$

is a fixed point of (19), and also a stable[14] steady state equilibrium for the optimum choice of working hours for some parameter values of the model[15].

Equation (22) constitutes a negative link between the labour supply and the rate of growth of wages for a unit human capital endowment. The intuition behind this result is the fact that it is more profitable to invest in the children's schooling if the growth rate of wages is higher. As a result, parents allocate more of their time endowment for the schooling of the next generation, and less for working.

The steady state value of the optimum human capital can be calculated by substituting Equation (22) into Equation (4):

---

14 An *unstable* steady state equilibrium is a steady state equilibrium such that unless the economy starts at that point, there is no sequence of rational expectations that delivers it as the asymptotic optimum choice for the generations at time infinity. A steady state equilibrium is stable if it is not unstable. The solution in (22) is stable if and only if $\dfrac{(1-\theta)\cdot e^r}{\gamma\cdot\theta\cdot e^g} < 1.$

15 The other fixed point for $\gamma > 1$ is when $\ell=1$. Although it is in some cases a stable steady state equilibrium for the optimum choice of working hours, it is not economically interesting. Therefore, this possibility is ignored for the rest of this paper.

$$\bar{h} = m \cdot \left[\frac{\gamma \cdot \theta e^{g}}{(1-\theta) \cdot e^{r} + \gamma \cdot \theta \cdot e^{g}} \right]^{\gamma} \tag{23}$$

As expected, $\bar{h}$ is increasing in the growth rate of wages[16]. Again, the intuition behind it is that when the growth rate of wages are higher, it becomes more rewarding to invest on the schooling of the next generation. The steady state human capital value in (23) is socially inefficient because of three reasons: First, there is an inter-generation co-ordination failure in the sense that parents take the children's choice of working hours as given. Second, individuals ignore the external effect of their effort on the aggregate human capital level. Third, there is a co-ordination failure among the parents of the same generation since they take the strategies of other parents as given in their maximisation problem. The last two of these three inefficiencies are due to the assumption that the aggregate technology is decreasing returns to scale, and the consequence is an over investment in human capital of the next generation. The effect of the first inefficiency is ambiguous and dependent on the schooling parameter γ. For most parameter values the consequence is a likely over investment on the schooling of the next generation. These observations for the possible inefficiencies, of course, are relevant only when the changes in human capital do not affect the rate of growth of incomes. Therefore, one should revise these observations also for their effects on the endogenous growth rate.

The steady state growth path of the parent's optimum level of financial capital, and the consumption scheme for the parenthood and retirement can be calculated from Equations (2), (3) and (18):

---

16 This positive relationship between the growth rate and the investment on human capital accumulation is analogous to the Samuelson's idea of *acceleration principle* for physical capital investment.

$$c_{t,P} = \omega_t \cdot \bar{h} \cdot \bar{\ell} \cdot \frac{(1-\theta) \cdot e^{r} + \theta \cdot e^{g}}{e^{r} + e^{g_c}} \tag{24}$$

$$c_{t,R} = e^{g_c} \cdot c_{t,P} \tag{25}$$

$$k_t = \omega_t \cdot \bar{h} \cdot \bar{\ell} \cdot \frac{(1-\theta) \cdot e^{g_c} - \theta \cdot e^{g}}{e^{r} + e^{g_c}} \tag{26}$$

It is now clear from Equation (24) that the consumption during parenthood grows at the same rate as wages grow from one generation to another. Yet, the growth rate of consumption from parenthood to retirement in (25) is different as it was established earlier in Equation (18).

The aggregate income, stock of financial capital and consumption can be found by summing individual values over the population:

$$I_t = \sum_{P} \omega_t \cdot \bar{h} \cdot \bar{\ell} + \sum_{R} k_{t-1} \cdot e^{r} \tag{27}$$

$$K_t = \sum_{P} k_t \tag{28}$$

$$C_t = \sum_{P} c_{t,P} + \sum_{R} c_{t-1,R} \tag{29}$$

where, I is the aggregate income, K is the aggregate stock of money capital, and C is the aggregate consumption. By substituting the individual choices in (24), (25) and (26) into the equations (27), (28) and (29), one can re-formulate the aggregate values in terms of the wage level:

$$I_t = \omega_t \cdot \bar{h} \cdot \bar{\ell} \cdot L \cdot \frac{(2-\theta) \cdot e^{g_c} + e^{r} - \theta \cdot e^{g}}{e^{r} + e^{g_c}} \cdot e^{r} \tag{30}$$

$$K_t = \omega_t \cdot \bar{h} \cdot \bar{\ell} \cdot L \cdot \frac{(1-\theta) \cdot e^{g_c} - \theta \cdot e^{g}}{e^{r} + e^{g_c}} \tag{31}$$

$$C_t = \omega_t \cdot \bar{h} \cdot \bar{\ell} \cdot L \cdot \frac{(1-\theta) \cdot e^{r} + \theta \cdot e^{g}}{e^{r} + e^{g_c}} \cdot \left(e^{g_c - g} + 1\right) \tag{32}$$

Similarly, the steady state level of the aggregate factor inputs can be calculated from Equations (11) and (13):

$$\overline{H} = \overline{h} \cdot \overline{\ell} \cdot L \tag{33}$$

$$\overline{x} = \left(\frac{\alpha \cdot A}{r}\right)^{1/1-\alpha} \cdot \left(\overline{h} \cdot \overline{\ell} \cdot L\right)^{\beta/1-\alpha} \tag{34}$$

As a result, for a given capital input that is already innovated, one can derive the corresponding component of the aggregate profit at the steady state equilibrium:

$$\pi_t(z,\tau) \;=\; \overline{\pi}(\tau) \;=\; (1-\alpha-\beta)\cdot A \cdot \overline{H}^{\beta} \cdot \overline{x}^{\alpha} \cdot e^{\mu\cdot\tau} \tag{35}$$

It remains to solve for the range of capital inputs that are already innovated at a given date. At this stage, one needs to specify the mechanism that determines how the innovative decisions are made. In this paper, both the benefits and the costs of innovative activity are characterised at the aggregate level. In reality however, the decisions are made at the firm level operating in competitive markets. In this case, the excludebility of benefits from innovative activity is crucially important for firms to make their innovation decisions (see Romer(1990)).

Here, it is assumed, for simplicity, that a capital input is innovated whenever the asset value of future aggregate profits exceeds the rent flow of the modernisation costs. Formally, the capital input (z,τ) is innovated if and only if

$$\pi_t(z,\tau) \;\geq\; C_t(z,\tau) \cdot e^{r}.$$

It is possible in this set up that a capital input which was not profitable enough for innovation in one period, can be suitable for innovation in another[17]. On the other hand, the following proposition states that capital inputs of the same type are innovated all at the same time. That is, if a capital input is suitable for innovation, so are all the other capital inputs of the same type.

---

17 Once a capital input is innovated, there is no reason to dis-innovate since the modernisation cost is a sunk cost in its nature. Hence, the possibility of dis-innovation is disregarded.

Proposition 2

Let $C_t(z,\tau)$, and $\pi_t(z,\tau)$ be the modernisation cost, and the aggregate profit of the capital input (z,τ) at date t. Then

$$\pi_t(z,\tau) \geq C_t(z,\tau)\cdot e^r \quad \text{iff} \quad \pi_t(z,\tau^*) \geq C_t(z,\tau^*)\cdot e^r \quad \forall \tau,\tau^* \leq t.$$

Proof

See the Technical Appendix.

In other words, Proposition 2 states that countries operate in their *innovation possibility frontier* for all the types of capital inputs that are suitable for innovation. Accordingly, there is no gap between the countries of this model and the technological frontier, as is usually the case in conventional settings. Countries may differ, however, with respect to the type of capital inputs that are suitable for innovation. In particular, there is a critical type of capital input such that all the capital inputs that are below or equal to this critical type are innovated and the others are not. Formally,

Proposition 3

Assume that the capital input types are randomly distributed on a set of finite elements Ω. Furthermore, assume that the economy is on an equilibrium growth path where the working hours evolve according to a rational expectations sequence $\{\ell_{t+n}\}_{n=1}^{\infty}$ that is not periodic. Let $C_t(0,\tau) = 0$ $\forall$ $t,\tau>0$. Then, $\exists$ $z^* \in \Omega$, and T, such that, $\forall$ $t > T$

$$\pi_t(z,\tau) \geq C_t(z,\tau)\cdot e^r \quad \text{iff} \quad z \leq z^*.$$

Proof

See the Technical Appendix.

Propositions 2 and 3 characterise the steady state equilibrium dynamics of the capital inputs that are innovated for a given set of model parameters. Accordingly, the countries that are endowed with a higher level of human capital innovate a wider range of capital inputs. This is because the higher level of human capital increases the post-innovation profits in these countries, while decreasing the cost of innovation at the same time.

Given the set of capital inputs is determined as described in Propositions 2 and 3, one can calculate the wage level and the aggregate output:

$$\omega_t = \overline{h}^{-\beta/1-\alpha} \cdot \left(\frac{A \cdot \alpha^{\alpha}}{r^{\alpha}}\right)^{1/1-\alpha} \cdot \frac{\beta}{e^{\mu}-1} \cdot \left(\sum_{z \le z^*} e^{\mu \cdot s_z(t)} - Z^*\right) \tag{36}$$

$$Y_t = A \cdot \overline{H}^{\beta} \cdot \overline{x}^{\alpha} \cdot \frac{1}{e^{\mu}-1} \cdot \left(\sum_{z \le z^*} e^{\mu \cdot s_z(t)} - Z^*\right) \tag{37}$$

where Z^* is the number of capital input types that are suitable for innovation. $s_z(t)$ is the latest capital input of type z that is already innovated. Because, all countries in this model operate in their innovation possibility frontier, $s_z(t)$ also represents the world wide technological frontier for type- z capital inputs.

2.2.2 Determination of Long Run Growth

The main issue in this paper is to examine why and how countries that are subject to the same technological frontier can grow at different rates, when also the determinants of economic growth are endogenously determined. The recent endogenous growth theory is mostly concentrated on the effects of differentials in household's savings behaviour on the endogenous growth performance (see Rebelo(1991) for an example). In these models, countries can grow faster than others if they save more. This is because, they assume the effects of savings on the accumulation of augmentable factors are independent of the relative state of the economy compare to the rest of the world. This contradicts with the traditional view that the returns on savings are higher if the country is technologically lagging. It is hard to imagine why the rates of return on savings should be the same between two countries, if one is technologically advanced and has to invent before it can innovate, and the other is technologically lagging and can imitate the other. Once the endogenous growth models are modified for this view, the results reverse to an homogenous long run growth rate across all countries, just as it is in the traditional exogenous growth models (see Lucas (1993) for an example). On the other hand, empirical regularities seem to suggest that there are differences in the long run growth rates of per capita incomes across the countries. Then, it requires a theory of endogenous growth while maintaining the more realistic assumption that there is a high degree of technological diffusion across countries.

So far, we have established the behaviour of households on human capital investment for a given rate of economic growth, and the firms' behaviour on innovative activity for a given level of human capital. Growth in this model is generated by innovating new capital inputs. Every period new inventions arrive to the world wide technological frontier, and countries choose the ones that are suitable for innovation in their environments. Let 's' be the rate at which the new inventions arrive, and $\phi(z)$ is the probability that a new arrival is of type z. Then, the total number of capital inputs of type z that are already invented at date t is given by

$$S_z(t) = \phi(z) \cdot s \cdot t \tag{38}$$

Then, the aggregate output in (36) can be approximated, for large t, by

$$Y_t \cong A \cdot \overline{H}^{\beta} \cdot \overline{x}^{\alpha} \cdot \frac{1}{e^{\mu} - 1} \cdot s^* \cdot e^{\mu \cdot \phi^* \cdot s \cdot t} \tag{39}$$

where

$$\phi^* \equiv \max_{z \le z^*} \{\phi(z)\} \tag{40}$$

and

$$s^* \equiv \ \#\{z \mid \phi(z) = \phi^*\} \tag{41}$$

Here '#' is a set operator that returns the number of elements in its argument. Similarly, the asymptotic growth rate (i.e., t is large) of aggregate output can be approximated by

$$g^* = \frac{\dot{Y}}{Y} \cong \mu \cdot \phi^* \cdot s \tag{42}$$

Equation (42) establishes the second link between the steady state level of human capital and the endogenous rate of growth. In particular, modernisation costs and post-innovation profits are determined as a function of the aggregate level of human capital endowment. This, in turn, sets the highest capital input type that is suitable for innovation. The growth rate of aggregate output is asymptotically equivalent to the maximum of the growth rates of output among the capital input types that are suitable to be employed in production.

The steady state equilibrium in this general equilibrium model is the pair of human capital level and an asymptotic growth rate, $(\overline{h}^*, \overline{g}^*)$, that solves Equations (23) and (42). Because both equations establish a positive link between the two variables, the uniqueness of the equilibrium is not guaranteed. As a result, even countries with similar micro structures can exhibit differences in their growth rates of aggregate output.

Figure-2.1 illustrates the possibility of multiple equilibria for a given set of model parameters. The HA curve plots the household choice of human capital investment at the steady state against the growth rate in wages (i.e., Equation (23)). As the growth rate in wages increases, it becomes more attractive to invest in the schooling of the next generation rather than devoting time to working. The result is a higher steady state value of human capital. The broken line (EG curve) plots the inverse relationship between the growth rate of incomes and the level of human capital in Equation (42). As human capital rises, the set of suitable capital input types expands. At some critical human capital levels, a capital input type that grows faster than the preceding ones (i.e. a capital type $\underline{\underline{z}}$ such that $\phi(\underline{\underline{z}}) > \phi(z)$ for all $z < \underline{\underline{z}}$) becomes suitable for innovation. These are the points where EG curve is discontinuous since a marginal change in the human capital endowment raises the rate of steady state growth rate with a discrete jump. The '*' indicates a general equilibrium point that are all locally stable.

Comparative Statics

The countries in this model differ with respect to their micro structure parameters: The inter-generation transfer of wealth (θ), and the global environment (G) that affects modernisation costs. The former determines the household's behaviour concerning the accumulation of human capital, and the later influence the business sector decisions concerning innovative activity.

The changes in θ shift the HA curve. The partial derivative of the human capital in Equation (23) with respect to θ is always greater than zero. In turn, this implies that an increase (decrease) in the value of θ increases (decreases) the steady state human capital level for the same growth rate in wages. This translates to a rightward (leftward) shift of the HA curve. It is illustrated in Figure-2.2. Similarly, sufficient changes in G shift the EG curve. This is because modernisation costs decrease (increase) with an increase (decrease) in G, and this might cause an increase (decrease) in the critical capital input type for the same level of human capital endowment. It is also possible, however, that small changes in G have *no effect* on the economy since they may not be enough to change the critical capital input type. The effect of a sufficient size increase in G is illustrated in Figure-2.3.

Because there are multiple equilibria, temporary changes in parameter values of this model can have permanent effects by shifting the economy between different equilibrium points. Consider, for example, an economy that is in the steady state equilibrium point B of Figure-2.3. An increase in G shifts the EG curve to the left without having any effect on the HA curve. The economy is now in the basin of attraction of the steady state equilibrium point A. After a sufficient number of periods, G can be decreased to its original level without the economy going back to its original equilibrium point B.

Recently, Parante and Prescott (1994) and Lucas(1993) show how traditional growth models explain the *miracles* of east Asian economies. In this model the miracles are possible with sufficient changes in the micro structure parameters. When the steady state equilibrium moves from one point to another with a higher growth rate (i.e., from B to A in Figure-2.3 for example), the country suddenly finds itself far behind its new innovation possibility frontier. The result can be a miraculous rate of growth until the economy converges to its new steady state equilibrium.

From this point of view, the model establishes the link between the two previous explanations for *miracle making*. The modernisation costs of this paper are directly analogous to the barriers to the technology adaptation in Parante and Prescott (1994). It could also be interpreted as the critical productivity level ξ in Lucas (1993) at which the new goods are innovated. Then, one could argue that the miracles are initiated by the reasons as described in Parante and Prescott (1994), and Lucas (1993) provides the transition dynamics until the countries reach the new steady state equilibrium.

Exogenous Growth

The neo-classical approach treats growth as an exogenous phenomenon (Solow (1956)). This, however, does not necessarily imply that it is *manna from heaven*. Just as competitive prices are exogenous for a firm but endogenously determined at the market level, economic growth may be taken as exogenous for a given country although it might be endogenously determined for the world as a whole. In fact, this is more likely to be the case with a high degree of technological diffusion across countries, and that all countries are small compared to the world.

The same arguments apply to the model in this paper that treats the arrival of new inventions as an exogenous process. Here, the difference is that countries can attain different growth rates depending on their micro structures and human capital levels. This gives the endogenous nature of economic growth despite a world wide technological frontier to which all countries have equal access.

The assumption that the technological frontier is mutual among countries also provides many ways of re-producing the neo-classical exogenous growth model in this paper. To achieve this in an easy way, assume that $\beta=0$, and that capital inputs are uniformly distributed with respect to their types (i.e., $\phi(z)=\phi \ \forall \ z\in\Omega$). Furthermore, assume that there are no modernisation costs, and that all new inventions are immediately innovated. Let κ_t be the worth of aggregate physical capital stock used in production at time t;

$$\kappa_t = \sum_{(z,\tau)} x_t(z,\tau)\cdot e^{\mu\cdot\tau} \tag{43}$$

The steady state equilibrium values for the aggregate output and physical capital stock of this model can be approximated by;

$$Y_t \cong \frac{A}{e^{\mu}-1}\cdot s^* \cdot \overline{x}^{\alpha} \cdot e^{\mu\cdot\phi\cdot s\cdot t} \tag{44}$$

$$\kappa_t \cong \frac{1}{e^{\mu}-1}\cdot s^* \cdot \overline{x} \cdot e^{\mu\cdot\phi\cdot s\cdot t} \tag{45}$$

where s^* is the number of capital input types that are suitable for innovation, s is the rate at which the new inventions arrive, and ϕ is the probability that a capital input is type z as before. The crucial difference here is that ϕ is the same for all capital types. Then, merely by substituting (45) into (44), one can obtain the familiar neo-classical production function with exogenous technological progress;

$$Y_t = A_t \cdot \kappa_t^{\alpha} \tag{46}$$

The technology parameter A_t in (46) grows at the exogenous rate

$$g_a = \dot{A} / A = \mu(1\text{-}\alpha)s \tag{47}$$

which is independent of the other model variables. There are, however, two major differences: First, the financial capital mobility here replaces the Solow assumption, so that countries operate in their steady state equilibrium at all times. Second, there is no technological progress unless there is investment in capital inputs. The latter differs from the previous interpretation of the technology parameter A_t as learning by doing, since only then growth in total productivity was thought possible with no change in the aggregate physical capital (see Arrow (1969) for an example).

Empirical Support

This paper provides a long run growth model in which all countries are subject to the same technological frontier. As with any other theory of economic growth, it is constructed to satisfy the empirical regularities of balanced economic growth as they are posited by Kaldor (1961). The model, however, differs from others in its implications for the dynamics of incomes in a cross section of countries. First of all, the model predicts divergence of countries into clubs where each club differentiates according to the growth rate of incomes. This prediction is consistent with recent empirical findings by Quah (1993) who reports a divergence of incomes into a *two-camp-world;* the rich and the poor. Another implication of the model is that it predicts bilateral co-integration of incomes across the countries[18]. This is consistent with the findings of Bernard and Durlauf (1991) who find many bilateral co-integrations across a set of 15 industrialised countries.

This model also illustrates why and how non-productive factors may be correlated with incomes in the cross section. In Equation (38), aggregate income is dependent on s* which is itself a function of many variables that determine modernisation costs in a country. Indeed, including the traditional measures of productive inputs (i.e., human capital as measured with the number of years of schooling and physical capital as defined in (43)) does not remove the cross sectional correlation between the level of per-capita incomes and non-productive factors that affect modernisation costs. Recently Barro and Sallai-Martin (1992), Barro

[18] The neo-classical growth model predicts a bilateral co-integration of incomes with a (1,-1) as the co-integrating vector, whereas, most other endogenous growth models predict no co-integrating relationship.

and Lee(1993), and Barro(1994) have studied the impact of various non-productive factors on aggregate productivity across the countries. They find some of these variables (political stability for example) significantly correlated with level of per-capita incomes across the countries.

2.3 Conclusions

Traditional exogenous growth models assume a global technological frontier, and that all countries grow at the same rate as the technological frontier does. Therefore, they implicitly assume a perfect diffusion of technological knowledge across countries (perhaps with some degree of delay). On the contrary, recent endogenous growth literature takes the exact opposite position in which all countries generate their own economic growth and technological frontiers. In these models the effects of international interactions on the technological diffusion across countries are completely neglected. They do not provide any explanation as to why the state of technological knowledge in the rest of the world does not affect the accumulation of technological knowledge in a small open economy.

This chapter seeks to analyse how countries may grow at different rates in the long run even if they face the same technological frontier. It provides an overlapping generations model in which the technological frontier expands exogenously with the introduction of new capital inputs. The capital inputs are marketed all around the world from the same world wide prices. Therefore, all countries potentially face the same technological frontier, though some may not choose to utilise it to the full extent depending on country specific factors. This is because each innovation with a new capital input requires some modernisation costs, and it may not be profitable in some countries to adopt every new technology. Capital inputs differ in their relative modernisation costs. That is, for a given environment some capital inputs are cheaper to innovate than others. This also raises the possibility of multiple equilibria for a given environment, and the cross country differentials in long run growth rates. In this circumstance, countries can switch between two steady state equilibria permanently by temporary policies. It also provides an explanation to the recent East Asian miracles.

The predictions of the model are consistent with the empirical work in Chaters 1 and 3 of this thesis, as well as previous empirical research. In particular, contrary to other endogenous growth models, it shows the possibility of bilateral co-integration of per capita incomes across countries without imposing the restrictions that the exogenous growth models do. This prediction is confirmed in the empirical work by Bernard and Durlauf (1991) and Chapter 3 of this thesis. Besides, the model also provides an explanation as to why non-productive factors, as reported recently by Barro and his co-authors and Chapter 1 in this thesis, can be correlated with the growth rate and the level of per capita incomes across the countries.

Figure - 2.1

Simultaneous Determination of Endogenous Growth Rate and Human Capital Endowment

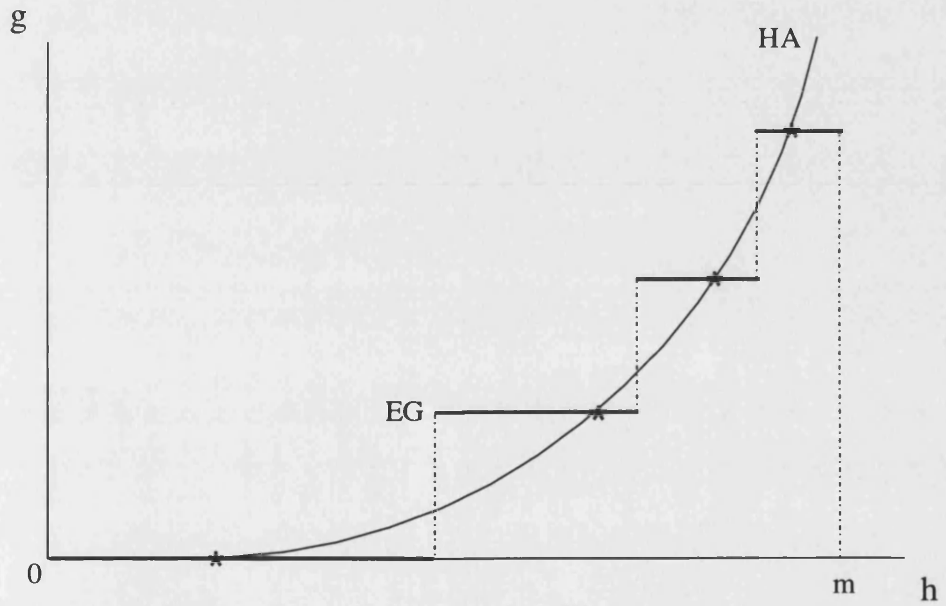


Figure - 2.2

The Effect of An Increase in Gratuity Parameter: θ

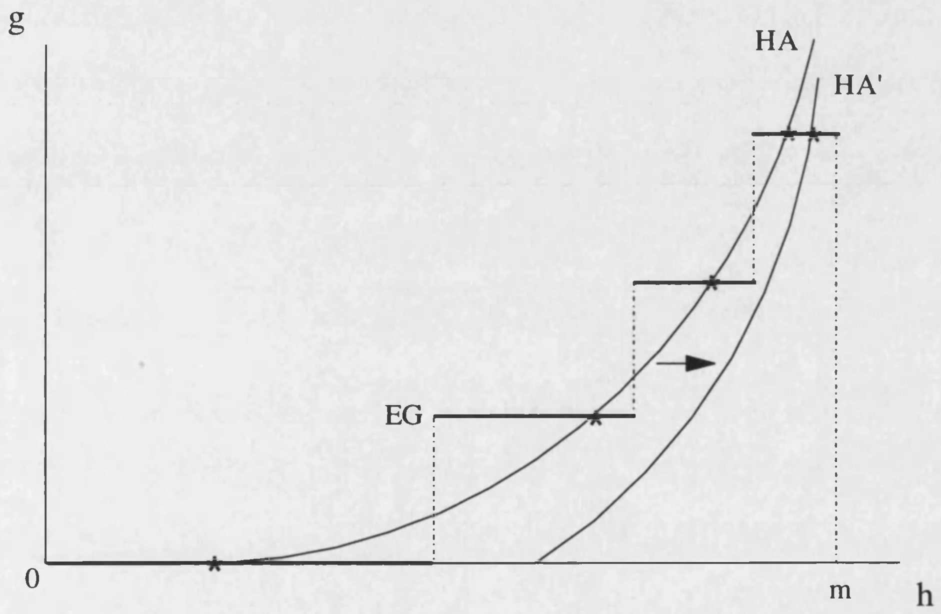
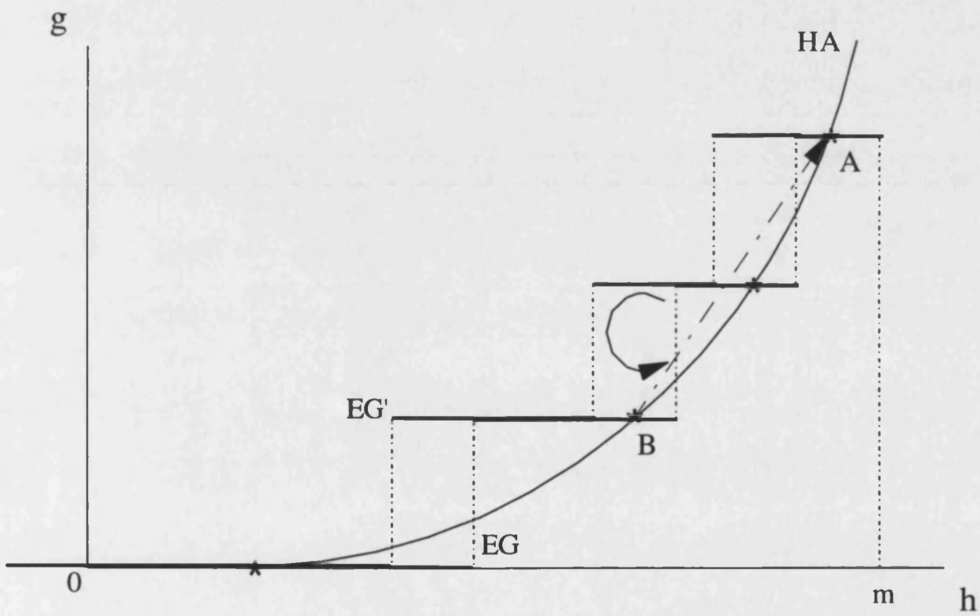


Figure - 2.3

The Effect of An Increase in Environmental Parameter: G



2.4 Technical Appendix

Proof of Proposition 1

Let $h_t \in [0,m]$ be the state of present human capital level. Then,

$$\ell_{t-1} = 1 - \left(\frac{h_t}{m}\right)^{1/\gamma} \tag{A1}$$

Choose an arbitrary $\ell_t \in [0,1]$. and generate the sequence

$$\begin{aligned}
\bar{\ell}_{t-1} &= \ell_{t-1} \\
\bar{\ell}_{t} &= \ell_{t} \\
\bar{\ell}_{t+n} &= \begin{cases} D_{t+n} & \text{if} \quad 0 < D_{t+n} < 1 \\ 1 & \text{if} \quad D_{t+n} \geq 1 \\ 0 & \text{if} \quad D_{t+n} \leq 0 \end{cases} \qquad \forall \text{ integer } n > 0
\end{aligned}$$

where

$$D_{t+n} = \frac{\omega_{t+n-1}\left(m \cdot \left(1-\bar{\ell}_{t+n-2}\right)^{\gamma}\right) \cdot (1-\theta) \cdot e^{r}}{\dfrac{\partial \omega_{t+n}\left(m \cdot \left(1-\bar{\ell}_{t+n-1}\right)^{\gamma}\right)}{\partial \bar{\ell}_{t+n-1}} \cdot \theta}$$

Then, the sequence $\left\{\bar{\ell}_{t+n}\right\}_{n=0}^{\infty}$ is rational expectations equilibrium, since it solves (19) $\forall$ t+n, such that $n \geq 0$.

Proof of Proposition 2

Assume that

$$\pi_t(z,\tau) \geq C_t(z,\tau)e^{r}$$

for some $\tau \leq t$ where t is the current calendar time. Then,

$$\sum_j \Gamma(z, h_t^j, G_t) \cdot \omega_t^j(h_t^j) \cdot e^{r-\mu \cdot s} \quad \leq \quad (1-\alpha-\beta) \cdot A \cdot H_t^{\beta} \cdot \bar{x}_t^{\alpha}$$

which implies

$$\pi_t(z,\tau^*) \geq C_t(z,\tau^*)e^{r}$$

for all $\tau^* \leq t$.

Proof of Proposition 3

At any calendar time t,

$$\pi_t(z,\tau) \geq C_t(z,\tau)e^r$$

if and only if

$$\sum_j \Gamma(z, h_t^j, G_t) \cdot \omega_t^j(h_t^j) \cdot e^{r-\mu \cdot s} \quad \leq \quad (1-\alpha-\beta) \cdot A \cdot H_t^\beta \cdot \overline{x}_t^\alpha \qquad \text{(A1)}$$

The left hand side of (A2) is increasing in z and equal to zero at z=0. The right hand side is strictly positive and independent of the type. Hence, there exist $z_t^* \in \Omega$ such that

$$\pi_t(z,\tau) \geq C_t(z,\tau)e^r \text{ if and only if } z \leq z_t^*.$$

Because, it is assumed that the economy is on an equilibrium growth path where the optimum working hours evolve according to a non-periodic rational expectations equilibrium sequence $\{\ell_{t+n}\}_{n=0}^{\infty}$, the following limits exist;

$$\lim_{t\to\infty} H_t \to \overline{H}$$
$$\lim_{t\to\infty} \overline{x}_t \to \overline{x}$$

This implies, the limit

$$\lim_{t\to\infty} z_t^* \to z^*$$

exist and well defined.

2.5 References

Acemoglu, D., `Market Imperfections, Investment in Skills and Growth', 1993, CEPR Conference Paper.

Arrow, K.J., `Classificatory Notes On The Production And Transmission of Technological Knowledge', 1969, American Economic Association: Papers And Proceedings, Vol .90, pp.29-35.

Barro, R. J., `Democracy and Growth', 1994, unpublished seminar paper.

Barro, R. J., and Lee, J.W., `Losers and Winners in Economic Growth', 1993, National Bureau of Economic Research, Working Paper, No.4341.

Barro, R. J., and Sala-i Martin, X., `Convergence', 1992a, Journal of Political Economy, Vol. 100, pp.223-51.

Bernard, A. B., and Durlauf S. N., `Convergence of International Output Movements', 1991, National Bureau of Economic Research, Working Paper No. 3717.

Kaldor, N., `Capital Accumulation and Economic Growth', 1961, In F. Lutz (ed.), The Theory of Capital, London: Macmillan.

Lucas, R. E., `On the Mechanics of Economic Development', 1988, Journal of Monetary Economics, Vol. 22, pp.3-22.

-------, `Making Miracle', 1993, Econometrica, Vol. 61, pp.251-72.

Mankiw, N. W., Romer D., and Weil D. N., `A Contribution to the Empirics of Economic Growth', 1992, Quarterly Journal of Economics, Vol. 107, pp.407-432.

Nelson, R. R., and Phelps E. S., `Investment in Humans, Technological Diffusion, and Economic Growth', 1966, American Economic Review, Papers and Proceedings, Vol. 56, pp.69-75.

Parente, S. L., Prescott, E. C., `Technology Adoption and Growth', 1991, NBER Working Paper:3733.

------, `Barriers to Technology Adoption and Development', 1994, Journal of Political Economy, Vol. 102, pp.298-321.

Quah, D., `Empirical cross-section dynamics in economic growth', 1993b, European Economic Review, Vol. 37, pp.426-34.

Rebelo, S., `Long-Run Policy Analysis and Long-Run Growth', 1991, Journal of Political Economy, Vol. 99, pp.500-21.

Rivera-Batiz, L., Romer, P.M., `Economic Integration and Endogenous Growth', 1991, Quarterly Journal of Economics, Vol.106, pp.531-55.

Romer, P. M., `Increasing Returns and Long-Run Growth', 1986, Journal of Political Economy, Vol. 94, pp.1002-37.

------, `Endogenous Technological Change', 1990, Journal of Political Economy, Vol. 98, pp.S71-S102.

Solow, R. M., `A Contribution to the Theory of Economic Growth', 1956 Quarterly Journal of Economics, Vol. 70, pp.65-94.

------, `Perspectives on Growth Theory', 1994, The Journal of Economic Perspectives, Vol. 8, pp.45-54.

Young, A., `Learning by Doing and the Dynamic Effects of International Trade', 1991, Quarterly Journal of Economics, Vol.106, pp.369-405.

CHAPTER 3

Permanent Impacts of World War II and 20th Century European Growth

3.1 Introduction

Following Nelson and Plosser (1982), it is now widely accepted that logarithm of per capita incomes is better described as a unit root process (as opposed to being stationary around a linear trend). This has raised two other important questions: Firstly, if incomes are non-stationary in the stochastic sense, then are there any stable long run relationship between the per capita incomes across countries? Secondly, how do we separate the permanent shocks to incomes from those that are temporary? Both questions are, in fact, closely related to one another. A stable long run relationship between incomes of two countries is possible only if the permanent innovations are the same (perhaps with different long run impacts). Similarly, if there exist a stable long run relationship between incomes across countries, then this should be informative in separating permanent shocks from those that are temporary.

The real business cycles hypothesis claims that most of the fluctuations in incomes are due to the technological progress, and that the transitional dynamics, if any, comes from an intrinsic persistence such as consumption smoothing and intertemporal substitution of labour. If confirmed by the data, this would imply that technological progress is what lies under the historical business cycles, and the other forces are relatively less important. It is this implication of the real business cycles theory what motivated many macro-econometricians to study alternative permanent-transitory decomposition of incomes.

Permanent disturbances are usually considered as a result technological progress and the other shocks as temporary.

A stable long run relationship in per capita incomes across countries is implied by the neo-classical growth models. These models consider technological progress as an exogenous and world-wide phenomenon. Therefore, in the long run, all countries are subject to the same permanent technology shocks. At this point it is easier to make the connection between the two questions addressed in this chapter. If neo-classical growth models are good approximations to real life economies, then the cross-country differences in incomes should be stable around a constant. In this case, the unit root process that is driving the permanent component of incomes could be interpreted as the world-wide technological progress. Furthermore, this permanent component could be expected as *causally prior* to the idiosyncratic temporary shocks in small open economies. This is because slight changes in incomes of small economies are not likely to influence the world-wide technological progress. In return, this should be informative in separating the permanent components of incomes by using cross-country observations.

Both questions are tackled in the recent empirical literature but not jointly. Many authors, including Beveridge and Nelson (1981), Harvey (1985), Watson (1986), Campbell and Mankiw (1987), Clark (1987), Evans (1989), Cochrane (1988), Blanchard and Quah (1989), King, Stock, Plosser, and Watson (1991), Lippi and Reichlin (1994) addressed to the second question, and proposed alternative ways of decomposing GDP into its permanent and transitory components. Among them, Blanchard and Quah (1989), and King et al. (1991) have adopted vector autoregression (VAR) models. They both use very similar identifying assumptions by imposing restrictions on the long run effects of the system disturbances. They differ only in their choice of additional variables, and that King et al. (1991) considers the more general case with cointegrated variables. In particular, Blanchard and Quah (1989) uses the unemployment rate in a bivariate VAR model, where as King et al. study a larger system with consumption, investment, nominal money supply, price deflator and the interest rate. These two approaches are closely related to the identification method used in this paper, and will be discussed in more detail later in the chapter.

The question of whether a stable long run relationship exists between the incomes across countries has also attracted considerable attention in the recent empirical literature (see Barro and Salla-i Martin (1992), Barro (1991), Quah (1993), Bernard and Durlauf (1995), and Uysal (1995) among others). Underlying it is the sharp claims by neo-classical exogenous growth models that per capita incomes converge in the long run if one controls for several country specific factors. Among them Bernard and Durlauf (1995) is closest to this paper. They study the cointegration properties of per capita incomes across 15 industrialised countries, and interpret the lack of bilateral cointegration as evidence against the convergence hypothesis of exogenous growth models.

This chapter attempts to answer both these questions simultaneously by studying 20th Century economic growth across six European countries. In Bernard and Durlauf (1995), it is shown that there is more than one unit root process that is driving incomes in these six countries. They interpret this as evidence in favour of endogenous growth models, and argue that it contradicts to the implications of neo-classical exogenous growth models. We start from the point where their work ended, and try to identify the distinct unit root components for each country. We find that World-War-II (WW2) is the main source of differences in permanent components. Then, we include dummies to control for the permanent *economic* consequences of this major event of the 20th Century. Our results indicate that the income differences across the six European countries has been stable during the period if one controls for WW2. This is consistent with the implications of augmented neo-classical growth models where the major re-structuring events, such as wars, are explicitly taken into consideration.

We also study the permanent-transitory decomposition of incomes for all six countries. In the light of real business cycles theory, the permanent component is interpreted as the world-wide technological progress. We find that the impact of a typical technology shock is increasing in time, and that the transition dynamics of permanent shocks contribute about 25% to temporary fluctuations. The dynamic patterns are consistent with the learning by doing and/or inter-industry technological diffusion ideas suggested by Griliches (1957), Jovanovic and Lach (1990), and Young (1991). Furthermore, a positive technology shock increases incomes more in the short run than in the long run. This could be due to a temporary boost in investment demand to adapt the new technology into the actual lines of production. We cannot, however, find any evidence in favour of the Schumpeterian ideas of creative-destruction[19].

The rest of this chapter is organised as follows: Section 3.2 is a theoretical discussion where we study alternative permanent-

[19] See Aghion and Howitt (1993) for a theoretical discussion.

transitory decomposition techniques in cointegrated systems. The identification approach that is proposed here is very close to that of King et al. (1991), and imposes restrictions on long run effects of the disturbances. The main difference is that we do not fully identify the disturbances and instead work with *composite shocks* to the system. In particular, we assume that temporary and permanent disturbances are orthogonal to each other, and obtain a set of permanent components only based on this assumption. Section 3.3 contains the empirical work where we concentrate on per capita incomes of six European countries. Section 3.4 includes some concluding remarks.

3.2 Permanent And Transitory Components In Cointegrated Systems

In this section, we study the decomposition of variables into their permanent and transitory components in cointegrated systems. The identification method developed here can be viewed as a generalised version of King et al. (1991).

Let x_t be an $m \times 1$ vector of random variables that are integrated of order one. Then, there exists a unique Wold decomposition for the stationary first differences:

$$\Delta x_t = \sum_{j=0}^{\infty} C_j \, \varepsilon_{t-j} \tag{2.1}$$

where

$$C_0 = I_m ,$$

$$E(\varepsilon_t \varepsilon_s') = \begin{cases} \Omega_\varepsilon & \text{if } t = s \\ 0 & \text{otherwise} \end{cases}$$

As it is shown formally in Granger and Engle (1987), cointegration implies that

$$C = \sum_{j=0}^{\infty} C_j$$

is deficient rank which is given by the number of distinct unit root processes that are driving the system (i.e., rank(C) = m-r where r<m is a positive integer and referred as the cointegrating rank). The Wold decomposition representation in (2.1) is uniquely identified only with the condition that C_0 is the identity matrix of appropriate order. Once this rather arbitrary condition is relaxed, there are many other representation formats that describe the same system. Formally, for any non-singular matrix R, one can transform the system by

$$\Delta x_t = \sum_{j=0}^{\infty} D_j^R \xi_{t-j}^R \tag{2.2}$$

where

$$D_j^R = C_j R$$

$$\xi_t^R = R^{-1} \varepsilon_t.$$

The matrix R is referred as the identification matrix and there are no strict rules for choosing one. It is, however, widely assumed that RR′ is equal to the covariance matrix of the Wold-decomposition residuals, so that the new system disturbances are orthogonal to each other. Depending on the context, this assumption can be quite reasonable and help to interpret the system disturbances. It is, however, not sufficient to identify R uniquely, and additional restrictions are generally needed. In practice, these restrictions are sometimes inspired by an underlying economic theory but often chosen quite arbitrarily. One early popular method has been to choose the unique Choleski decomposition of Ω_ε. However, it suffers from two major shortfalls; first the restrictions are quite arbitrary and therefore the system disturbances are not interpretable, and second, the analysis is sensitive to the ordering of the variables.

Many economists argue that additional identifying constraints should be imposed by using long run properties of an appropriate economic model. This is mostly motivated by two main reasons: First, short run constraints are very sensitive on the choice of period length, and second, most economic theories have robust implications only on the long run. Blanchard and Quah (1989) provide an identification method that uses such long run constraints. They choose the identification matrix by imposing restrictions on the ultimate effects of the disturbances so that a triangular long run impulse matrix is obtained. Formally, the identification matrix is chosen such that

$$\begin{aligned} S &= \sum_{j=0}^{\infty} D_j^R \\ &= C\,R \end{aligned}$$

is the new triangular long run impulse matrix, and given by the unique Choleski decomposition of $C\Omega_\varepsilon C'$. In many cases, the system disturbances that are identified with this method can have easy interpretations directly from economic theory[20]. It could be considered as the long run analogue of the popular Choleski decomposition since what is triangularised here is the matrix of long run impulses.

In cointegrated systems, this identification method is even more appealing. This is because cointegration naturally imposes such long run relationships on the ultimate effects of the system disturbances. If for example, a system has cointegration rank r, then we can set r disturbances as temporary to all the variables, and m-r disturbances as permanent. However, calculations to obtain R are not as straight forward. This is mainly because the long run impulse matrix C is

[20] Blanchard and Quah (1989) consider a bivariate model that is consist of income growth and unemployment rate as the variables. They identify the system so that one of the disturbances have no long run impact on incomes and interpret these disturbances as demand shocks.

singular for cointegrated systems. Hence, neither its inverse exists nor the Choletski decomposition of $C\Omega_\varepsilon C'$ is uniquely defined.

In this paper, the approach is slightly modified to accommodate for cointegrated systems. In particular, we decompose each variable into permanent and transitory components without fully identifying R. All we need is to assume that system disturbances are orthogonal to each other, and that only (m-r) of them have permanent effects. These two assumptions are sufficient for a decomposition of each variable into its permanent and transitory components. Formally, the identification matrix R is chosen so that

$$RR' = \Omega_\varepsilon \tag{2.3}$$

and the long run impulse matrix has a shape of a rectangular:

$$S = [S_1 \vdots 0] \tag{2.4}$$

where S_1 is a m x (m-r) matrix with all non-zero entries. Hence, the method identifies the first m-r disturbances as permanent and last r disturbances as temporary on all variables. Because the cointegration rank is r and the system disturbances are orthogonal to each other, such a decomposition naturally suits the system. Notice, however, that neither the long run impulse matrix in (2.4) is uniquely identified nor the identification matrix in (2.3) is fully determined for a given long run impulse matrix. Thus, if the individual disturbances and their dynamic effects on the variables are important, then additional identifying restrictions are needed. Such a case is considered in King et. al. (1991). They impose other constraints to uniquely determine S_1 in Equation (2.4).

This paper is not concerned with individual disturbances but a general decomposition into permanent and transitory components. Given that the identification matrix is chosen to satisfy Equations (2.3) and (2.4), each variable can be expressed as a sum of two composite components: one that consists of the first (m-r) permanent disturbances, and the other that consists of the last r transitory disturbances:

$$x_t = x_{P,t} + x_{T,t} \tag{2.5}$$

$$\Delta x_{P,t} \overset{\text{def}}{=} \sum_{j=0}^{\infty} D_1^R(j)\ \xi_1^R(t-j) \tag{2.6}$$

$$\Delta x_{T,t} \overset{\text{def}}{=} \sum_{j=0}^{\infty} D_2^R(j)\ \xi_2^R(t-j) \tag{2.7}$$

Here subscript 1 (2) indicates the first m-r (last r) columns of D^R and the corresponding system disturbances. Since R satisfies (2.4), $x_{P,t}$ and $x_{T,t}$ consist of the permanent and transitory components respectively. We will refer to such a decomposition as System-Based Permanent-Transitory decomposition.

Because $\Delta x_{i,t}$ (i=P,T) are stationary processes, each entry in them can also be expressed with a unique Wold-decomposition representation:

$$\Delta x_{P,t}^k = \sum_{j=0}^{\infty} \Psi_P^k(j)\ u_P^k(t-j) \tag{2.8}$$

$$\Delta x_{T,t}^k = \sum_{j=0}^{\infty} \Psi_T^k(j)\ u_T^k(t-j) \tag{2.9}$$

where

$$\Psi_i^k(0) = 1 \quad \forall k, i \in \{P,T\}.$$

Here superscript-k represents the k^{th} variable in Δx_i. The innovations u_P^k and u_T^k can be interpreted as *typical permanent* and *typical temporary* shocks to the k^{th} variable. These fundamental representations of the permanent and transitory component of each variable are unique, and are independent of R so long as it satisfies the conditions in (2.3) and

(2.4)[21] (see Technical Appendix-A). Hence, this paper studies these typical composite shocks rather than being concerned with each individual disturbance.

One of the important objectives of a permanent-transitory decomposition is to study the dynamic impacts of different type of disturbances on each variables. These dynamic impacts are referred as the impulse response functions, and are quite informative measures of persistence for different type of disturbances. In general, the analysis is quite sensitive to the choice of the identification matrix R. However, by using the uniqueness of the representations in (2.8) and (2.9), one can study the impulse response functions of typical permanent and typical temporary shocks for each country without fully identifying R. In particular, the cumulative sequence of moving average coefficients in (2.8) and (2.9), $\left\{ \sum_{\ell=0}^{j} \Psi_{P}^{k}(\ell) \right\}_{j=0}^{\infty}$ and $\left\{ \sum_{\ell=0}^{j} \Psi_{T}^{k}(\ell) \right\}_{j=0}^{\infty}$, give the dynamic response of variable k to a typical permanent and a typical temporary shock respectively.

3.3 Empirical Results

Identifying Permanent and Temporary Components in Incomes

In this section, we study the trend and cyclical components of incomes in six European countries[22] during the 20th Century: Austria, Belgium, Denmark, France, Italy, and Netherlands. First, we obtain a set of permanent and transitory components by applying the decomposition method described in Section 3.2. This rather mechanical algorithm picks-up the consequences of WW2 as permanent in all six countries. Then, we use a dummy variable to control for the

---

[21] This is why the components are not superscripted with R in Equations (2.6) and (2.7).

[22] We study only these six countries for compatibility with Bernard and Durlauf (1995).

asymmetric permanent effects of this major, but isolated, event. In this case, we find only one unit root process in the system which can be interpreted as world wide technological progress. Next, we study the dynamic properties of incomes by using the same decomposition methods, but with the WW2-dummy included. The results indicate that the dynamic pattern of the permanent shocks are consistent with the learning-by-doing and/or inter-industry technological diffusion ideas.

The data used in this empirical exercise are the logarithm of annual real per capita GDP in 1980 PPP-adjusted dollars (see Figure-3.1). It is the same data set that is used by Bernard and Durlauf (1995), and a detailed description can be found in their paper. For the most part, it is obtained from Maddison (1982) and Maddison (1989), and adjusted to confirm with the current national borders.

Bernard and Durlauf (1995) examine the data for unit roots and cointegration properties. Their results are also reported in Table-3.1 and Table-3.2 of this paper for completeness[23]. The presence of the unit roots are examined by using augmented Dickey-Fuller tests, and cointegration rank is search by Johansen (1991) maximum likelihood procedure.

Table-3.1

Unit Root Tests by ADF(4): 5% critical value: -3.4639

	ADF statistics		ADF statistics
Austria	0.1118	France	-0.2128
Belgium	0.4665	Italy	0.3280
Denmark	0.5856	Netherlands	-0.2569

[23] The test results are somewhat different due to the differences in calculational techniques.

According to Table-3.1, we cannot reject the null hypothesis that per capita incomes follow a unit root process for any of the 6 European Countries. Table-3.2 shows that the cointegrating rank is 3, and that there are 3 distinct common stochastic trends that are driving the system[24].

Table-3.2

Cointegration Tests:

Likelihood Ratio	5% Critical Value	No. of C.Vector(s)
138.05	104.94	None*
84.40	77.74	≤ 1*
56.99	54.64	≤ 2*
33.83	34.55	≤ 3
16.09	18.17	≤ 4
1.30	3.74	≤ 5

* rejected at the 5% level

Based on these two tables, we assume three of the six disturbances to the system have permanent effects on incomes. Without loss of generality, we choose the first three disturbances as permanent and the last three disturbances as temporary. This, however, is not enough to fully identify the system. In fact, there are infinitely many long run impulse matrices that would specify the first three disturbances as permanent and the remaining three as temporary. Hence, we do not go into the interpretation of individual

[24] Notice that choosing the cointegrating rank based on a test statistics may not be appropriate for our present purposes. Instead, one could use a consistent selection criterion to avoid asymptotic TYPE-I errors. Such criteria are studied in Uysal (1996).

disturbances, and instead study the system-based permanent-transitory decomposition as is discussed in Section 3.2.

The system-based permanent components of incomes are defined by the sum of the components that are due to the three permanent disturbances:

$$\Delta Y_P^k(t) = \sum_{j=0}^{\infty} D_{k1}^R(j)\, \xi_1^R(t-j) + \sum_{j=0}^{\infty} D_{k2}^R(j)\, \xi_2^R(t-j) + \sum_{j=0}^{\infty} D_{k3}^R(j)\, \xi_3^R(t-j) \quad (3.1)$$

Here ΔY_P^k represents the first difference of system-based permanent component of incomes in country k. The disturbances and corresponding coefficients are superscripted with R to indicate that they are associated with an identification matrix R. Having obtained the permanent components from Equation (3.1) with an arbitrary choice of R, we can estimate the unique (and independent of R) Wold-decomposition representation of the system-based permanent components for each country:

$$\Delta Y_P^k = \sum_{j=0}^{\infty} \Psi_P^k(j)\ u_P^k(t-j) \quad (3.2)$$

We construct the impulse response function of a typical permanent shock from the estimated MA coefficients in Equation (3.2). Figure-3.2 illustrates these impulse response functions for all six countries. According to the figure, the impact of a typical permanent shock increases over time, and they vary quite differently across the six countries. For Austria, for example, the ultimate impact of a typical permanent shock is about three times as big as it is for Denmark. It should be noted, however, the typical permanent shocks are different for each country, although they are constructed from the same system disturbances (because of differences in loadings).

A similar treatment is applied to the remaining three temporary disturbances of the system (i.e., ξ_4 - ξ_6). The impulse response functions of typical temporary shocks are illustrated in Figure-3.3 for all six countries[25]. According to these impulse response functions, the effect of a typical temporary shock is not very persistent, and its half life is less than five years for all six countries. Furthermore, the contribution of temporary shocks to the unanticipated annual fluctuations are not significantly bigger than that of permanent shocks since the typical innovations are of comparable sizes.

Another important observation from Figures 3.2 and 3.3 is that we cannot find any evidence supporting the Schumpeterian ideas of creative-destruction. According to this view, new positive innovations initially cause reductions in productivity due to the loss of skills with the replacement of old technology. Hence, if this was a dominant and significantly important phenomenon, one would expect that a typical positive shock (permanent, or temporary, or both) should have a negative effect on incomes initially, and only afterwards become positive. On the contrary, we find that the impulse response functions of both permanent and temporary disturbances never crosses below the zero line, and therefore a typical positive shock never reduces incomes.

Historical decomposition of the sample data into its permanent and transitory components can also be quite insightful. Figures 3.4 and 3.5 plot such historical decomposition of per capita incomes for all six European countries. It is now more transparent from these figures that much of the fluctuations in incomes are due to the permanent components. Furthermore, the major events during the century have left negative permanent effects on per capita incomes. In particular, the

[25] The ultimate effects are somewhat different than zero due to calculational approximations.

permanent components are below zero during the period from 1915 to 1970. The situation is especially worsened during the World War-II when the de-trended incomes dropped by more than 60%.

At first, a war having permanent effect on incomes may seem rather contradictory to the neo-classical growth models. This is because, the basic Solow growth model predicts that the physical capital that is destroyed during the war-time are re-constructed after the war until full recovery is achieved (i.e., convergence to a steady state). However, one can easily augment the basic model to obtain different permanent impacts of a war on all countries involved. This is because, not only the physical capital are destroyed during a war, but also the structure of countries changes after the war. For example, there are still restrictions imposed on the losers of WW2 on their spending and size of the army. Furthermore, all six countries in this paper have joined to the European Union after the war which led significant changes in their international transactions. Some might have benefited more than the others. The legislation, law and even the borders of these countries have altered during the 20th Century[26]. Naturally, some of this still ongoing restructuring may have altered the long term relative steady state level of incomes permanently across the six countries.

World-War-II and Conditional Convergence

Structural breaks in the data will bias the cointegration tests towards finding too many unit root processes. Hence, it could be the case that there is only one unit root process driving the incomes in these six European countries, and that there has been a structural break in relative steady states as a result of WW2. If confirmed by the data,

[26] Even though the data is corrected for changes in national borders, a non-uniform split of population can cause a permanent shift in per capita incomes.

this would be consistent with the neo-classical growth models, and the single unit root process would then be interpreted as world wide technological progress.

Table-3.3

Cointegration Test with WW2 dummy included:
Structural brake in relative incomes in 1939

Likelihood Ratio	5% Critical Value	No. of CE(s)
158.91	104.94	None*
100.05	77.74	$\leq 1^*$
70.52	54.64	$\leq 2^*$
43.99	34.55	$\leq 3^*$
21.47	18.17	$\leq 4^*$
2.63	3.74	≤ 5

* rejected at the 5% level

Table-3.3 reports the new cointegration tests with a pre-war dummy variable included to allow for a structural break at the beginning of WW2 (see the Technical Appendix-B for a formal description). According to these results, we reject the null hypothesis that there are more than one unit root processes in this system at the 5% level. The dummy variable that we use allows for a structural break in the deterministic component of incomes during 1939. The cointegration test results, however, are sensitive to the choice of the year of structural break. In particular, we cannot reject the null hypothesis that there are two unit root processes in the system with structural break dummies in 1938 or 1940.

In the rest of this empirical exercise, we will assume that there is only one unit root process in the system and that there has been a structural break in relative incomes in 1939. Furthermore, we interpret this unit root process as the world wide technological process, and assume that the cointegration matrix parameters satisfy the conditions for the stationarity of per capita income ratios as is implied by the neo-classical growth models. Table-3.4 includes the results of likelihood ratio test statistics for both conditional and strict convergence hypotheses. The formal definitions for both convergence concepts are given in Technical Appendix-B. Intuitively, conditional convergence means that the differences in log-per-capita-incomes are stable around a constant, where as strict convergence forces that constant to be zero. According to these test statistics we cannot reject the conditional convergence hypothesis at the 10% level, but the strict convergence hypothesis is rejected.

Table-3.4

Likelihood Ratio Tests for Convergence Hypotheses

Conditional Convergence,	$\chi^2(5)$:	9.1
Strict Convergence,	$\chi^2(10)$:	52.5

When there is only one unit root process and interpreted as permanent innovations to the world wide technological progress, the assumption about the orthogonality of permanent and temporary disturbances seems more reasonable. This is because the world wide technological progress is not likely to be exogenous with respect to the idiosyncratic shocks to relatively small economies. Therefore, one would expect it to be Granger-causally prior to the incomes of individual countries. This, however, implies the orthogonality of permanent and temporary components of incomes when the transitory dynamics correlated with the technology shocks are also included in the permanent components (see Technical Appendix-C for a formal discussion). Hence, we believe that one could treat the permanent components obtained in this exercise as the impact of world wide technological progress on the domestic economies with a higher confidence.

Figures 3.6 and 3.7 show the impulse response functions for typical permanent and temporary shocks respectively. According Figure-3.6, the permanent components are smaller compared to the earlier results. This is because the contribution of the WW2 to the permanent component has been considerable, and that controlling for this major event reduces the size of the remaining permanent component. The general dynamic pattern of these impulse response functions are consistent with the learning-by-doing and/or inter-industry technological diffusion ideas. Furthermore, a positive technology shock causes incomes to overshoot in the medium run in all six countries. This would be consistent with an initial boost in the investment demand to adopt the new technology. We also find some weak evidence for the creative-destruction hypothesis in the case of Austria where a positive permanent shock reduces the incomes for up to a year in this country. It is also interesting to notice that there are big differences on the transitory effects of permanent shocks. In Denmark for example, the initial overshooting is relatively smaller, where as in France it is bigger and last longer. There is not yet an adequate theoretical study for explaining such cross country differences in temporary effects of technology shocks. But the differences in factor input markets, industrial composition, and business legislation and law are probably important determinants for a first approximation.

Not surprisingly, temporary shocks are bigger and more persistent in this case where the permanent effects of WW2 are separated out (see Figure-3.7). This is because, all disturbances except one are forced to be temporary in all countries. In addition, any transitory dynamics due to WW2 is also added to the temporary disturbances. Consequently, a typical temporary shock accounts for more of the income fluctuations than it did before. However, one should be wary of interpreting all this temporary behaviour as a result of demand side fluctuations[27]. Clearly it includes the dynamics due to temporary supply side shocks (such as changes in weather, natural disasters etc.) as well as the recovery from the physical destruction of WW2.

Figures 3.8 and 3.9 report the historical decomposition of incomes under the conditional convergence restrictions and with WW2 dummy included. The permanent components do not have the big fall during the WW2 years in this case. Instead, a bigger share of the adverse impacts of the War is covered with the temporary fluctuations. Especially those countries that felt more of the heat during the War experienced a larger temporary drop in their incomes. They all, however, recover quickly back to the pre-war levels by 1950. This speed of recovery is quite faster than one would expect in the light of Solow growth model, and suggest that the economic losses during the War were larger than that can be explained by physical destruction alone[28].

---

[27] Blanchard and Quah (1989) interprets temporary shocks as demand disturbances.
[28] With the usual calibration of Solow growth model, the half-life of a temporary deviation from the steady state is about 11 years.

It is one of the motivations of the ongoing empirical research that to estimate the contribution of technological progress to transitory fluctuations of incomes. Table-3.5 lists the share of permanent shocks in conditional variance of forecast errors for several horizons.

Table-3.5

Share of Permanent Shocks in Transitory Fluctuations (%)

Horizon	Austria	Belgium	Denmark	France	Italy	Nether.
0	9.9	12.9	0.0	1.6	17.6	1.5
5	9.2	22.4	31.2	26.0	17.0	18.8
10	21.7	23.6	32.1	30.2	24.7	24.3
20	22.6	23.8	32.1	30.2	25.0	24.5
40	22.6	23.8	32.1	30.2	25.0	24.5

According to Table-3.5, the short run contribution of permanent shocks to the transitory dynamics of incomes varies substantially across the countries. In the case of Denmark, for example, the short run transitory fluctuations are completely accounted by the temporary shocks. This is because the technology shocks almost immediately reach their final effect without having extended transitory dynamics (see Figure-3.6 for the impulse response function of a technology shock in Denmark). On the contrary, in Italy more than 15% of the transitory dynamics in the current year comes from the technology shocks. The long run contribution of technology shocks to the transitory dynamics is more comparable across countries, and is about 25%.

3.4 Conclusions

This chapter is an empirical exercise to study 20th Century economic growth in six European countries. It searches for the long run relationships between the income levels of these six countries, and decomposes their incomes into permanent and transitory components.

Our analysis identifies WW2 as the major event of the Century that had different permanent effects in all countries. Hence, formal cointegration tests find more than one unit root processes if one does not control for a structural break during the War. If, on the other hand, the permanent effect of WW2 on incomes is controlled for, the data is consistent with the implications of neo-classical growth models. In particular, we find that the income differences has been stationary during the century. This is consistent with the conditional convergence ideas across these six European countries. Furthermore, we also test for unconditional convergence of incomes. That is, the incomes differ by temporary deviations only which they converge back in time. The data rejects the unconditional convergence hypothesis.

We also examine the data for the behaviour of typical permanent and transitory shocks. We find that the impulse response function of permanent shocks is consistent with the stochastic technological progress where the dynamic pattern is interpreted as learning-by-doing or inter-industry technological diffusion ideas. The functions also indicate an investment demand boost in the medium term. The overall contribution of the permanent shocks to the transitory dynamics in incomes is about 25%. Finally, we find that temporary shocks are quite big in magnitude and the effects are persistent for up to 5 years.

Figure - 3.1

European Incomes During 20th Century

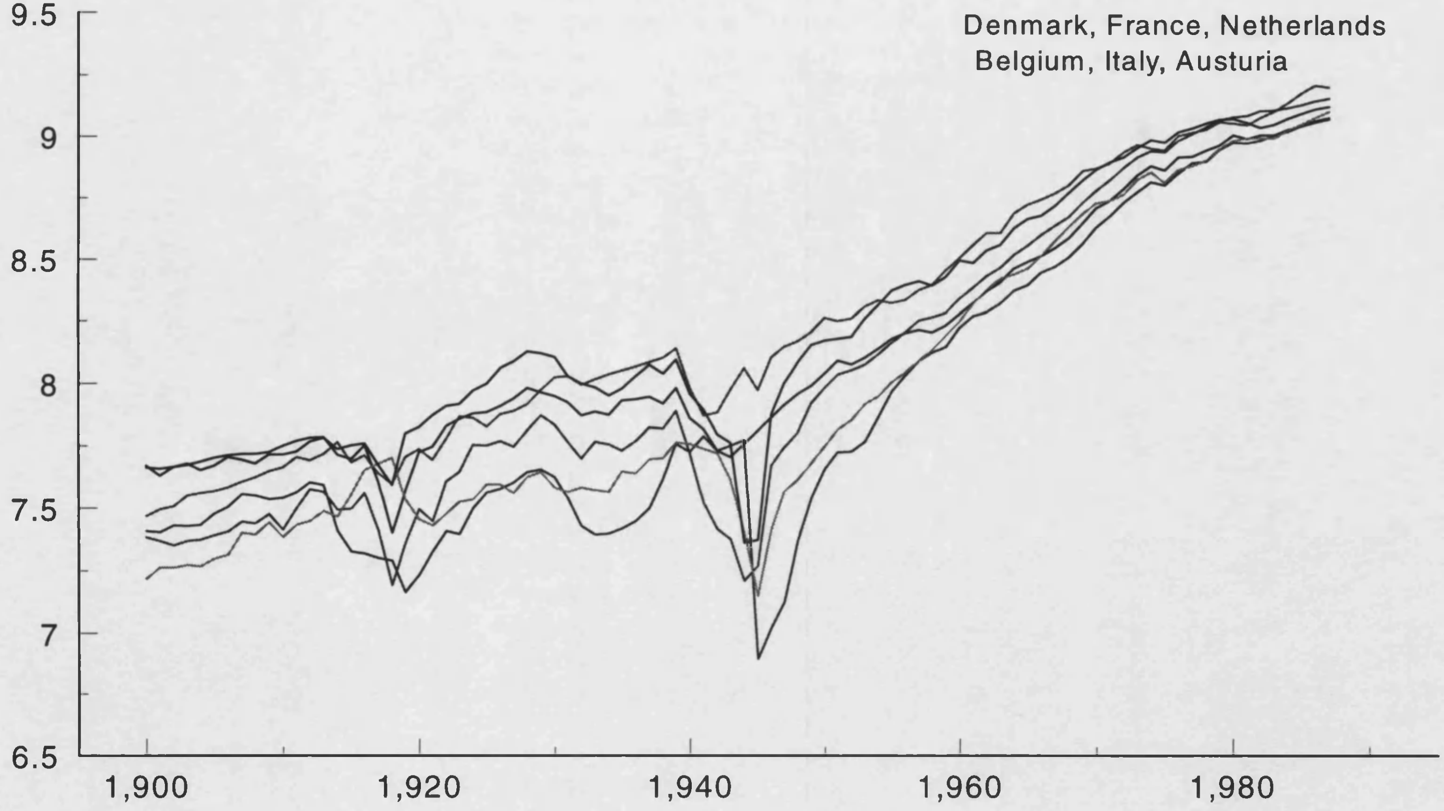


Figure-3.2
Impulse Response Functions: Permanent Shocks

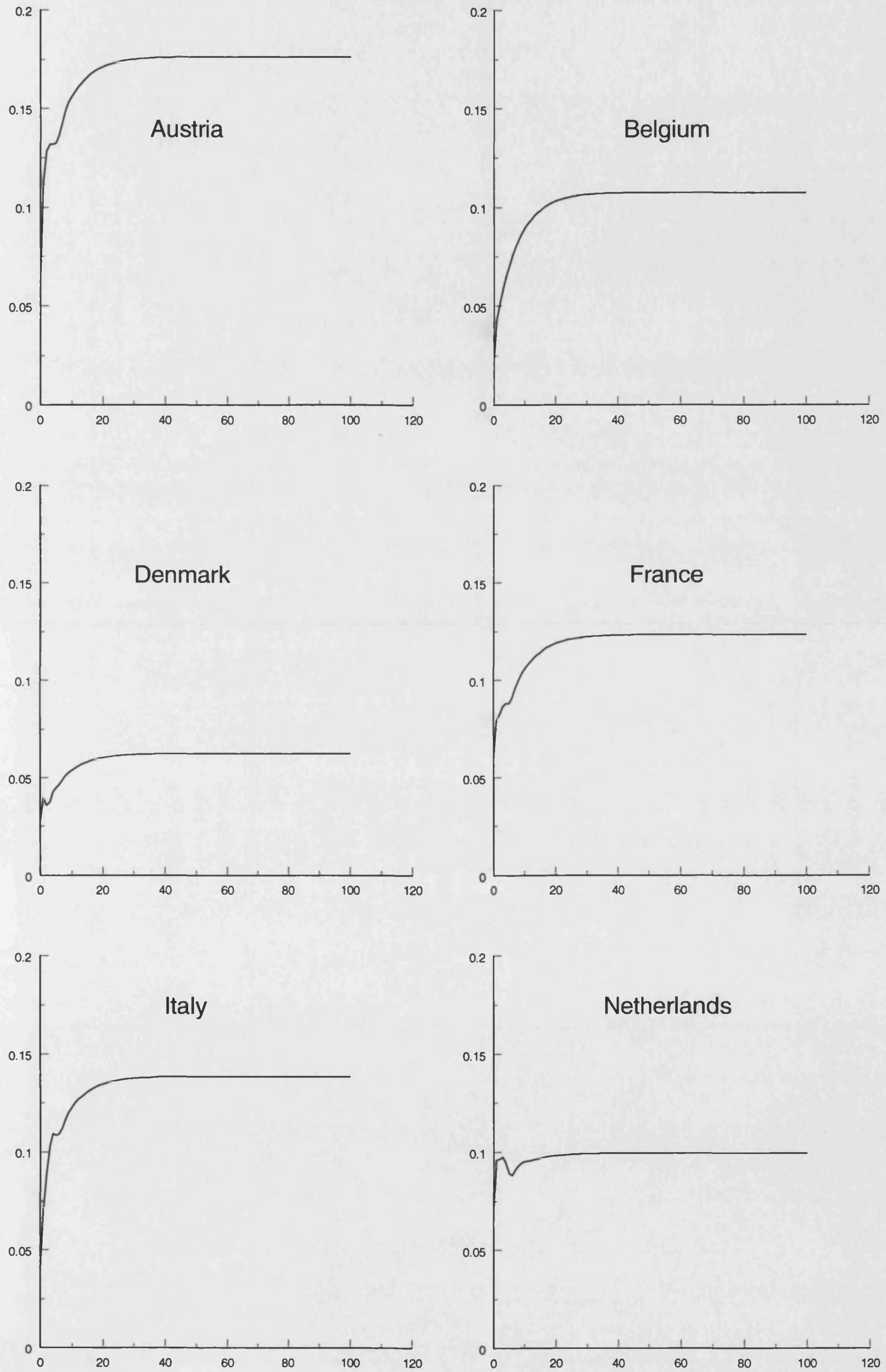


Figure-3.3
Impulse Response Functions: Temporary Shocks

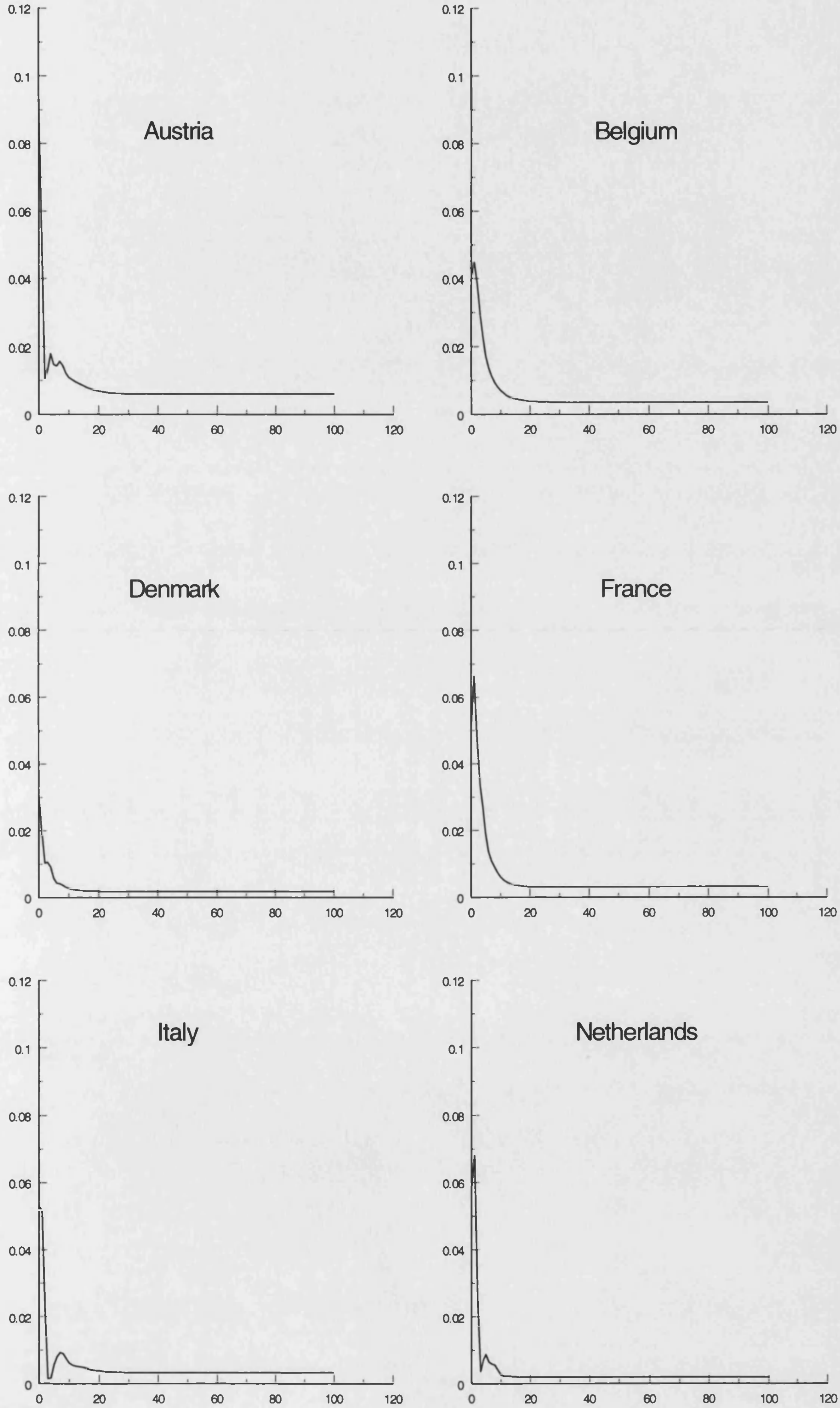


Figure-3.4

20th Century Euro-Growth: Permanent Components

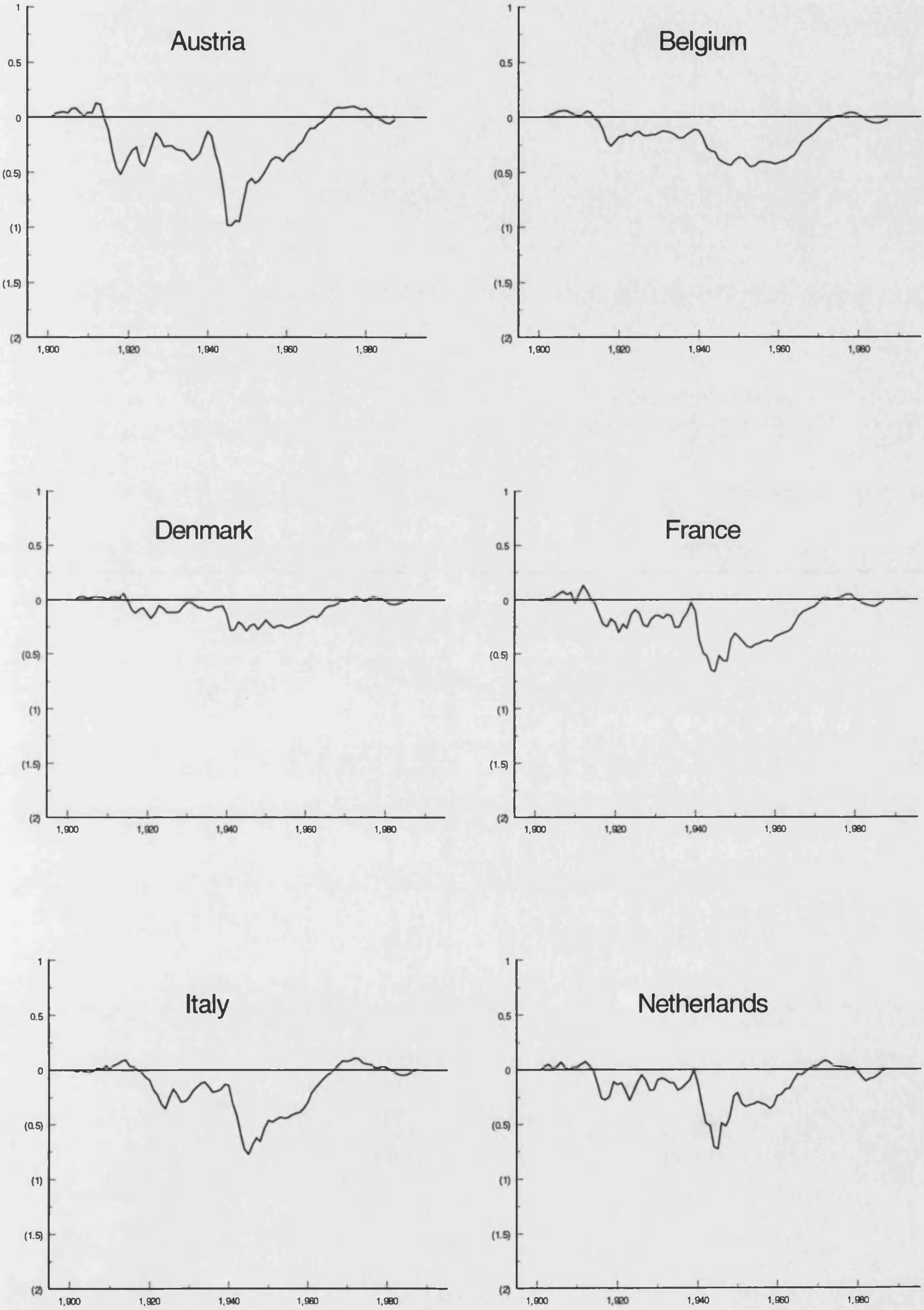


Figure-3.5

20th Century Euro-Growth: Temporary Components

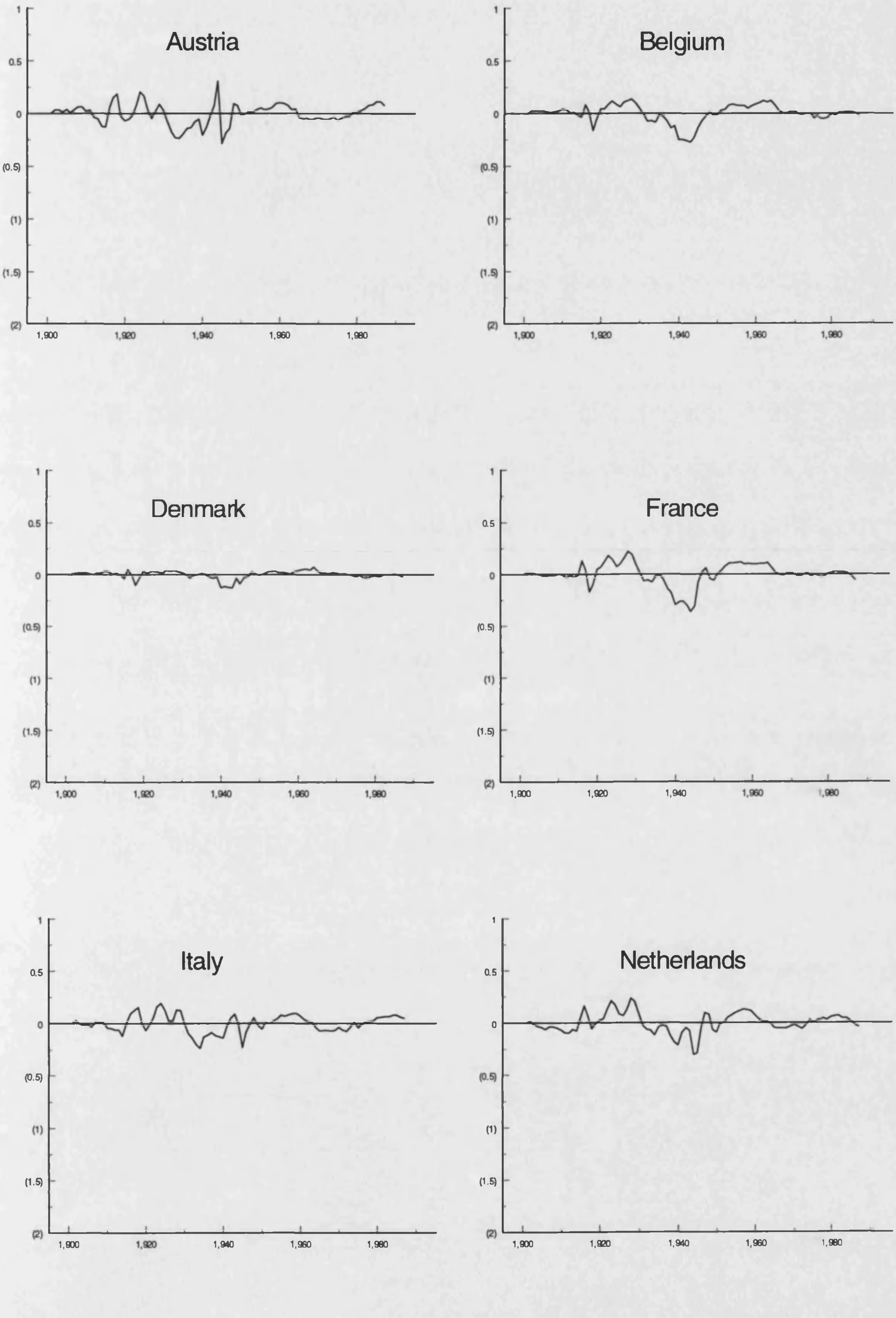


Figure-3.6

Impulse Response Functions (with WW2 Dummy): Permanent Shocks

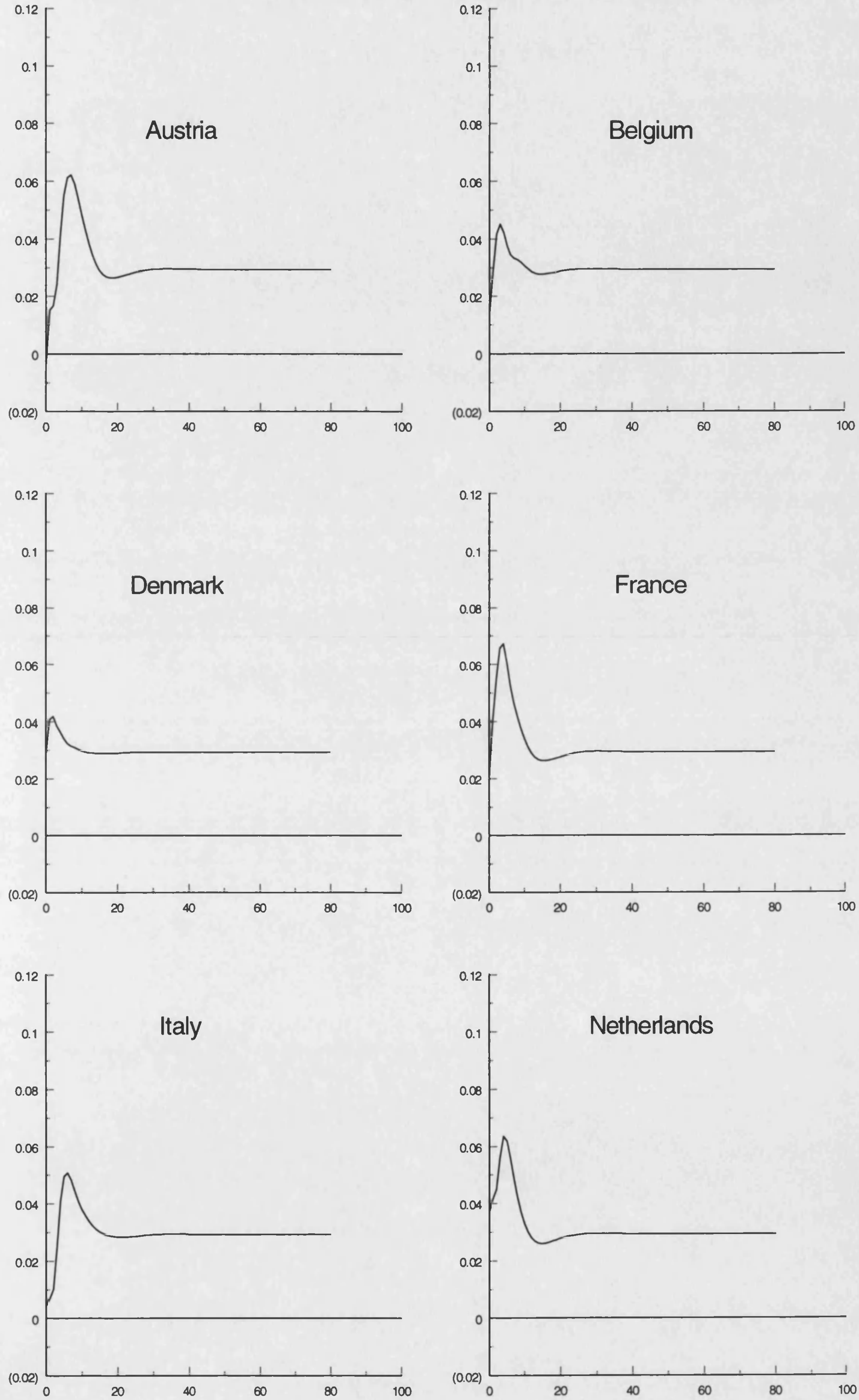


Figure-3.7
Impulse Response Functions (with WW2 Dummy): Temporary Shocks

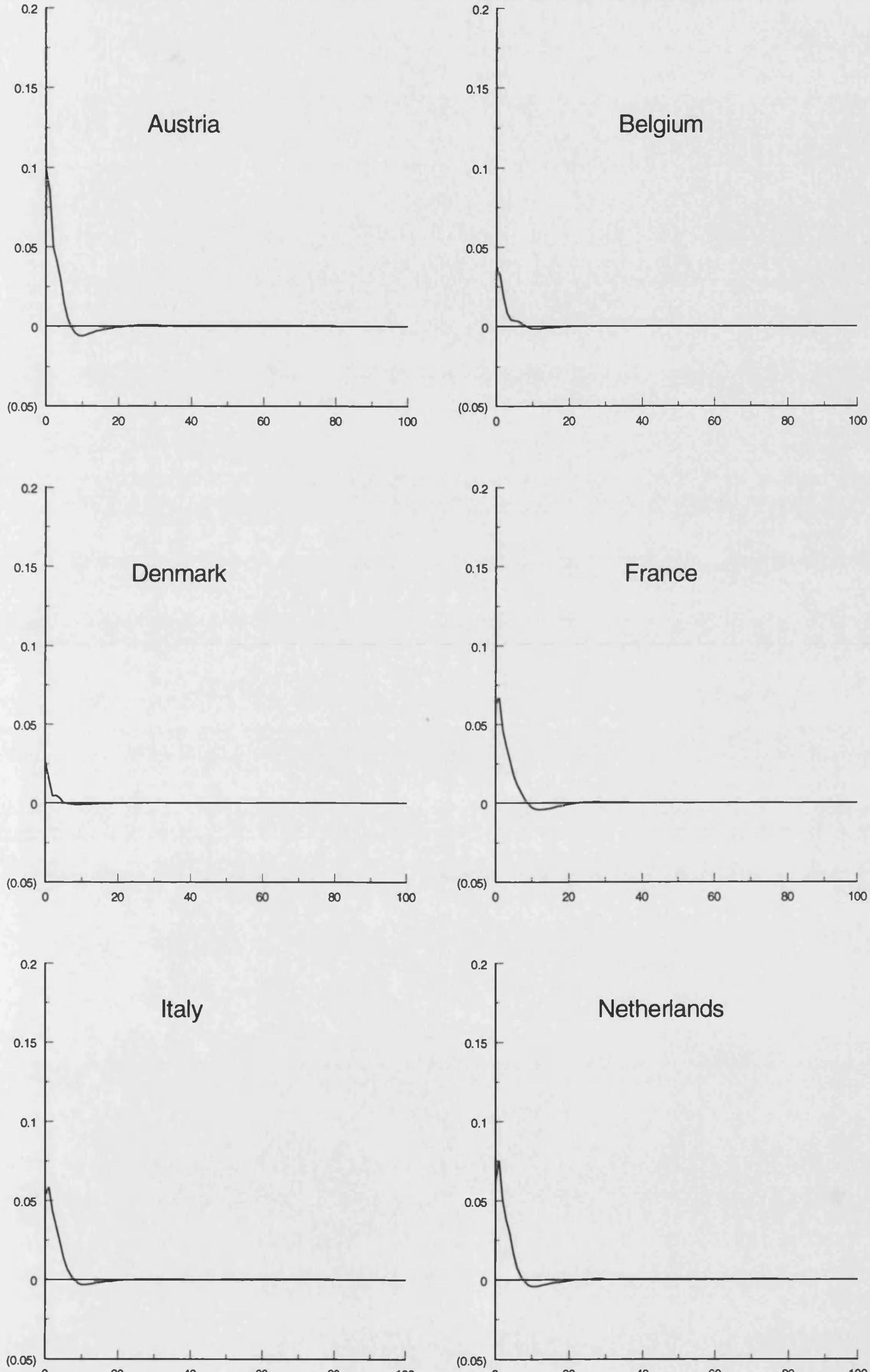


Figure-3.8

20th Century Euro-Growth (with WW2 Dummy): Permanent Components

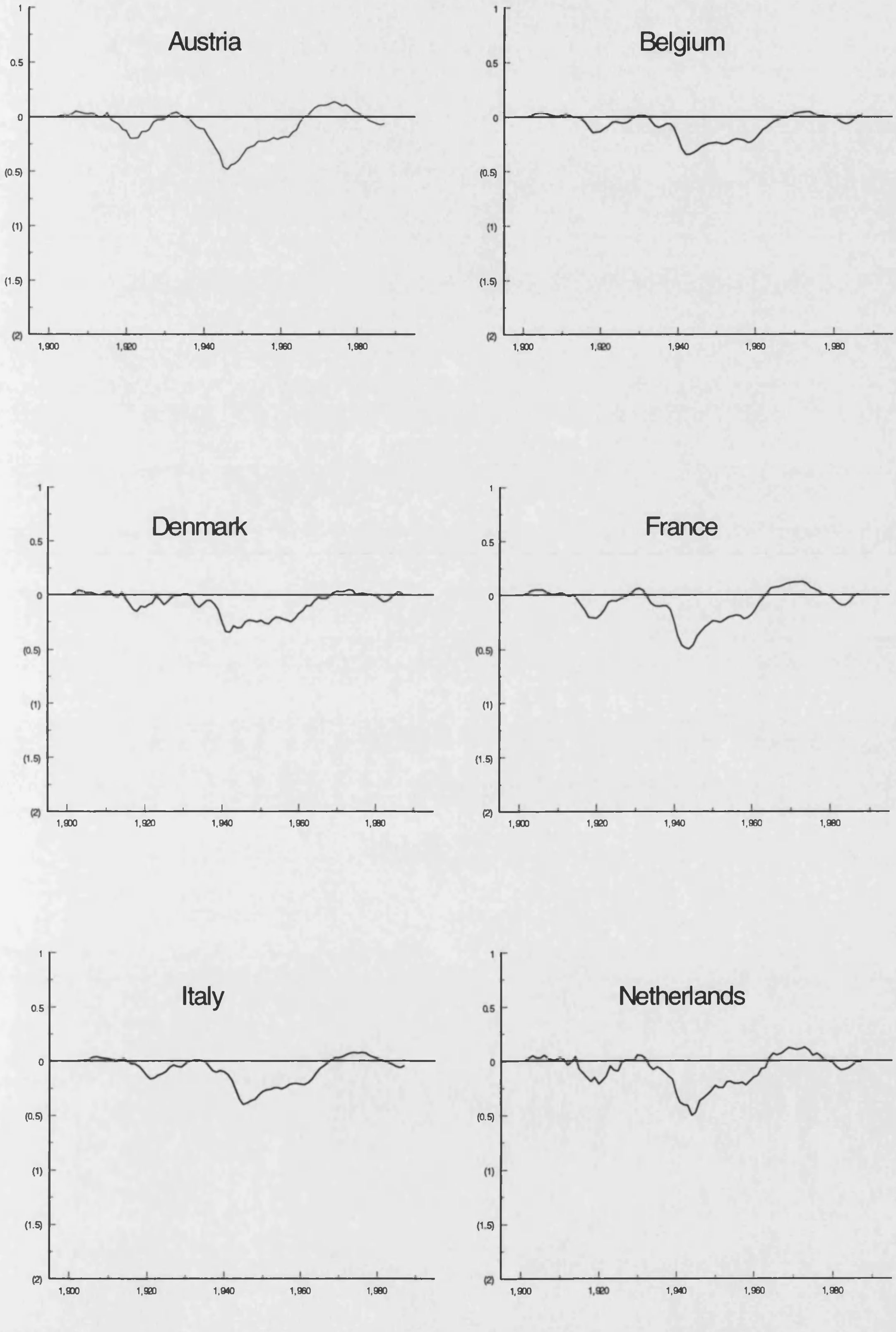
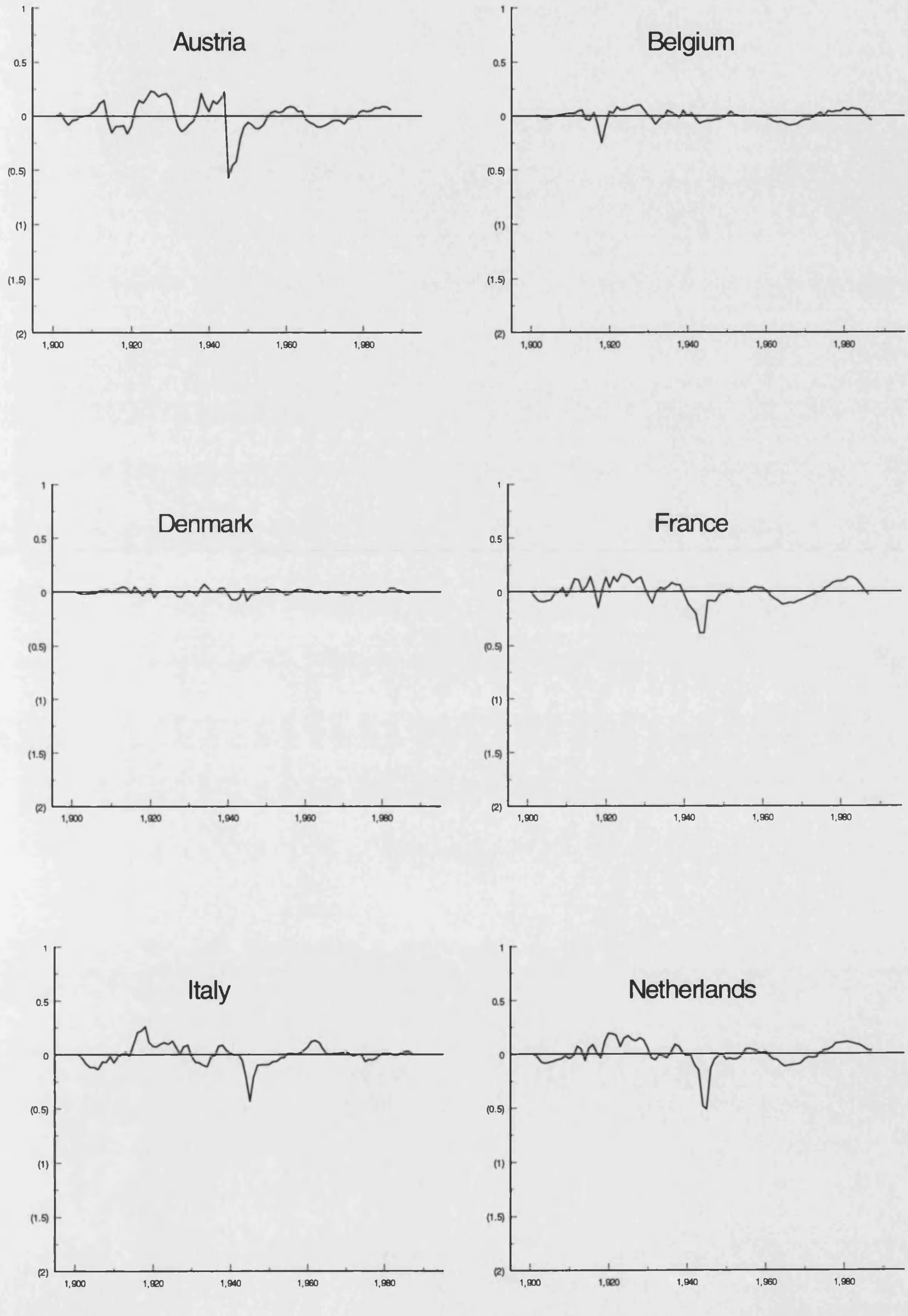


Figure-3.9

20th Century Euro-Growth (with WW2 Dummy): Temporary Components



3.5 Technical Appendix-A

UNIQUENESS OF SYSTEM BASED PERMANENT TRANSITORY DECOMPOSITION:

Consider an m-dimensional multivariate system given by

$$\Delta x_t = \sum_{j=0}^{\infty} C_j \, \varepsilon_{t-j} \tag{A.1}$$

where

$$C_0 = I_m ,$$

$$E(\varepsilon_t \varepsilon_s') = \begin{cases} \Omega_\varepsilon & \text{if } t = s \\ 0 & \text{otherwise} \end{cases}$$

Suppose the system is cointegrated with r distinct cointegrating vectors and that the rank of

$$C = \sum_{j=0}^{\infty} C_j$$

is m-r.

Let R be a non-singular identification matrix of the system so that

$$\Delta x_t = \sum_{j=0}^{\infty} D^R(j) \, \xi^R_{t-j} \tag{A.2}$$

where

$$D^R(j) = C_j \, R \tag{A.3}$$

$$\xi^R_t = R^{-1} \, \varepsilon_t \tag{A.4}$$

$$RR' = \Omega_\varepsilon \tag{A.5}$$

and the new long run impulse matrix

$$S = \sum_{j=0}^{\infty} D^R(j)$$
$$= C\,R$$

is such that it selects the last r disturbances as temporary to the system:

$$S = [S_1 \vdots 0] \qquad \text{(A.6).}$$

Because the rank of C is m-r, such an identification matrix, R, exists though it is not unique.

Now, consider the following system-based permanent-transitory decomposition:

$$x_t = x_{P,t} + x_{T,t} \tag{A.7}$$

$$\Delta x_{P,t} \overset{\text{def}}{=} \sum_{j=0}^{\infty} D_1^R(j)\ \xi_1^R(t-j) \tag{A.8}$$

$$\Delta x_{T,t} \overset{\text{def}}{=} \sum_{j=0}^{\infty} D_2^R(j)\ \xi_2^R(t-j) \tag{A.9}$$

where subscript 1 (2) indicates the first m-r (last r) columns of $D^R(j)$ and the corresponding system disturbances. Since R satisfies (A.6), $x_{P,t}$ and $x_{T,t}$ consist of the permanent and transitory components respectively (i.e., the effect of the innovations to $x_{T,t}$ eventually disappear, where as the ultimate impact of the innovations to $x_{P,t}$ is non-zero).

Because Δx_i (i=P,T) are stationary processes under some regularity conditions, each entry in them can also be expressed with a unique Wold-decomposition representation:

$$\Delta x_P^k = \sum_{j=0}^{\infty} \Psi_P^k(j)\ u_P^k(t-j) \tag{A.10}$$

$$\Delta x_T^k = \sum_{j=0}^{\infty} \Psi_T^k(j)\ u_T^k(t-j) \tag{A.11}$$

where

$$\Psi_i^k(0) = 1 \quad \forall k, i \in \{P, T\}.$$

Here, the superscript k represents the k^{th} variable in Δx_i.

The following theorem states that the moving average coefficients in (A.10) and (A11) are independent of the choice of R so long as (A.5) and A(6) are satisfied. Hence, it proves the uniqueness of the system based permanent transitory decomposition given in (A.7).

THEOREM A.1.

Consider the m-dimensional system of integrated variables in (A.1) with r cointegrating vectors. Let R be an identification matrix that satisfies the conditions in (A.5) and (A.6). Then, the moving average coefficients in the (A.10) and (A.11) are independent of R.

PROOF OF THEOREM A.1.

Because

$$\Delta x_{P,t} \overset{\text{def}}{=} \sum_{j=0}^{\infty} D_1^R(j)\ \xi_1^R(t-j)$$

$$\Delta x_{T,t} \overset{\text{def}}{=} \sum_{j=0}^{\infty} D_2^R(j)\ \xi_2^R(t-j)$$

we have

$$\begin{aligned} E\left(\Delta x_{P,t}\, \Delta x'_{P,t-s}\right) &= \sum_{j=0}^{\infty} D_1^R(j)\, D_1^{R'}(j+s) \\ &= \sum_{j=0}^{\infty} C(j)\, R E_1 E_1' R' C'(j+s), \end{aligned}$$

and

$$\begin{aligned} E\left(\Delta x_{T,t}\, \Delta x'_{T,t-s}\right) &= \sum_{j=0}^{\infty} D_2^R(j)\, D_2^{R'}(j+s). \\ &= \sum_{j=0}^{\infty} C(j)\, R E_2 E_2' R' C'(j+s), \end{aligned}$$

where

$$E_1 = \begin{bmatrix} I_{m-r} \\ 0 \end{bmatrix}$$

$$E_2 = \begin{bmatrix} 0 \\ I_r \end{bmatrix}.$$

To prove the Wold Decomposition representations in (A.10) and (A.11) are unique, it suffices to show that above autocovariance matrices are independent of R. However, the autocovariance matrices for both temporary and permanent components would be independent of R if and only if $R E_2 E_2' R'$ is independent of R. To show this is the case when conditions (A.5) and (A.6) are satisfied, we need to following decomposition of the matrices:

$$R = \begin{bmatrix} R_1 & R_2 \\ R_3 & R_4 \end{bmatrix} \qquad S = \begin{bmatrix} S_{11} & 0 \\ S_{12} & 0 \end{bmatrix} \qquad C = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} \begin{bmatrix} I_{m-r} & \delta_3 \end{bmatrix}$$

$$\Omega_\varepsilon = \begin{bmatrix} \Omega_1 & \Omega_2 \\ \Omega_3 & \Omega_4 \end{bmatrix} \qquad C \Omega_\varepsilon C' = \begin{bmatrix} \omega_1 & \omega_2 \\ \omega_3 & \omega_4 \end{bmatrix}$$

Here, R_1, S_{11}, and δ_1 are (m-r) x (m-r) matrices[29]. Then,

$$C R = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} \begin{bmatrix} I_{m-r} & \delta_3 \end{bmatrix} \begin{bmatrix} R_1 & R_2 \\ R_3 & R_4 \end{bmatrix} = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} \begin{bmatrix} R_1 + \delta_3 R_3 & R_2 + \delta_3 R_4 \end{bmatrix}$$

By equating the last expression to S, one obtains,

$$R_1 = \delta_1^{-1} S_{11} - \delta_3 R_3 \tag{A.12}$$

$$R_2 = -\delta_3 R_4 \tag{A.13}$$

$$R = \begin{bmatrix} -\delta_3 \\ I_r \end{bmatrix} \begin{bmatrix} R_3 & R_4 \end{bmatrix} + \begin{bmatrix} \delta_1^{-1} S_{11} & 0 \\ 0 & 0 \end{bmatrix} \tag{A.14}$$

(A.14) and (A.5) can be used to obtain the following;

$$R_3 R_3' + R_4 R_4' = \Omega_4 \tag{A.15}$$

$$R_3 R_3' = (\Omega_2 + \delta_3 \Omega_4)' \delta_1' \omega_1^{-1} \delta_1 (\Omega_2 + \delta_3 \Omega_4) \tag{A.16}$$

Notice that

---

[1] Notice that C can be decomposed in this way since its rank is (m-r).

$$
\begin{aligned}
R E_2 E_2' R' &= \begin{bmatrix} R_2 R_2' & R_2 R_4' \\ R_4 R_2' & R_4 R_4' \end{bmatrix} \\
&= \begin{bmatrix} \delta_3 R_4 R_4' \delta_3' & \delta_3 R_4 R_4' \\ R_4 R_4' \delta_3' & R_4 R_4' \end{bmatrix}
\end{aligned}
$$

Hence, it is independent of R by using the relationships in (A15) and (A.16). This concludes the proof of Theorem A.1.

Notice that the proof of the theorem does not impose any conditions on S_{11} other than its non-singularity. Indeed, for any long run impulse matrix with S_{11} non-singular, there exist a matrix R satisfies both (A.5) and (A.6). Then, we can choose S_{11} such that in its k^{th} row all entries are zero except the first and still satisfy the conditions (A.5) and (A.6). Such a long run impulse matrix would identify all the innovations to the k^{th} variable as temporary except the first disturbance which is permanent. We refer to such a decomposition as the variable-based permanent-temporary decomposition. Clearly, the permanent component of the k^{th} variable obtained this way is at most as big as the permanent component obtained in a system-based permanent-transitory decomposition. This is because the system-based permanent components are unique and contain the combined effects of the first (m-r) disturbances to the system. However, the permanent component in a variable-based decomposition is consist of only the first of these (m-r) disturbances. Because the disturbances are orthogonal to each other, their joint variance should be greater than the variance of any single one alone. Hence, the permanent component in a system based decomposition should be at least as big as the permanent component in a variable based decomposition with equality only when m-r=1. At this point it is also easy to notice that the permanent component in a variable-based decomposition is Granger-causally prior to the other variables. This is because it is consist of only one of the system disturbances and hence cannot be further decomposed.

3.6 Technical Appendix-B

STRUCTURAL BREAK, PRE-WAR DUMMY, AND CONVERGENCE

We assume the log-per-capita-incomes consist of a deterministic linear trend component and a stochastic component that is integrated of order one;

$$y(t) = y^{s}(t) + y^{d}(t) \tag{B.1}$$

where

$$y^{d}(t) = \theta + m \cdot t \tag{B.2}$$

$$\Delta y^{s}(t) = \Psi \Delta y^{s}(t-1) + \alpha\beta y^{s}(t-2) + \varepsilon(t). \tag{B.3}$$

Here, superscripts "s" and "d" indicates stochastic and deterministic components respectively. "m" is the vector of long run growth rates and θ is the vector of relative incomes. A structural break is formulated by allowing these structural form parameters of the deterministic component to change. In particular, we assume

$$y^{d}(t) = \theta_1 + m_1 t + \theta_2 d(t) + m_2 t\, d(t) \tag{B.4}$$

where

$$d(t) = \begin{cases} 1 & \text{if } t \geq 1939 \\ 0 & \text{otherwise} \end{cases} \tag{B.5}$$

is the pre-war dummy that controls for the timing of the structural break. Notice that both θ_1 and θ_2 are identifiable only up to "r" parameters where "r" is the number of the cointegrating vectors. In our experiment with six countries and five cointegrating vectors, it suffices to assume that last entries in both θ_1 and θ_2 are zero. This is equivalent to measuring relative incomes with respect to that of Netherlands and assuming that no shift took place in this country's per capita income because of the War.

Cointegration tests are formulated in terms of reduced form representations. The reduced form equivalent of the model described in (B.1), (B.3), (B.4) and (B.5) is given by

$$\begin{aligned}\Delta y(t) = \mu + g\,t + \gamma_1 \Delta d(t) + \gamma_2 \Delta d(t-1) \\ + \gamma_3 d(t-2) + \Psi \Delta y(t-1) + \alpha \beta y(t-2) + \varepsilon(t)\end{aligned} \tag{B.6}$$

where μ, g, γ_1, γ_2, and γ_3 are restricted functions of the structural form parameters. Because the cointegration tests are studied on unrestricted reduced form representations, and the critical values are not robust on imposition of such restrictions, we search the cointegration rank by using the unrestricted version of (B.6).

In the light of the decomposition of incomes in (B.3) and (B.4), one can also define two convergence concepts as is implied by the neo-classical growth models;

- *Conditional Convergence*: Per capita incomes of two countries is said to satisfy the conditional convergence hypothesis if β in (B.3) is in the space spanned by the vector (1, -1).
- *Strict Convergence*: Per capita incomes of two countries is said to satisfy the strict convergence hypothesis if β in (B.3) is in the space spanned by the vector (1, -1), and that $\beta(\theta_1\text{-}\theta_2) = 0$.

In other words, we say that per capita incomes of two countries satisfy the conditional convergence hypothesis if the difference between the stochastic components is stationary around a constant. Furthermore, we say that there is a strict convergence of incomes after the War between these two countries if in addition that constant is reduced to zero. These definitions of convergence concept are closely related to those discussed in Bernard and Durlauf (1995). Interested readers are referred to their paper for further discussion.

3.7 Technical Appendix-C

GRANGER CAUSALLY PRIOR WORLD WIDE TECHNOLOGICAL PROGRESS AND ORTHOGONALITY OF PERMANENT AND TRANSITORY COMPONENTS:

Let ω_t be the world wide technological progress that drives the permanent component of per capita incomes. Furthermore, let τ_t be the temporary fluctuations. Suppose, they have the following canonical MA representations:

$$\Delta\omega_t = \sum_{j=1}^{\infty} \Gamma_j u_{t-j} \qquad \text{(C.1)}$$

$$\begin{bmatrix} \Delta\omega_t \\ \tau_t \end{bmatrix} = \sum_{j=1}^{\infty} \begin{bmatrix} \Phi_{1,j} & \Phi_{2,j} \\ \Phi_{3,j} & \Phi_{4,j} \end{bmatrix} \begin{bmatrix} \xi_{1,t-j} \\ \xi_{2,t-j} \end{bmatrix} \qquad \text{(C.2)}$$

where

$$u_t \sim IID(0,\sigma^2),$$

$$\begin{bmatrix} \xi_{1,t} \\ \xi_{2,t} \end{bmatrix} \sim IID(0,\Omega_\xi),$$

and $\Gamma_1 = 1$, Φ_1=1, Φ_2=0, Φ_3=1, Φ_4=0. Here, $\Delta\omega_t$ can be interpreted as innovations to the world wide technological progress in terms of new inventions in production techniques. We allow for serial dynamics in this process, since one new invention usually leads to new opportunities for the discovery of others.

For relatively small economies, it is reasonable to assume that world wide technological progress is Granger causally prior to incomes, and in particular, to its temporary components. This, on the other hand, implies that

$$V\left(\omega_{t+k} \middle| \{\omega_{t-j}\}_{j=1}^{\infty}\right) = V\left(\omega_{t+k} \middle| \{\omega_{t-j}, \tau_{t-j}\}_{j=1}^{\infty}\right). \qquad \text{(C.3)}$$

Furthermore, because these k-step ahead forecast error variances are optimal they cannot be improved upon by using previous observations.

This implies that the 1-step ahead forecast error of $\Delta\omega_t$ based on (C.1), u_t, should be equal to the 1-step ahead forecast error of $\Delta\omega_t$ based on (C.2), $\xi_{1,t}$. Hence, using the uniqueness of the Wold decomposition representations in (C.1) and (C.2), $\Delta\omega_t$ is Granger-causally-prior to τ_t if, and only if,

$$\Phi_{2,j} \quad = \quad 0, \text{ for all } j \geq 0.$$

In other words, (C.1) and (C.2) can be combined to obtain

$$\begin{bmatrix} \Delta\omega_t \\ \tau_t \end{bmatrix} = \sum_{j=1}^{\infty} \begin{bmatrix} \Gamma_j & 0 \\ \Phi_{3,j} & \Phi_{4,j} \end{bmatrix} \begin{bmatrix} u_{t-j} \\ \xi_{2,t-j} \end{bmatrix} \tag{C.4}$$

where,

$$\begin{bmatrix} u_t \\ \xi_{2,t} \end{bmatrix} \sim \text{IID}(0,\Omega_\xi).$$

Furthermore, the representation in (C.4) can be transformed into

$$\begin{bmatrix} \Delta\omega_t \\ \tau_t \end{bmatrix} = \sum_{j=1}^{\infty} \begin{bmatrix} \sigma\Gamma_j & 0 \\ C_{1,j} & C_{2,j} \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma} u_{t-j} \\ e_{t-j} \end{bmatrix} \tag{C.5}$$

where,

$$\begin{bmatrix} \sigma\Gamma_j & 0 \\ C_{1,j} & C_{2,j} \end{bmatrix} = \begin{bmatrix} \Gamma_j & 0 \\ \Phi_{3,j} & \Phi_{4,j} \end{bmatrix} R^{-1} \tag{C.6}$$

$$\begin{bmatrix} \frac{1}{\sigma} u_t \\ e_t \end{bmatrix} = R \begin{bmatrix} u_t \\ \xi_{2,t} \end{bmatrix}, \tag{C.7}$$

and R is the unique Choleski decomposition of Ω_ξ.

The representation in (C.5) can be used to decompose the stochastic part of per capita incomes into three main components:

$$y_t \quad = \quad \omega_t \quad + \quad \tau_{u,t} \quad + \quad \tau_{e,t} \tag{C.8}$$

where

$$\tau_{u,t} = \frac{1}{\sigma}\sum_{j=1}^{\infty} C_{1,j}\, u_{t-j} \tag{C.9}$$

$$\tau_{e,t} = \sum_{j=1}^{\infty} C_{2,j}\, e_{t-j} \tag{C.10}$$

Here $\tau_{u,t}$ can be interpreted as the temporary fluctuations in incomes due to the world wide technological progress. It could be as a result of a learning process, inter-industry technological diffusion, temporary adjustment of labour in response to the permanent changes in wages, and a short term boost in physical capital investment demand. $\tau_{e,t}$ is orthogonal to the other two components and includes other demand side fluctuations as well as temporary technology shocks.

Consequently, Granger causality of world wide technological progress implies that incomes can be decomposed into orthogonal permanent and transitory components where the permanent components include world wide technological progress and its transitory effects on incomes (i.e., $\omega_t + \tau_{u,t}$). When this result combined with the uniqueness of the system based permanent temporary decomposition (see Technical Appendix-A), it is only natural to interpret the permanent components as the consequences of world wide technological progress on domestic incomes. It should be noted, however, this requires implicit assumptions on the nature of the temporary shocks across the six countries in our empirical exercise. In particular, we assume that the temporary shocks to these six countries can be adequately expressed by five main factors (similar implicit assumptions exist also in other empirical work).

3.8 References

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

Barro, R. J., & Salla-i Martin, X. (1992). Convergence. Journal of Political Economy, 100, 223-251.

Bernard, A. B., & Durlauf, S. N. (1995). Convergence in International Output. Journal of Applied Econometrics, 10, 97-108.

Beveridge, S., & Nelson, C. R. (1981). A New Approach to the Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle. Journal of Monetary Economics, 7, 151-174.

Blanchard, O. J., & Quah, D. (1989). The Dynamic Effect of Aggregate Demand and Supply Disturbances. The American Economic Review, 79, 655-673.

Campbell, J. Y., & Mankiw, N. G. (1987). Are Output Fluctuations Transitory. Quarterly Journal of Economics, 102, 798-814.

Clark, P. K. (1987). The Cyclical Component of the U.S. Economic Activity. Quarterly Journal of Economics, 102, 798-814.

Cochrane, J. H. (1988). How Big is the Random Walk Component in GNP. Journal of Political Economy, 96, 893-920.

Evans, G. W. (1989). Output and Unemployment Dynamics in the United States: 1950-1985. Journal of Applied Econometrics, 4, 213-237.

Granger, C. W. J., & Engle, R. F. (1987). Co-integration And Error Correction: Representation, Estimation, And Testing. Econometrica, 55, 251-76.

Griliches, Z. (1957). Hybrid Corn: an Exploration in the Economics of Technological Change. Econometrica, 25, 501-522.

Harvey, A. C. (1985). Trends and Cycles in Macroeconomic Time Series. Journal of Business and Economic Statistics, 3, 216-227.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models. Econometrica, 59, 1551-80.

Jovanovic, B., & Lach, S. (1990). Entry, Exit and Diffusion with learning by Doing. The American Economic Review, 79, 690-699.

King, R. G., Stock, J. H., Plosser, C. I., & Watson, M. (1991). Stochastic Trends and Economic Fluctuations. The American Economic Review, 81, 819-839.

Lippi, M., & Reichlin, L. (1994). Diffusion of Technical Change and the Decomposition of Output into Trend and Cycle. Review of Economic Studies, 61, 19-30.

Maddison, A. (1982). Phases of Capitalist Development. Oxford University Press.

Maddison, A. (1989). The World Economy in the 20th Century. Development Centre of the Organisation for Economic Co-operation and Development.

Nelson, C. R., & Plosser, C. I. (1982). Trends and Random Walks in Macroeconomic Time Series. Journal of Monetary Economics, 10, 129-162.

Quah, D. (1993). Empirical Cross Section Dynamics in Economic Growth. European Economic Review, 37, 426-34.

Quah, D. (1992). The Relative Importance of Permanent and transitory Components: Identification and some Theoretical Bounds. Econometrica, 60, 107-118.

Uysal, E. (1995). Endogeneity in Long Run Growth Performance. Unpublished draft copy.

Watson, N. M. (1986). Univariate Detrending Methods with Stochastic Trends. Journal of Monetary Economics, 18, 49-75.

Young, A. (1991). Learning by Doing and the Dynamic Effects of International Trade. Quarterly Journal of Economics, 151, 369-405.

CHAPTER 4

Criteria for Simultaneous Selection of Autoregressive Order with Cointegrating Rank

4.1 Introduction

Recent developments in multivariate analysis of integrated time series has found many applications in empirical research. Discoveries regarding to the cointegration properties of economic variables helped explain the long run behaviour of an economic system, and discriminate among alternative models. The standard approach for such cointegration analysis is to apply one of the cointegration tests that have been proposed in theoretical econometrics literature (see Johansen(1988), Philips(1991), Park(1992), Saikkonen(1992), Shin(1994) among others). However, in some empirical research the main purpose of study is merely to analyse the properties of a multivariate system for a given cointegrating rank. The determination of cointegrating rank itself and its statistical significance only have secondary importance. In such cases, it is more appropriate to choose the cointegrating rank based on a consistent criterion rather than a test statistics to avoid the asymptotic TYPE-I errors. This paper modifies the order selection criteria proposed in the earlier literature to cointegrating systems in a way that the order of the system and the cointegrating rank are selected simultaneously and consistently.

We consider an m-dimensional autoregressive (AR) process $\{X_t\}$ given by

$$A(L)X_t = \left(I_m - \sum_{j=1}^{p} A_j L^j\right)X_t = \gamma D_t + \varepsilon_t \qquad (1.1)$$

where $|A(z)| = 0$ has r<m roots outside the unit circle, and the remaining (m-r) roots have modulus one. We assume ε_t are serially uncorrelated and normally distributed with a positive definite covariance matrix Ω_ε. D_t is the set of explanatory variables with finite second moments. It may, for example, include seasonal dummies, or a constant intercept. It could also include I(0) stochastic variables that are uncorrelated with the disturbances. One can generalise D_t also to accommodate for variables that are higher order of magnitude. Though it complicates the algebra considerably, and left out of this paper. This formulation characterises the stochastic component of variables in X_t as integrated of order one, I(1), and the system as cointegrated, CI(1,0), with cointegration rank r.

The selection criteria considered in this paper are in the form of

$$\Phi(k,s) = \log\left(|\hat{\Omega}(k,s)|\right) + \kappa(k)\frac{f(T)}{T} + \sigma(s)\frac{g(T)}{T} \tag{1.2}$$

for $k = 0,1,2,......,K$ and $s = 0,1,...,m$. Here $\hat{\Omega}(k,s)$ is the maximum likelihood estimate of the residual covariance matrix, Ω_ε, for a fit of order k and cointegrating rank s. K is a finite number that is *apriori* believed to be greater than the actual order `p' of the autoregressive system. The estimates for (p,r) are obtained by minimising the criterion in (1.2) with respect to the pair (k,s).We show in the next section that these estimates will be weakly consistent if $K \geq p$ and the functions $\kappa(.)$, $\sigma(.)$, $f(.)$ and $g(.)$ are strictly increasing so that $f(T) \to \infty$, $g(T) \to \infty$, $f(T)/T \to 0$, and $g(T)/T \to 0$. This formulation is quite general, and several widely used order selection criteria such as Akaike(1973), Hannan and Quinn (1979), and Schwarz (1978) are only its particular cases. With the exception of the last additive term for cointegrating rank, a similar formulation is studied by Paulsen (1984).

4.2 Simultaneous Determination of Order with Cointegration Rank:

The model in (1.1) can be re-written as

$$\left(I_m - \sum_{j=1}^{p-1} \Gamma_j L^j\right) \Delta X_t = \gamma D_t + \Gamma_p X_{t-p} + \varepsilon_t \qquad (2.1)$$

where Δ is the first difference operator, and $\Gamma_j = -I_m + \sum_{i=1}^{j} A_i$ for j = 1,2,.,k.

We assume there is no multicointegration as studied by Granger and Lee (1989), so that the rank of Γ_p coincides with the cointegration rank, r. The formulation in (2.1) is called the error correction mechanism (ECM) representation, and it is more convenient to study the cointegrating properties of the system. It also enables us for a direct comparison of our results with the earlier cointegration tests since it has been the basis for most previous studies.

Because Γ_p is rank $r < m$, there exist two $m \times r$ matrices α and β' such that

$$\left(I_m - \sum_{j=1}^{p-1} \Gamma_j L^j\right) \Delta X_t = \gamma D_t + \alpha\beta X_{t-p} + \varepsilon_t \qquad (2.2).$$

The maximum likelihood estimation of the parameters in (2.2) for known order and cointegrating rank are considered by Johansen (1991). In addition, he proposes a likelihood ratio test statistics that compares (2.2) with the unrestricted form of (2.1). In this paper, we too use the maximum likelihood estimation of a hypothesised model for comparability, but the main results can be shown to hold also for other estimation techniques (such as Granger and Engle (1987) two-step procedure).

The selection process requires a search of model fitting to T observations of the system for alternative values of autoregressive order and cointegrating rank. For ease of notation, we will denote a hypothesised model of order k and cointegrating rank s with M(k,s). Let $\hat{\beta}(k,s)$ be the maximum likelihood estimate of the cointegrating matrix when the model M(k,s) is fitted to the data. We denote $\hat{\Gamma}\left(h, \hat{\beta}(k,s)\right)$ and $\hat{\alpha}(h,\beta(k,s))$ as the OLS estimate of the remaining parameters of model M(h,s) for given $\hat{\beta}(k,s)$. Clearly, for h=k, $\hat{\Gamma}\left(k, \hat{\beta}(k,s)\right)$ and $\hat{\alpha}(k,\beta(k,s))$ are numerically equivalent to the maximum likelihood estimates of the remaining parameters of the model M(k,s), and will be denoted by $\hat{\Gamma}(k,s)$ and $\hat{\alpha}(k,s)$. A similar notation is adopted for the estimates of the residual covariance matrix.

$$\hat{\Omega}\left(h, \hat{\beta}(k,s)\right) = \frac{1}{T} e(h,\beta(k,s))\, e(h,\beta(k,s))' \tag{2.3}$$

where

$$e(h,\beta(k,s)) \quad = \quad \Delta x - \hat{\Gamma}\left(h, \hat{\beta}(k,s)\right) z_h - \alpha(h,\beta(k,s))\, \hat{\beta}(k,s)\, x_h .$$

Here Δx is the mxT matrix of observations of Δx_t. Similarly x_h and z_h are matrices of observations for x_{t-h} and

$$z_t(h) \quad = \quad \begin{bmatrix} D_t \\ \Delta x_{t-1} \\ \vdots \\ \Delta x_{t-h+1} \end{bmatrix}$$

respectively. As before, $\hat{\Omega}(k,s)$ and e(k,s) denote the special case when h = k. Notice that the number of observations is unchanged with different lagging. This is because we assume T is equal to the actual number observations minus the loss due to K (maximum amount) times lagging. Hence, we have the same number of observations for all values of autoregressive order. This simplifies the algebra considerably and make the comparisons among alternative models easier. In practice, however, it would be inefficient to drop out more observations than necessary. Nevertheless, our results continue to hold even if the maximum number of observations are used for each autoregressive order. This is because the effect of finite number of observations on the residual covariance matrix estimate would be at most $O_p(T^{-1})$. This will be more transparent later in the paper. We can now state the main result of this paper.

THEOREM 1. Let the time series $\{x_t\}$ be given by (1.1) with r cointegrating vectors. Also let $\hat{\Omega}(k,s)$ be the maximum likelihood estimate of the disturbance covariance matrix, Ω_ε, under the null of p=k and r=s. Then the pair (k,s) which minimises

$$\Phi_T(k,s) = \log\left(|\hat{\Omega}(k,s)|\right) + \kappa(k)\frac{f(T)}{T} + \sigma(s)\frac{g(T)}{T}$$

for $k = 0,1,2,\ldots\ldots,K$ and $s = 0,1,\ldots,m$, is a weakly consistent estimate for the pair (p,r) if $K \geq p$, and the functions $\kappa(.)$, $\sigma(.)$, f(T) and g(T) are strictly increasing so that $f(T) \to \infty$, $g(T) \to \infty$, $f(T)/T \to 0$, and $g(T)/T \to 0$.

PROOF. The proof of this theorem is considerably long. Hence we start by giving the intuition behind it first:

We show that

$$\ln|\hat{\Omega}(k,s)| < \ln|\hat{\Omega}(p,r)|$$

iff $k \geq p$ and $s \geq r$. If this is the case, hen the difference between them is $O_p(T^{-1})$. Hence the penalty term, $\kappa(k)f(T)/T + \sigma(s)g(T)/T$, dominates in selection procedure, and asymptotically chooses (p,r) since f(T) and g(T) are both unbounded. Otherwise (i.e., $k < p$ or $s < r$) the difference between the estimated covariance matrices is $O_p(1)$ so that (p,r) is again selected for asymptotically the penalty term approaches to zero.

For a formal proof, we first re-state an hypothesised model M(k,s) in its ECM format;

$$\Delta X_t = \gamma D_t + \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{p+1} \Delta X_{t-k+1} + \alpha_s \beta_s X_{t-k} + \varepsilon_t \qquad (2.4)$$

There are eight cases to be compared with the true model M(p,r):

1. $k > p,\ s < r$
2. $k > p,\ s > r$
3. $k > p,\ s = r$
4. $k = p,\ s < r$
5. $k = p,\ s > r$
6. $k < p,\ s < r$
7. $k < p,\ s > r$
8. $k < p,\ s = r$

The strategy is as follows; First, we show in Lemma-1 below that if $k \geq p$, then

$$\Pr\{\ \Phi_T(k,s) > \Phi_T(k,r)\ \} \to 1 \quad \text{as } T \to \infty.$$

Next, Lemma-2 shows that, for $k > p$

$$\Pr\{\ \Phi_T(k,r) > \Phi_T(p,r)\ \} \to 1 \quad \text{as } T \to \infty.$$

Therefore these two Lemmas eliminate cases 1,2,3,4, and 5. For the remaining three cases 6,7, and 8, k is less than p. Clearly,

$$\ln|\hat{\Omega}(k,m)| \ > \ \ln|\hat{\Omega}(p,m)|$$

and it suffices to show that the difference is $O_p(1)$. This is done in Lemma 3.

LEMMA 1. Assume that M(p,r) is the true data generating mechanism (DGM) in (1.1). Then, for all $k \geq p$

$$\Pr\{\ \Phi_T(k,s) > \Phi_T(k,r)\ \} \to 1 \quad \text{as } T \to \infty.$$

PROOF OF LEMMA-1. The proof of this lemma is based on the earlier results developed in Johansen (1991). Notice that, since

$$X_{t-p} = \Delta X_{t-p} + \cdots + \Delta X_{t-k+1} + X_{t-k},$$

we can re-state the true DGM as

$$\begin{aligned} \Delta X_t &= \gamma D_t + \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{p-1} \Delta X_{t-p+1} \\ &\quad + \alpha\beta \Delta X_{t-p} + \cdots + \alpha\beta \Delta X_{t-k+1} + \alpha\beta X_{t-k} + \varepsilon_t \quad (2.5) \end{aligned}$$

Hence, for k > p, the models M(k,r) are not mis-specified though inefficient if the restrictions on the coefficients of ΔX_{t-p} to ΔX_{t-k+1} are not imposed in estimation.

It is shown in Johansen (1991) that

$$\ln|\hat{\Omega}(k,s)| = \ln|S_{00}(k)| + \sum_{i=1}^{s} \ln(1-\lambda_i)$$

where λ_i are the ordered eigenvalues of $S_{k0}\, S_{00}^{-1}(k)\, S_{0k}$ with respect to S_{kk}. Here, following the same notation,

$$S_{00} = \frac{1}{T} \Delta x \left(I - z_k'(z_k z_k')^{-1} z_k\right) \Delta x'$$

$$S_{0k} = \frac{1}{T} \Delta x \left(I - z_k'(z_k z_k')^{-1} z_k\right) x_k'$$

$$S_{k0} = \frac{1}{T} x_k \left(I - z_k'(z_k z_k')^{-1} z_k\right) \Delta x'$$

$$S_{kk} = \frac{1}{T} x_k \left(I - z_k'(z_k z_k')^{-1} z_k\right) x_k'.$$

where x_k, z_k, and Δx are as defined on page 142. Furthermore, he shows that for the true DGM model in (2.5) the smallest m-r eigenvalues are $O_p(T^{-1})$ and their precise asymptotic distribution is dependent on the specification of the dummy matrix D_t. The remaining biggest r eigenvalues are positive and asymptotically strictly grater than zero. Hence, for s > r and k ≥ p, we have

$$\ln|\hat{\Omega}(k,r)| - \ln|\hat{\Omega}(k,s)| = -\sum_{i=r+1}^{s} \ln(1-\lambda_i)$$

$$= \sum_{i=r+1}^{s} \lambda_i + \sum_{i=r+1}^{s} O_p(\lambda_i^2)$$

$$= O_p(T^{-1})$$

Therefore,

$$\Phi_T(k,s) = \ln|\hat{\Omega}(k,s)| + \kappa(k)\frac{f(T)}{T} + \sigma(s)\frac{g(T)}{T}$$

$$= \Phi_T(k,r) + (\sigma(s)-\sigma(r))\frac{g(T)}{T} + o_p\left(\frac{g(T)}{T}\right)$$

Thus, for $k \geq p$ and $s > r$, $\Pr\{\Phi_T(k,s) > \Phi_T(k,r)\} \to 1 \quad \text{as } T \to \infty$ as desired. Similarly, for $s < r$, we have

$$\ln|\hat{\Omega}(k,s)| - \ln|\hat{\Omega}(k,r)| = -\sum_{i=s+1}^{r} \ln(1-\lambda_i) > 0$$

$$= O_p^+(1)$$ [30]

Therefore

$$\Phi_T(k,s) = \ln|\hat{\Omega}(k,s)| + \kappa(k)\frac{f(T)}{T} + \sigma(s)\frac{g(T)}{T}$$

$$= \Phi_T(k,r) + O_p^+(1) + (\sigma(s)-\sigma(r))\frac{g(T)}{T}.$$

Thus, for $k \geq p$ and $s < r$, $\Pr\{\Phi_T(k,s) > \Phi_T(k,r)\} \to 1 \quad \text{as } T \to \infty$. This concludes the proof of Lemma-1.

REMARK- Notice that Lemma 1 already establishes a consistent criterion to select the cointegrating rank for a given upper bound of the autoregressive order that is grater than the order of the true DGM.

---

30 $O_p^+(.)$ represents a positive definite matrix of the specified order of magnitude.

LEMMA 2. Assume that M(p,r) is the true DGM in (1.1). Then, for $k > p$

$$\Pr\{\ \Phi_T(k,r) > \Phi_T(p,r)\ \} \rightarrow 1 \quad \text{as } T \rightarrow \infty .$$

PROOF OF LEMMA-2. The proof of this lemma is done in three steps. First we show that the stochastic difference between $\hat{\Omega}\left(k,\hat{\beta}(p,r)\right)$ and $\hat{\Omega}(p,r)$ is at most $O_p(T^{-1})$. To see this, notice that

$$\hat{\Omega}\left(k,\hat{\beta}(p,r)\right) - \hat{\Omega}(p,r) \;=\; \left(\hat{\Psi} - \tilde{\Psi}\right)\frac{v_p \cdot v_p'}{T}\left(\hat{\Psi} - \tilde{\Psi}\right)' \tag{2.6}$$

where

$$v_p \;=\; \begin{bmatrix} z(k) \\ \hat{\beta}(p,r)\, x_k \end{bmatrix}$$

$$\hat{\Psi} \;=\; \left[\ \hat{\Gamma}\left(k,\hat{\beta}(p,r)\right) \quad \hat{\alpha}\left(k,\hat{\beta}(p,r)\right)\right]$$

$$\tilde{\Psi} \;=\; \left[\ \hat{\Gamma}(p,r) \quad \hat{\alpha}(p,r)\hat{\beta}(p,r) \quad \cdots \quad \hat{\alpha}(p,r)\ \right].$$

Theorem-A.1 in the appendix shows that both $\hat{\Psi}$ and $\tilde{\Psi}$ converges to the true coefficients of the model M(p,r), $\Psi \;=\; \left[\ \Gamma_1 \quad \cdots \quad \Gamma_p \quad \alpha\beta \quad \cdots \quad \alpha\ \right]$, at rate $O_p(T^{-1/2})$. Furthermore, by straight forward application of the results in Philips and Durlauf (1986),

$$\frac{v_p \cdot v_p'}{T} \;\rightarrow\; \Sigma$$

where Σ is a positive semidefinite matrix. Consequently, the right hand side of (2.6) is at most $O_p(T^{-1})$.

The next step to prove this lemma is to show that the order of the stochastic difference between $\hat{\Omega}\left(k,\hat{\beta}(p,r)\right)$ and $\hat{\Omega}(k,r)$ is at most $O_p(T^{-1})$. Notice that

$$\hat{\Omega}\left(k,\hat{\beta}(p,r)\right) = \frac{1}{T}(v+\varepsilon)\left(I_T - v_p'(v_p v_p')^{-1} v_p\right)(v+\varepsilon)' \qquad (2.7)$$

$$\hat{\Omega}(k,r) = \frac{1}{T}(v+\varepsilon)\left(I_T - v_k'(v_k v_k')^{-1} v_k\right)(v+\varepsilon)' \qquad (2.8)$$

where

$$v = \begin{bmatrix} z(k) \\ \beta\, x_k \end{bmatrix}$$

$$v_k = \begin{bmatrix} z(k) \\ \hat{\beta}(k,r)\, x_k \end{bmatrix}.$$

The Lemma-A.1 in the appendix shows that one can reorganise the variables on the right hand side of (2.7) and (2.8) to obtain

$$\hat{\Omega}(k,r) = \frac{1}{T}\varepsilon\varepsilon' + O_p(T^{-1})$$

$$\hat{\Omega}\left(k,\hat{\beta}(p,r)\right) = \frac{1}{T}\varepsilon\varepsilon' + O_p(T^{-1}).$$

Hence the stochastic difference between $\hat{\Omega}\left(k,\hat{\beta}(p,r)\right)$ and $\hat{\Omega}(k,r)$ is at most $O_p(T^{-1})$ as desired. Consequently,

$$\hat{\Omega}(k,r) = \hat{\Omega}\left(k,\hat{\beta}(p,r)\right) + O_p(T^{-1})$$

$$\hat{\Omega}(k,r) = \hat{\Omega}(p,r) + O_p(T^{-1})$$

$$\hat{\Omega}(k,r) = \hat{\Omega}(p,r)\left(I_m + O_p(T^{-1})\right).$$

Hence,

$$\begin{aligned} \Phi_T(k,r) &= \ln|\hat{\Omega}(k,r)| + \kappa(k)\frac{f(T)}{T} + \sigma(r)\frac{g(T)}{T} \\ &= \Phi_T(p,r) + \left(\kappa(k)-\kappa(p)\right)\frac{f(T)}{T} + \ln|I_m + O_p(T^{-1})| \\ &= \Phi_T(p,r) + \left(\kappa(k)-\kappa(p)\right)\frac{f(T)}{T} + o_p\left(\frac{f(T)}{T}\right) \end{aligned}$$

Thus, for $k > p$, $\Pr\{\Phi_T(k,r) > \Phi_T(p,r)\} \to 1$ as $T \to \infty$. This concludes the proof of Lemma-2.

LEMMA 3. Assume that M(p,r) is the true DGM in (1.1). Then, for $k < p$

$$\Pr\{\Phi_T(k,s) > \Phi_T(p,r)\} \to 1 \quad \text{as } T \to \infty \ \forall s.$$

PROOF OF LEMMA-3. Notice that

$$\hat{\Omega}(k,m) - \hat{\Omega}(p,m) = \left(\hat{\Pi} - \tilde{\Pi}\right)\frac{\omega_p \cdot \omega_p'}{T}\left(\hat{\Pi} - \tilde{\Pi}\right)' \tag{2.9}$$

where

$$\omega_p = \begin{bmatrix} z_p \\ x_p \end{bmatrix}$$

$$\hat{\Pi} = \begin{bmatrix} \hat{\Gamma}(p,m) & \hat{\beta}(p,m) \end{bmatrix}$$

$$\tilde{\Pi} = \begin{bmatrix} \hat{\Gamma}(k,m) & \hat{\beta}(k,m) & \cdots & \hat{\beta}(k,m) \end{bmatrix}.$$

Therefore, the order of the stochastic difference between $\hat{\Omega}(k,m)$ and $\hat{\Omega}(p,m)$ is less than $O_p(1)$ iff the right hand side of (2.9) asymptotically converges to zero. Because $k < p$, $\frac{z_p \cdot z_p'}{T}$ converges to a positive definite matrix, this is possible iff

$$\hat{\Gamma}(p,m) - \left[\hat{\Gamma}(k,m) \quad \hat{\beta}(k,m) \quad \cdots \quad \hat{\beta}(k,m)\right] \to 0 \text{ as } T \to \infty.$$

However this implies

$$\Gamma_j = \alpha\beta \quad \text{for } p \geq j \geq k$$

which contradicts to the fact that the true autoregressive order is p. Hence, the right hand side of (2.9) is $O_p(1)$ and asymptotically converges to a positive semidefinite matrix[31]. Therefore, by a straight forward application of a theorem from Magnus and Neudecker (1988)[32],

$$\ln|\hat{\Omega}(k,m)| > \ln|\hat{\Omega}(p,m)|,$$

and the order of the stochastic difference is $O_p(1)$. Consequently, because

$$\ln|\hat{\Omega}(k,s)| \geq \ln|\hat{\Omega}(k,m)|,$$

we obtain

$$\begin{aligned}\Phi_T(k,s) &= \ln|\hat{\Omega}(k,s)| + \kappa(k)\frac{f(T)}{T} + \sigma(s)\frac{g(T)}{T} \\ &= \Phi_T(p,m) + O_P^+(1) + (\sigma(s)-\sigma(m))\frac{g(T)}{T}.\end{aligned}$$

By combining this with the earlier result of Lemma-1 that

$$\Phi_T(p,m) = \Phi_T(p,r) + (\sigma(m)-\sigma(r))\frac{g(T)}{T} + o_p\left(\frac{g(T)}{T}\right),$$

we obtain, for $k < p$ and for all s

$$\Phi_T(k,s) = \Phi_T(p,r) + O_P^+(1).$$

Hence, $\Pr\{\Phi_T(k,s) > \Phi_T(p,r)\} \to 1$ as $T \to \infty$ if $k < p$. This concludes both the proof of Lemma-3 and the proof of Theorem-1.

4.3 Simulations:

Theorem-1 is quite general and the results hold for a wide range of cost functions, $\kappa(k)f(T)/T + \sigma(s)g(T)/T$. In practice, however, only three

---

31 More precisely it is at least $O_p(1)$. But, as a straight forward application of the results that are established in Philips and Durlauf (1986), $\hat{\Pi}\frac{\omega_p\omega_p'}{T}\hat{\Pi}$, $\tilde{\Pi}\frac{\omega_p\omega_p'}{T}\hat{\Pi}$, and $\tilde{\Pi}\frac{\omega_p\omega_p'}{T}\tilde{\Pi}$ are all $O_p(1)$. Hence the right hand side of (2.9) is also at most $O_p(1)$ since it is just a linear combination of these three terms and their transpose.

32 Theorem 22 on page 21 in Chapter 1. We include the theorem in the appendix for completeness.

alternatives are generally used for the selection of the autoregressive order. In this paper, we too concentrate on these three alternatives:

1. $\text{G-AIC}(k,s) = \ln|\hat{\Omega}(k,s)| + m^2(k-1)\frac{2}{T} + (2ms - s^2)\frac{2}{T}$

2. $\text{G-HQC}(k,s) = \ln|\hat{\Omega}(k,s)| + m^2(k-1)\frac{2\ln\ln T}{T} + (2ms - s^2)\frac{2\ln\ln T}{T}$

3. $\text{G-BIC}(k,s) = \ln|\hat{\Omega}(k,s)| + m^2(k-1)\frac{\ln T}{T} + (2ms - s^2)\frac{\ln T}{T}$

Here m is the size of the multivariate system. The first is a generalisation of the criterion that is known as the Akaike Information Criterion, AIC, and proposed by Akaike (1973 and 1974). Although it does not satisfy the conditions of Theorem-1, it is included in our comparative simulation exercise for its widespread use. The second is a generalisation of the criterion originally proposed by Hannan and Quinn (1979). It is strongly consistent for $m > 1$, and based on the low of iterative logarithm. The third one is a generalisation of BIC criterion that is first studied by Schwarz (1978) and probably most popular in practical applications.

Since the results of Theorem-1 are of asymptotic nature, some Monte Carlo simulation experiments are carried out for measuring the relative performance of these criteria in finite samples.

The process we study is a simple bivariate system:

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

where $\begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$ are independently normally distributed with identity covariance matrix. We experiment with sample sizes 60, 80, 100, and 200 for each criteria for 10000 times. For comparative reasons, we also estimate the cointegrating rank with Johansen's maximum likelihood procedure for the given autoregressive order that are obtained with the same criteria for s=m. The results are summarised in Tables 4.1 to 4.4.

Table-4.1 contains the results for the sample size of 60 observations. For each (s,k) pair, two entries are listed. The first is the number of trials that the criterion selects the corresponding cell as the estimate for the cointegrating rank and the autoregressive order. The second number is analogous except the autoregressive order is selected by the same criterion for s=m, and the cointegrating rank is estimated with Johansen's maximum likelihood procedure for that order. According to this simulation exercise, the best performing method is the generalised-Schwarz-criterion, G-BIC. It selects the true cointegrating rank with the true autoregressive order in 95.7% of the cases. Its overall performance for selecting the correct cointegrating rank is about 96%. On the other hand, not very surprisingly, the performance of the generalised-Akaike-Information-criterion, G-AIC, is quite poor (it selects the correct cointegrating rank with correct order for less than 50% of the cases). This is mainly due to its inconsistency as a selection criterion.

The performance of the Johansen's maximum likelihood procedure is as what would be expected from the theory. It selects the correct cointegrating rank for about 92% to 94% of the cases. This is somewhat less than the 5% critical level due to a small sample bias. A similar finding about the over-rejection of the Johansen's procedure in small samples is also reported by Cheung and Lai (1993). Tables 4.2 to 4.4 are analogous with different sample sizes, and the results are similar to that of Table-4.1. In all cases the G-BIC outperforms the others. This is probably due to its stronger penalty term for over-parametrization. As it is shown in the proof of Theorem-1, the order of magnitude between the likelihood functions of just and over-parametrized models is $O_p(T^{-1})$.

Table-4.1

T=60 10000 simulations							
	s↓	k→ 1		2		3	
G-AIC	0	4823	7570	158	1145	4	538
	1	0	349	0	135	0	113
	2	3593	96	888	30	534	24
G-HQC	0	7945	8829	133	445	4	92
	1	959	409	0	65	0	24
	2	743	111	167	14	49	11
G-BIC	0	9568	9368	32	65	0	2
	1	316	431	0	12	0	1
	2	78	115	6	5	0	1

Table-4.2

T=80 10000 simulations							
	s↓	k→ 1		2		3	
G-AIC	0	4862	7756	175	1121	3	475
	1	0	341	0	108	0	80
	2	3676	86	831	19	453	14
G-HQC	0	8275	9021	104	362	7	53
	1	869	403	0	39	0	13
	2	625	100	96	7	24	2
G-BIC	0	9725	9414	13	50	0	3
	1	217	423	0	8	0	0
	2	40	100	5	2	0	0

Table-4.3

T=100 10000 simulations							
	s↓	k→ 1		2		3	
G-AIC	0	4859	7814	163	1048	6	509
	1	0	367	0	77	0	61
	2	3770	97	748	18	454	9
G-HQC	0	8497	9098	100	283	4	42
	1	776	427	0	29	0	5
	2	547	109	58	4	18	3
G-BIC	0	9771	9416	13	27	0	0
	1	178	440	1	4	0	0
	2	37	112	0	1	0	0

Table-4.4

T=200 10000 simulations							
	s↓	k→ 1		2		3	
G-AIC	0	4996	7945	159	1079	6	443
	1	0	321	0	66	0	42
	2	3726	78	727	17	386	9
G-HQC	0	9006	9318	70	175	2	12
	1	605	385	0	13	0	2
	2	292	95	23	0	2	0
G-BIC	0	9928	9501	2	10	0	0
	1	63	389	0	3	0	0
	2	7	97	0	0	0	0

Therefore, a good selection criteria should contain a penalty term that exceeds some critical level even for smaller samples. This can be done by either a faster rate of divergence of the functions f(T) and g(T), or by strengthening the penalty for excess parametrization by changing the functions $\kappa(.)$ and $\sigma(.)$. Intuitively however, there are potential benefits of preferring the later over the former. This is because, the faster the functions f(T) and g(T), the higher the chances of selecting a mis-specified (under-parametirized) model for relatively larger samples. From this point of view, the appropriate choice for the functions f(T) and g(T) is $2\ln\ln T$ as proposed by Hannan and Quinn (1979). This is because it is the slowest rate of divergence of the cost functions while still maintaining the strong consistency property of the criteria. From previous experience with autoregressive selection criteria, it seems the best choice for the functions $\kappa(.)$ and $\sigma(.)$ is such that they give extra degrees of freedom between the competing models. This has been the basis of our choice for the criteria studied above. This way the cost function and the stochastic gain from over-parametrization become comparable in magnitude even for small samples. On the other hand, in systems with integrated variables the asymptotic distribution of likelihood ratio statistics is not standard χ^2-distribution with extra degrees of freedom. Nonetheless, Johansen (1988) shows that $(0.85 - 0.58/\mathrm{df})\chi^2(\mathrm{df})$ is a good approximation. Here df is the extra degrees of freedom between competing models. Hence, we also examine the following criterion for the simultaneous selection of the cointegrating rank with autoregressive order;

$$\Phi_T(k,s) = \ln|\hat{\Omega}(k,s)| + m^2(k-1)\frac{2\ln\ln T}{T} + 1.7\left(2ms - s^2\right)\frac{2\ln\ln T}{T}$$

This is the same as G-HQC except for the penalty term for excess parametrization of the cointegrating rank is multiplied by 1.7. Its comparative small sample performance is reported in Table-4.5.

According to Table-4.5, this modified version of G-HQC out-performs all the others in selecting cointegrating rank with autoregressive order. Even for the small sample size of 60 observations, the correct cointegrating rank is selected for 98.4% of the cases. This confirms with our earlier intuitive discussion about the desirable qualities of a selection criteria.

Table-4.5

$$\Phi_T(k,s) = \ln|\hat{\Omega}(k,s)| + m^2(k-1)\frac{2\ln\ln T}{T} + 1.7\left(2ms - s^2\right)\frac{2\ln\ln T}{T}$$

10000 simulations

	s ↓	k → 1		2		3	
T=60	0	9630	8829	197	445	12	92
	1	124	409	1	65	0	24
	2	22	111	12	14	2	11
T=80	0	9712	9021	153	362	17	53
	1	102	403	2	39	0	13
	2	7	100	5	7	2	2
T=100	0	9773	9098	128	283	14	42
	1	67	427	2	29	0	5
	2	15	109	1	4	0	3
T=200	0	9874	9318	83	175	2	12
	1	36	385	1	13	0	2
	2	4	95	0	0	0	0

4.4 Conclusions

In some empirical research, the main purpose of analysis is to study the properties of data for a given model. In such cases, using consistent selection criteria is more appropriate than choosing the model based on a test statistics. This paper generalises the autoregressive order selection criteria for its simultaneous selection with cointegrating rank. We consider four alternatives. The Monte Carlo simulations indicate that a modified version of the Hannan-Quinn criterion out-performs the other techniques considered in this paper. However, the simulations also indicate that the performance of criteria is quite sensitive to the specification of the penalty terms, and more research is needed to generalise the results obtained in this paper.

4.5 Appendix

THEOREM A.1.

Consider the DGM in (2.2). Let $\hat{\beta}$ be a T-consistent estimate of β. Furthermore, let $\hat{\Psi}$ be the OLS estimate of the remaining coefficients of the model M(k,r) that are obtained by using $\hat{\beta}$. Then, for $k \geq p$

$$\sqrt{T}\,(\hat{\psi} - \psi) \xrightarrow{d} N(0, \Omega_{\Psi})$$

where $(\hat{\psi} - \psi) = \text{vec}(\hat{\Psi} - \Psi)$, and $\Omega_{\Psi} = \Sigma^{-1} \otimes \Omega_{\varepsilon}$. Here

$$\Psi = \begin{bmatrix} \Gamma_1 & \cdots & \Gamma_p & \alpha\beta & \cdots & \alpha \end{bmatrix},$$

$$\Sigma^{-1} = \text{plim}\frac{\hat{v} \cdot \hat{v}'}{T}$$

where $\hat{v}$ is the matrix of observations for

$$\hat{v}_t = \begin{bmatrix} D_t \\ \Delta x_{t-1} \\ \vdots \\ \Delta x_{t-k+1} \\ \hat{\beta} x_{t-k} \end{bmatrix}.$$

PROOF OF THEOREM A.1:

For $k \geq p$, the DGM can be re-stated by

$$\Delta x_t = \Psi v_t + \varepsilon_t$$

where

$$. \; v_t = \begin{bmatrix} D_t \\ \Delta x_{t-1} \\ \vdots \\ \Delta x_{t-k+1} \\ \beta x_{t-k} \end{bmatrix}.$$

The OLS estimate of Ψ is given by

$$\hat{\Psi} = \Delta x\, \hat{v}'(\hat{v}\hat{v}')^{-1}$$
$$= \Psi v \hat{v}'(\hat{v}\hat{v}')^{-1} + \varepsilon \hat{v}'(\hat{v}\hat{v}')^{-1}$$

One can re-organise the variables on the right hand side to obtain

$$\hat{\Psi} = \Psi - \Psi\frac{(\hat{v}-v)\hat{v}'}{T}\left(\frac{\hat{v}\hat{v}'}{T}\right)^{-1} + \frac{\varepsilon v'}{T}\left(\frac{\hat{v}\hat{v}'}{T}\right)^{-1} + \frac{\varepsilon(\hat{v}-v)'}{T}\left(\frac{\hat{v}\hat{v}'}{T}\right)^{-1}$$

Hence

$$\begin{aligned}
\sqrt{T}(\hat{\psi}-\psi) \quad &= \quad -\left(\left(\frac{\hat{v}\cdot\hat{v}'}{T}\right)^{-1} \otimes \Psi\right)\mathrm{vec}\left(\frac{(\hat{v}-v)\cdot\hat{v}'}{\sqrt{T}}\right) \\
&+ \quad \left(\left(\frac{\hat{v}\cdot\hat{v}'}{T}\right)^{-1} \otimes I\right)\mathrm{vec}\left(\frac{\varepsilon\cdot(\hat{v}-v)'}{\sqrt{T}}\right) \qquad \text{(A.1)} \\
&+ \quad \left(\left(\frac{\hat{v}\cdot\hat{v}'}{T}\right)^{-1} \otimes I\right)\mathrm{vec}\left(\frac{\varepsilon\cdot v'}{\sqrt{T}}\right)
\end{aligned}$$

Notice that

$$\hat{v}-v = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ (\hat{\beta}-\beta)x_k \end{bmatrix}.$$

Thus, because $(\hat{\beta}-\beta)$ is $O_p(T^{-1})$.

$$\frac{(\hat{v}-v)\hat{v}'}{\sqrt{T}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \left(\dfrac{(\hat{\beta}-\beta)x_k\,\hat{v}'}{\sqrt{T}}\right) \end{bmatrix} \to 0,$$

and

$$\frac{\varepsilon(\hat{v}-v)'}{\sqrt{T}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \left(\frac{\varepsilon x_k'(\hat{\beta}-\beta)'}{\sqrt{T}}\right) \end{bmatrix} \rightarrow 0.$$

Furthermore, we have

$$\frac{\hat{v}\cdot\hat{v}'}{T} \rightarrow \Sigma.$$

Consequently, the first two terms on the right hand side of (A.1) converges to zero asymptotically. Moreover

$$\mathrm{vec}\left(\frac{\varepsilon\cdot v'}{\sqrt{T}}\right) = \left(\frac{v}{\sqrt{T}}\otimes I_m\right)\mathrm{vec}(\varepsilon) \rightarrow N(0,\Sigma\otimes\Omega_\varepsilon).$$

Hence, the last term on the right hand side of (A.1) converges in distribution to $N(0,\Sigma^{-1}\otimes\Omega_\varepsilon)$. This concludes the proof of Theorem A.1.

LEMMA A.1.

Consider the DGM in (2.2). Let

$z_t(h)=\begin{bmatrix} D_t \\ \Delta x_{t-1} \\ \vdots \\ \Delta x_{t-h+1} \end{bmatrix}$, $\hat{v}=\begin{bmatrix} z(k) \\ \hat{\beta}\, x_k \end{bmatrix}$, and $v=\begin{bmatrix} z(k) \\ \beta\, x_k \end{bmatrix}$. Here $\hat{\beta}$ is a T-consistent estimate of the cointegrating matrix, β. Then,

$$\frac{1}{T}(v+\varepsilon)\left(I_T-\hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\right)(v+\varepsilon)' = \frac{1}{T}\varepsilon\varepsilon' + O_p(T^{-1}).$$

PROOF OF LEMMA A.1:

$$\frac{1}{T}(v+\varepsilon)\left(I_T-\hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\right)(v+\varepsilon)' = \frac{1}{T}v\left(I_T-\hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\right)v'$$
$$+ \frac{1}{T}\varepsilon\left(I_T-\hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\right)v'$$
$$+ \frac{1}{T}v\left(I_T-\hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\right)\varepsilon' \tag{A.2}$$
$$+ \frac{1}{T}\varepsilon\left(I_T-\hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\right)\varepsilon'$$

Notice that

$$\frac{1}{T}v\left(I_T-\hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\right)v' = \frac{1}{T}v(v-\hat{v})' - v\hat{v}'(\hat{v}\hat{v}')^{-1}\frac{1}{T}\hat{v}(v-\hat{v})'.$$

Furthermore, because $\hat{\beta}$ is T-consistent,

$$\frac{1}{T}v(v-\hat{v})' = \frac{1}{T}vx_k'(\hat{\beta}-\beta)' = O_p(T^{-1}),$$

$$\frac{1}{T}\hat{v}(v-\hat{v})' = -\frac{1}{T}(\hat{\beta}-\beta)x_kx_k'(\hat{\beta}-\beta)' + \frac{1}{T}vx_k'(\hat{\beta}-\beta)' = O_p(T^{-1})$$

$$v\hat{v}'(\hat{v}\hat{v}')^{-1} = I + (v-\hat{v})\hat{v}'(\hat{v}\hat{v}')^{-1} = I + O_p(T^{-1})$$

Hence the first term of the right hand side in (A.2) is $O_p(T^{-1})$.

Similarly,

$$\frac{1}{T}\varepsilon\left(I_T-\hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\right)v' = \frac{1}{T}\varepsilon(v-\hat{v})' - \varepsilon\hat{v}'(\hat{v}\hat{v}')^{-1}\frac{1}{T}\hat{v}(v-\hat{v})',$$

and

$$\frac{1}{T}\varepsilon(v-\hat{v})' = \frac{1}{T}\varepsilon x_k'(\hat{\beta}-\beta)' = O_p(T^{-1}),$$

$$\varepsilon\hat{v}'(\hat{v}\hat{v}')^{-1} = \varepsilon v'(\hat{v}\hat{v}')^{-1} - \varepsilon(v-\hat{v})'(\hat{v}\hat{v}')^{-1} = O_p(T^{-1/2}).$$

Hence the second and the third terms of the right hand side in (A.2) are both at most $O_p(T^{-1})$.

It remains to show that

$$\frac{1}{T}\varepsilon\left(I_T - \hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\right)\varepsilon' \;=\; \frac{1}{T}\varepsilon\varepsilon' - \frac{1}{T}\varepsilon\hat{v}'(\hat{v}\hat{v}')^{-1}\hat{v}\varepsilon'$$

is $\frac{1}{T}\varepsilon\varepsilon' + O_p(T^{-1})$. This simply follows from

$$\frac{1}{T}\varepsilon\hat{v}' = \frac{1}{T}\varepsilon v' - \frac{1}{T}\varepsilon x_k'(\hat{\beta} - \beta)' = O_p(T^{-1/2}).$$

THEOREM 22 (Magnus and Neudecker (1988), pp 21).

Let A be positive definite and B positive semidefinite. Then

$$|A+B| \;\geq\; |A|$$

with equality if and only if B = 0.

PROOF OF THEOREM-22.

Let Λ be a positive definite diagonal matrix such that

$$S'AS = \Lambda, \qquad S'S = I.$$

Then, $SS' = I$ and

$$A + B = S\Lambda^{1/2}(I + \Lambda^{-1/2}S'BS\Lambda^{-1/2})\Lambda^{1/2}S'$$

and hence

$$\begin{aligned} |A+B| &= |S\Lambda^{1/2}|\cdot|I + \Lambda^{-1/2}S'BS\Lambda^{-1/2}|\cdot|\Lambda^{1/2}S'| \\ &= |S\Lambda^{1/2}\Lambda^{1/2}S'|\cdot|I + \Lambda^{-1/2}S'BS\Lambda^{-1/2}| \\ &= |A|\cdot|I + \Lambda^{-1/2}S'BS\Lambda^{-1/2}| \end{aligned}$$

If B=0 then $|A+B| = |A|$. If $B \neq 0$, then the matrix $\Lambda^{-1/2}S'BS\Lambda^{-1/2}$ will be positive semidefinite with at least one positive eigenvalue. Hence $|\Lambda^{-1/2}S'BS\Lambda^{-1/2}| > 1$ and $|A+B| > |B|$.

4.6 References

Akaike, H., 1973, "Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models, Biometrika, Vol.60, pp.255-65.

Akaike, H., 1974, "A New Look At The Statistical Model Identification", IEEE Transactions on Automatic Control, AC-19, pp.716-23.

Cheung, Y.W., and Lai, K.S., 1993, "Finite-Sample Sizes of Johansen's Likelihood Ratio Tests For Cointegration", Oxford Bulletin of Economics and Statistics, Vol.55, pp.313-28.

Granger, C.W.J., and Engle, R.F., 1987, "Co-integration And Error Correction: Representation, Estimation, And Testing", Econometrica, Vol.55, pp.251-76.

Granger, C.W.J., and Lee, T., 1989, "Multicointegration", Advances in Econometrics, Vol.8, pp.71-84.

Hannan, E.J., and Quinn, B.G., 1979, "The Determination of the Order of an Autoregression", Journal of the Royal Statistical Society, Vol.B-41, pp.190-5.

Johansen, S., 1988, "Statistical Analysis of Cointegration Vectors", Journal of Economic Dynamics and Control, Vol. 12, pp.231-254.

Johansen, S., 1991, "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", Econometrica, Vol.59, pp.1551-80.

Park, J.Y., 1992, "Canonical Cointegrating Regressions", Econometrica, Vol.60, pp.119-43.

Paulsen, J., 1984, "Order Determination of Multivariate Autoregressive Time Series With Unit Roots", Journal of Time Series Analysis, Vol.5, pp.115-27.

Phillips, P.C.B., 1991, "Optimal Inference In Cointegrated Systems", Econometrica, Vol.59, pp.283-306.

Phillips, P.C.B., and Durlauf, S.N., 1986, "Multiple Time Series Regression with Integrated Processes", Review of Economic Studies, Vol.53, pp.473-95.

Quinn, B.G., 1980, "Order Determination for a Multivariate Autoregression", Journal of the Royal Statistical Society, Vol.B-42, pp.182-5.

Saikkonen, P., 1992, "Estimation and Testing of Cointegrated Systems by an Autoregressive Approximation", Econometric Theory, Vol.8, pp.1-27.

Schwarz, G., 1978, "Estimating the Dimension of A Model", Annals of Statistics, Vol.6, pp.461-4.

Shibata, R., 1980, "Asymptotically Efficient Selection Of The Order Of The Model For Estimating Parameters Of A Linear Process", Annals of Statistics, Vol.8, pp.147-64.

Shin, Y., 1994, "A Residual-Based Test of the Null of Cointegration Against the Alternative of No Cointegration", Econometric Theory, Vol.10, pp.91-115.