Neglected Heterogeneity and Dynamics in Cross-Country Savings Regressions[*]

Nadeem U. Haque
International Monetary Fund

M. Hashem Pesaran
Trinity College, Cambridge

Sunil Sharma
International Monetary Fund

January 1999

Abstract

This paper examines the extent to which conclusions of cross-country studies of private savings are robust to allowing for the possible heterogeneity of saving behaviour across countries and the inclusion of dynamics. It provides a review of the econometric implications of neglected slope heterogeneity and dynamics for the fixed effects estimators routinely used in such studies, and illustrates the nature and extent of the biases involved by a re-examination of time series data from 21 OECD countries previously analysed in the literature. The paper shows that neglecting heterogeneity and dynamics in cross-country savings regressions can lead to misleading inferences about the key determinants of savings behavior. If differences across countries are ignored, one can overestimate the influence of certain factors on the private savings rates and at the same time obtain highly significant, but spurious, nonlinear effects for some of the potential determinants. The results indicate that among the many variables considered in the literature only the fiscal variables – the general government surplus as a proportion of GDP and the ratio of government consumption to GDP– seem to be the key determinants of private savings rates in the industrial countries in the post world war II period.

JEL Classifications: E21, C23.

Key Words: Saving Behavior, Cross-Country Studies, Slope Heterogeneity, Dynamics, Panel Data Models.

1. Introduction

There are many important empirical issues that the aggregate private savings behavior literature attempts to address. The first is the still unresolved question of whether higher private savings precede increases in growth in the gross domestic product (GDP) or whether higher growth in GDP is triggered by other factors and higher private savings are merely a result of higher income levels as well as GDP growth rates. Evidence provided by Carroll and Weil (1994) suggests that growth Granger causes savings, and even taking into account the shortcomings of Granger causality, this can be interpreted as saying that higher growth precedes higher savings. This still allows for the possibility that while the initial impetus to growth should be sought in other factors, the process of (rapid) development may involve a virtuous circle whereby higher growth leads to higher domestic savings which in turn through raising domestic capital formation leads to even higher growth rates.

The second key question is whether private savings are affected by fiscal policy variables, in particular the public sector fiscal balances and the aggregate level and composition of government expenditures. Recent research seems to suggest that aggregate private savings do offset movements in public savings, although strict Ricardian equivalence does not hold. The focus of the current research has been on identifying the magnitude of this offset and evaluating the channels through which such behavior manifests itself at the microeconomic level. In this context, openness of economies and access to international markets have also been examined to understand how access to foreign savings may influence aggregate private savings.

The third set of issues is related to analyzing the effect of demographic changes and the role played by the provision of social insurance. In particular, interest has focussed on how changes in the dependency ratio are likely to affect private savings and whether such behavior is modified by the presence and design of pension systems and social security.

Other possible determinants of the savings behavior have also been considered; such as the rate of inflation and its variability, changes in the terms of trade, borrowing constraints, the depth of the financial system, and macroeconomic and political stability. Whether such factors have short-term effects or whether the influences are longer lasting has generally not been addressed in the literature. In part this is due to a lack of a theoretical framework for analyzing such factors and in part because good empirical proxies for variables such as financial depth and political and economic stability are not available.

Most recent studies have used multi-country panel data to examine the various issues raised in the literature about aggregate private savings behavior. An

[1]

early cross-country study of household saving in developing countries was provided by Schmidt-Hebbel, Webb and Corsetti (1992). Other prominent examples are Carroll and Weil (1994) mentioned above and related recent studies by Attanasio, Picci and Scorcu (1998), Ogaki, Ostry and Reinhart (1995), Edwards (1996), Masson, Bayoumi and Samiei (1995), Dayal-Gulati and Thimann (1997), and Callen and Thimann (1997). Carroll and Weil (1994) base their analysis on data from 64 countries (OECD and developing) over the period 1958-87. They report fixed-effects regressions of growth on the savings rates and *vice versa* using non-overlapping 5-year averages. Attanasio, Picci and Scorcu (1998) also focus on the growth-investment-saving inter-linkages and estimate country-specific savings/growth rates regressions for a panel of 123 industrial and developing countries over the period 1961-94. Ogaki, Ostry and Reinhart (1995) on the other hand provide an empirical examination of the interest-elasticity of savings in 13 low-income and middle-income countries over 1968-92. Dayal-Gulati and Thimann (1997) consider saving rates across 5 Asian and 9 Latin American countries over the period 1975-95, and report separate static fixed-effects regressions (estimated by least squares and instrumental variables techniques) for the Asian group and Latin American group of countries. Their results show considerable heterogeneity in the parameter estimates across the two country groupings. Callen and Thimann (1997) consider the household saving behaviour across 21 OECD countries over 1975-95, and report static fixed-effects regressions. They do not allow for the possibility of slope heterogeneity in their analysis, but attempt to take account of error variance heterogeneity and a limited degree of dynamics in some of their specifications.

The studies by Edwards (1996) and Masson, Bayoumi and Samiei (1995, 1998) consider a wider class of variables in their analysis of saving behavior. Edwards (1996) uses a 36 country data set (of which eleven are industrialized countries) over the period 1970-92, while Masson, Bayoumi and Samiei (hereafter MBS) examine the determinants of the private savings rate using time series observations on 21 OECD countries over 1971-93, and 40 developing countries over 1982-93. Both of these papers build on the extensive theoretical and empirical literature that exists on savings and consumption bevaviour. The study by MBS focuses on six possible determinants of private savings: fiscal variables (general government budget surplus), demographics (the dependency ratio measured as the ratio of the population that is under 20 and 65 and over to that aged between 20-64 years), GDP growth and level of GDP per capita relative to U.S. GDP, the interest rate, the inflation rate and changes in the terms of trade. It reports separate static panel regressions for their sample of industrial and developing countries. Based on the panel regressions and a number of associated cross-section regressions, MBS then conclude that general government surplus, the dependency ratio, GDP

growth, and the relative per capita GDP have significant effects on the private savings rate, although with differing degrees of precision. The evidence on the effects of the interest rate and the level of inflation on the savings rate is much weaker and depends on the sample of countries under consideration.

The modelling strategy in these and other cross-country studies of savings rates has generally relied on pooled OLS or static fixed-effects regressions. Little attempt has been made to introduce dynamics, differentiate short-run and long-run parameters or examine the extent to which the diversity in institutional structures across countries can be taken into account while assessing the behavior of aggregate private savings.[1] Clearly, saving and wealth accumulation is a dynamic process and static specifications are unlikely to capture essential features of such processes. Also, cross-country parameter homogeneity restrictions are, at best, likely to hold for long-run parameters with the short-run responses differing across countries. Moreover, while economic theory provides some guidance on long-run parameters it is typically silent on short-run dynamics and the exact nature of the adjustment processes. And most importantly, unless the heterogeneity across countries in short-run responses is taken into account (and allowed for to the extent permitted by the data), the estimation of parameters of interest can be contaminated, vitiating the conclusions drawn about the importance of various factors in determining aggregate private savings.

The papers mentioned above clearly address an important issue of special-concern to policy makers and academics alike. They attempt to isolate the key determinants of the saving rates across a large number of industrialized and developing countries. However, the econometric implementation generally relies on fixed-effects estimation of static panel data models, this raises the question of robustness of their results to the neglected dynamics and the possible heterogeneity across countries of short-run and long-run effects of the various likely determinants of the private savings rate. In this paper we examine the robustness of the conclusions to these two sources of misspecification using the MBS data.

Initially, we consider the consequences of allowing for slope heterogeneity, but continue to use the static specification of the savings rate equations used in the literature.[2] We examine the importance of slope heterogeneity using two different methods: (i) We explicitly model the cross-sectional variations of the slope coefficients in terms of country-specific wealth-GDP ratios and income per capita and estimate augmented fixed-effects regressions. We show that once slope hetero-

---

[1] The studies by Carroll and Weil (1994) and Attanasio, Picci and Scorcu (1998) do allow for dynamics but focus on growth-saving linkages and do not consider the effect of other possible variables such as the dependency ratio, fiscal policy variables, inflation and other variables considered in the studies by Edwards (1996) and Masson, Bayoumi and Samiei (1995, 1998).

[2] However, even the "static" models implicitly involve some dynamics through the wealth variable included in many specifications.

[3]

geneity is taken into account the quadratic relationship between private savings and income found in MBS (1998) is no longer statistically significant. The theoretical reasons behind this result are discussed in some detail, with the relevant derivations relegated to an appendix. (ii) Our second approach to the slope heterogeneity problem follows Pesaran and Smith (1995) and compares the average estimates of the slope coefficients from individual country regressions (referred to as the "mean group estimates") to those obtained using the fixed-effects procedure. We also consider different groupings of the coefficient estimates according to whether they are positive or negative and whether they are statistically significant, thus providing a simple and effective way of reporting the extent of the heterogeneity that seems to exist in the underlying parameters.

We then allow for dynamics in individual savings rate equations, and consider two types of estimators: the mean group estimator of the long-run coefficients, and a new pooled estimator that allows for short-run heterogeneity whilst restricting the long-run coefficients to be the same across groups (see Pesaran, Shin and Smith, 1999). As shown in Pesaran and Smith (1995) and Pesaran, Smith and Im (1996), in the presence of dynamics and slope heterogeneity the use of standard panel techniques, such as the fixed-effects estimator or the Anderson-Hsiao estimator, leads to inconsistent estimates and potentially misleading inferences even for large N and T panels. In contrast, the mean group and the pooled estimators of the long-run coefficients (the latter under the assumed homogeneity of the long-run parameters) are consistent when both N (the number of countries) and T (the number of time periods) are reasonably large. Our results show that only the effects of the ratio of general government budget surplus to GDP, and the ratio of general government current expenditure to GDP are reasonably robust to the presence of slope heterogeneity and yield plausible estimates for the off-setting effects of government budget surpluses and government consumption expenditures on private savings across the OECD countries.

The plan of the paper is as follows: Section 2 sets up the econometric framework used in cross-country savings studies and reviews the theoretical literature on slope heterogeneity both in static and dynamic panel data models. Section 3 provides a detailed empirical examination of the private savings equation analyzed by MBS (1995, 1998). Section 4 presents a summary of the main results, discusses their economic interpretation and their relevance for policy.

2. Cross-Country Savings Regressions: Neglected Heterogeneity

Almost all cross-country studies of savings behavior in the literature make use of linear reduced form models where the savings rate (private or national) is

explained in terms of a number of factors such as GDP growth, the real interest rate, inflation, wealth effects, fiscal indicators and demographic variables. In its most general form such a country-specific savings equation can be written as:

$$\lambda_i(L)s_{it} = \mu_i + \sum_{j=1}^{k} \beta_{ij}(L)x_{ijt} + \varepsilon_{it}, \qquad (2.1)$$

$$= \mu_i + \boldsymbol{\beta}_i'(L)\mathbf{x}_{it} + \varepsilon_{it}$$

where s_{it} is the savings rate of country $i = 1, 2, ..., N$ during period $t = 1, 2, ..., T$, $\mathbf{x}_{it} = (x_{i1t}, x_{i2t}, ..., x_{ikt})'$ is a $k \times 1$ vector of factors deemed to be important in the determination of the savings rate in country i . The parameters μ_i represent the fixed-effects, and $\lambda_i(L)$ and $\boldsymbol{\beta}_i(L)$ are the distributed lag functions of orders p and q_i defined by

$$\lambda_i(L) = 1 - \lambda_{i1}L - \lambda_{i2}L^2 - ... - \lambda_{ip_i}L^{p_i}, \qquad (2.2)$$

$$\boldsymbol{\beta}_i(L) = \boldsymbol{\beta}_{i0} + \boldsymbol{\beta}_{i1}L + ... + \boldsymbol{\beta}_{iq_i}L^{q_i}, \qquad (2.3)$$

where $\boldsymbol{\beta}_{is} = (\beta_{i1s}, \beta_{i2s}, ..., \beta_{iks})'$. Finally, ε_{it}, the disturbance term for country i, is intended to capture the effect of a multitude of (remaining) unobserved effects on the saving rate. These disturbances are assumed to be distributed independently (both over time and across countries) with a zero mean and a constant variance, σ_i^2. They are also usually assumed to be distributed independently of the regressors, $\mathbf{x}_{it}$.

For each country (2.1) represents an autoregressive distributed lag model of order $(p, q_1, q_2, ..., q_k)$, or simply $ARDL(p, q_1, q_2, ..., q_k)$. All the cross-country studies of savings rates are special cases of this $ARDL$ model. Edwards (1996) distinguishes between private and public savings, and applies a static version of (2.1) to a panel of 36 countries. For the private savings rate he uses pooled OLS which assumes that all the coefficients (slopes as well as the intercepts) are homogenous across countries. For the public savings rate equation he allows for heterogenous intercepts but continues to assume slope homogeneity. In one set of regressions Edwards includes Latin American dummy variables for each regressor in the panel; thus allowing for a limited degree of parameter heterogeneity. Such dummy variable regressions still impose error variance homogeneity across all the countries in the sample as well as parameter homogeneity within the two country groupings. MBS (1995, 1998) also use a homogenous static version of (2.1) in their analysis of private saving rates. They consider two separate country groups, one consisting of 21 OECD economies and the other 40 developing economies. The latter is then grouped into high, middle and low income countries. Therefore, they allow for the parameters to differ between the country groupings, but not across the

[5]

countries within a given group. For each country group they only allow for fixed-effects and assume that all the slope coefficients and error variances are the same across countries. MBS (1998) also experiment with a limited dynamic specification where (implicitly) a common factor restriction is imposed on the distributed lag functions $\lambda_i(L)$ and $\boldsymbol{\beta}_i(L)$. In particular they impose (without testing) the restrictions:

$$\lambda_i(L) = 1 - \lambda L, \text{ and } \beta_{ij}(L) = \beta_{j0}(1 - \lambda L), \tag{2.4}$$

for all i and j.

In an attempt to address the possible simultaneity problem, many of these cross-country regression studies also provide instrumental variable (IV) estimates using lagged values of the regressors as instruments. Such a procedure is valid in the case of static models. But in the case of the saving process where the dynamics of adjustments can play an important role the application of the IV procedure is problematic as any lagged value that is used as an instrument could itself potentially be an important factor in explaining the short dynamics of the saving process. This difficulty with the IV technique is related to the problem of identification and of restrictions in macroeconometric specifications, a problem that was considered in Sims (1980).[3]

The extent to which the results based on cross-country savings regressions are robust to neglected slope heterogeneity depends on (i) whether there are any systematic relationships between the slope coefficients and the regressors, (ii) the importance of the dynamics in the saving process, and (iii) the degree of slope heterogeneity itself. The overall outcome can only be judged in an empirical context by comparing the fixed-effects estimators reported in the literature with those estimators that allow for slope heterogeneity and dynamics. But before proceeding with the empirical analysis it is instructive to examine, at a theoretical level, the consequences of neglected slope heterogeneity on the fixed-effects estimators in the context of simple models.

2.1. Neglected Heterogeneity in Static Models

Consider the following static fixed-effects panel data model:

$$y_{it} = \mu_i + \boldsymbol{\beta}' \mathbf{x}_{it} + u_{it}, \tag{2.5}$$

where $\boldsymbol{\beta}$ is the vector of unknown slope coefficients assumed to be the same across all the countries in the panel, and u_{it} are unobserved random shocks assumed to be distributed independently (both over time and across countries) and identically

---

[3] Edwards (1996) and MBS (1995), for example, estimate their panel data model by instrumental variable (IV) techniques, using lagged values of the regressors as instruments.

with a zero mean and a constant variance, σ^2. It is also usually assumed that the errors, u_{it}, and the regressors, $\mathbf{x}_{it}$, are independently distributed. The only source of heterogeneity allowed for in this framework are the fixed-effects, μ_i.

The slope and error variance homogeneity assumption that underlie the fixed-effects formulation may be a reasonable first-order approximation for analysis of household or firm behaviour in a given locality or region, but it is less likely to hold across countries that are at different stages of their economic development, have differing institutions, customs or social norms. This is particularly true of the short run effects since institutional constraints are likely to be more binding in the short-run and hence can cause the private savings rate across countries to respond differently to its various determinants such as output growth, the real interest rate, inflation, and fiscal deficits.

The consequences of neglecting slope heterogeneity can be quite serious, particularly when the slopes vary systematically across countries. To see this consider the following simple static panel data model with cross-sectionally varying slopes and intercepts:

$$y_{it} = \mu_i + \beta_i x_{it} + u_{it}. \tag{2.6}$$

For simplicity of exposition we confine our attention to a model with a single time-varying regressor and continue to maintain the error variance homogeneity assumption. The coefficients, β_i, are allowed to vary freely across countries but are otherwise assumed to be fixed (over time). It proves useful to decompose β_i into a common component, β, and a remainder, η_i, that varies across countries:

$$\beta_i = \beta + \eta_i. \tag{2.7}$$

The nature of the slope heterogeneity can now be characterized in terms of the properties of η_i. In the panel literature, following the work of Swamy (1970), it is often assumed that η_i are identically independently distributed (i.i.d.) as well as distributed independently of x_{it} and u_{it} for all i and t, with a zero mean and a constant variance, σ_η^2. This specification is known as the random coefficients model (RCM) discussed extensively by Hsiao (1986, Ch. 6). The RCM only allows for a minimal degree of slope heterogeneity. For example, it rules out systematic relationships between β_i and the sample means or variances of the x_{it} process. To cover such possibilities we shall also consider the case where the β_i's are treated as fixed; in exactly the same way that the intercepts, μ_i, are treated in the fixed-effects model, (2.5).

Consider an investigator that ignores the heterogeneity of the slope coefficients in (2.6) and instead estimates the following augmented homogenous slope model:

$$y_{it} = \alpha_i + \delta_x x_{it} + \delta_z z_{it} + v_{it}, \tag{2.8}$$

[7]

where z_{it} is an additional regressor spuriously thought to be important by the investigator. For expositional simplicity we assume that for each i, $\mathbf{w}_{it} = (x_{it}, z_{it})'$ is covariance stationary with the variance-covariance matrix

$$\Omega_i = \begin{pmatrix} \omega_{ixx} & \omega_{ixz} \\ \omega_{izx} & \omega_{izz} \end{pmatrix}. \tag{2.9}$$

It is now easily established that the fixed-effects (FE) estimator of the slope coefficient, $\boldsymbol{\delta} = (\delta_x, \delta_z)'$, is generally inconsistent. The inconsistency of the FE estimator in this set-up is derived in the appendix. Denoting the FE estimator of $\boldsymbol{\delta}$ by $\hat{\boldsymbol{\delta}}_{FE}$, we have

$$\plim_{N,T\to\infty} \left(\hat{\delta}_{x,FE} - \beta \right) = \frac{E(\omega_{izz})Cov(\omega_{ixx}, \beta_i - \beta) - E(\omega_{ixz})Cov(\omega_{ixz}, \beta_i - \beta)}{E(\omega_{ixx})E(\omega_{izz}) - (E(\omega_{ixz}))^2},$$
$$\tag{2.10}$$

$$\plim_{N,T\to\infty} \left(\hat{\delta}_{z,FE} \right) = \frac{E(\omega_{ixx})Cov(\omega_{ixz}, \beta_i - \beta) - E(\omega_{ixz})Cov(\omega_{ixx}, \beta_i - \beta)}{E(\omega_{ixx})E(\omega_{izz}) - (E(\omega_{ixz}))^2}, \tag{2.11}$$

where[4]

$$Cov(\omega_{ixx}, \beta_i - \beta) = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{ixx}(\beta_i - \beta) \right),$$

$$Cov(\omega_{ixz}, \beta_i - \beta) = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{ixz}(\beta_i - \beta) \right),$$

$$E(\omega_{ixx}) = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{ixx} \right), \quad E(\omega_{izz}) = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{izz} \right),$$

$$E(\omega_{ixz}) = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{ixz} \right).$$

The above results have a number of interesting implications:

1. The fixed-effects estimator of δ_z is robust to slope heterogeneity if the incorrectly included regressors, z_{it}, are on average orthogonal to x_{it}, namely when $E(\omega_{ixz}) = 0$. However, in the presence of slope heterogeneity the fixed-effects estimator of δ_x continues to be inconsistent even if z_{it} and x_{it}

---

[4]In the case where the slopes are randomly distributed across i, the limits in the definitions of $Cov(\omega_{ixx}, \beta_i - \beta)$ and $Cov(\omega_{ixz}, \beta_i - \beta)$ need to be replaced by probability limits.

are on average orthogonal. The direction of the asymptotic bias of $\hat{\delta}_{x,FE}$ depends on the sign of $Cov(\omega_{ixx}, \beta_i - \beta)$. The bias of $\hat{\delta}_{x,FE}$ is positive when $Cov(\omega_{ixx}, \beta_i - \beta) > 0$ and *vice versa*.[5]

2. The fixed-effects estimators, $\hat{\boldsymbol{\delta}}_{x,FE}$ and $\hat{\delta}_{z,FE}$, are both consistent if

$$Cov(\omega_{ixz}, \beta_i - \beta) = Cov(\omega_{ixx}, \beta_i - \beta) = 0. \qquad (2.12)$$

Clearly, these conditions are met under slope homogeneity. In the present application where the regressors are assumed to be strictly exogenous, the fixed-effects estimators converge to their true values under the RCM where the slope coefficients and the regressors are assumed to be independently distributed.[6] Notice, however, that since the β_i's are assumed to be fixed over time, then any systematic dependence of η_i on $\mathbf{w}_{it}$ over time is already ruled out under model (2.6). The random coefficients assumption imposes further restrictions on the joint distribution of η_i and the cross-sectional distribution of $\mathbf{w}_{it}$. This result provides a straightforward extension of the misspecification results already familiar from the regression analysis of single series (i.e. when $N = 1$), and which forms the basis of the "general-to-specific" methodology advocated by David Hendry.[7]

3. In general, where $E(\omega_{ixz}) \neq 0$ and $Cov(\omega_{ixz}, \beta_i - \beta) \neq 0$ and/or $Cov(\omega_{ixx}, \beta_i - \beta) \neq 0$, the fixed-effects estimators, $\hat{\delta}_{x,FE}$ and $\hat{\delta}_{z,FE}$, are both inconsistent. Therefore, if the slope coefficients are fixed but vary systematically across the groups, the application of the general-to-specific methodology to standard panel data models can lead to misleading results (spurious inference). An important example is provided by the case when attempts are made to check for the presence of nonlinearities by testing the significance of quadratic terms in static panel data models using fixed-effects estimators. In the context of our simple specification this would involve setting $z_{it} = x_{it}^2$, and a test of the significance of z_{it} in (2.8) will yield sensible results only if the conditions defined by (2.12) are met. In general, it is possible to falsely reject the linearity hypothesis when there are systematic relations between the slope coefficients and the cross-sectional distribution of the regressors. Therefore, results from nonlinearity tests in panel data models should be interpreted with care. The linearity hypothesis may be rejected not because of the existence of a genuine nonlinear relationship between y_{it} and x_{it}, but

---

[5]Notice that $E(\omega_{ixx})E(\omega_{izz}) - (E(\omega_{ixz}))^2 > 0$, unless of course x_{it} and z_{it} are perfectly collinear for all i, which we rule out. See Assumption A.3.

[6]See, for example, Swamy (1970) and Hsiao (1986, Ch. 6).

[7]See, for example, Hendry (1993, Ch's. 6 and 19), and Gilbert (1986).

due to slope heterogeneity. Finally, it is worth noting that since the β_i's are fixed for each i, the nonlinear specification

$$y_{it} = \alpha_i + \delta_x x_{it} + \delta_z x_{it}^2 + v_{it}, \tag{2.13}$$

can not be reconciled with (2.12), unless it is assumed that β_i varies proportionately with x_{it}. Clearly, it is possible to allow the slopes, β_i, to vary systematically with some aspect of the cross-sectional distribution of x_{it} without requiring β_i to be proportional to x_{it}, and hence time-varying. For example, it could be that

$$\beta_i = \gamma_0 + \gamma_1 \bar{x}_i, \tag{2.14}$$

where $\bar{x}_i = T^{-1} \sum_{t=1}^{T} x_{it}$. This specification retains the linearity of (2.12) for each i, but can still yield a statistically significant effect for x_{it}^2 in (2.13) if slope heterogeneity is ignored and fixed-effects estimates of (2.13) are used in inference. This feature of fixed-effects regressions under heterogeneous slopes is illustrated in Figure 1. The figure shows scatter points and associated regression lines for three countries with slopes that differ systematically with $\bar{x}_i$. It is clear that the pooled regression based on the scatter points from all the three countries will exhibit strong nonlinearities, although the country-specific regressions are linear. For an empirical illustration of this see section 3.1 below.

## 2.2. Neglected Heterogeneity in Dynamic Models

The neglect of slope heterogeneity is yet more serious in dynamic panel data models. As shown in Pesaran and Smith (1995), it generates a correlation between the regressors and the error term as well as serial correlation in the disturbances, and hence introduces a bias in the fixed-effects estimator, even for sufficiently large T and N. Also, unlike the static case, the inconsistency of the fixed-effects estimators does not disappear even if we assume that the slope coefficients are distributed independently of the exogenous regressors and the disturbances. To get some idea of the size and the direction of the asymptotic bias of the fixed-effects estimators, suppose that the y_{it}'s are generated according to the following simple dynamic version of (2.6):

$$y_{it} = \mu_i + \lambda_i y_{i,t-1} + \beta_i x_{it} + u_{it}, \quad u_{it} \sim iid\ (0, \sigma_i^2), \quad i = 1, 2, ..., N; \ t = 1, 2, ..., T. \tag{2.15}$$

To simplify the exposition in what follows we assume that $\lambda_i = \lambda$, and focus on the fixed-effects estimators of λ and β (denoted by $\hat{\lambda}_{FE}$ and $\hat{\beta}_{FE}$) in

$$y_{it} = \mu_i + \lambda y_{i,t-1} + \beta x_{it} + v_{it}, \tag{2.16}$$

[10]

Figure 2.1: Fixed-Effects Regression under Heterogeneous Slopes
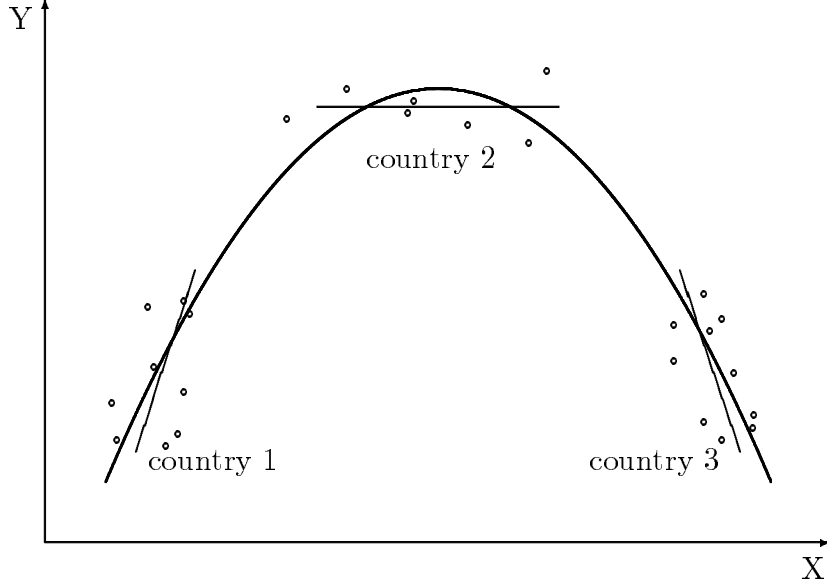
where
$$v_{it} = \eta_i x_{it} + u_{it}. \tag{2.17}$$

We also assume that the η_i's are distributed independently of x_{it} and u_{it} for all i and t, and that they have a zero mean and a constant variance, σ_η^2. It is now easily seen that the fixed-effects estimator of λ and β in (2.16) will be inconsistent even if both T and N are allowed to increase without bounds. The inconsistency of the fixed-effects estimators is due to the non-zero correlation between v_{it} and $y_{i,t-1}$, and the fact that v_{it}'s are serially correlated, unless x_{it} is serially uncorrelated. Assuming the dynamic processes (2.16) are stable (namely, $\mid \lambda \mid < 1$), and have been in operations for a sufficiently long time we can obtain

$$Cov(v_{it}, y_{i,t-1}) = E([(v_{it} - \bar{v}_i)(y_{i,t-1} - \bar{y}_{i,-1})] = \sigma_\eta^2 \sum_{j=0}^{\infty} \lambda^j \gamma_x(j+1),$$

where $\bar{v}_i = T^{-1} \sum_{t=1}^{T} v_{it}$, $\bar{y}_{i,-1} = T^{-1} \sum_{t=1}^{T} y_{i,t-1}$, and $\gamma_x(j)$ is the $j-th$ order autocovariance function of the x_{it} process. In this simple example $Cov(v_{it}, y_{i,t-1}) = 0$ only if x_{it} is a serially uncorrelated process, which is very unlikely in the case of most economic time series.[8] The aymptotic bias of the fixed-effects estimators of

---

[8] In the more general case where the coefficients of the lagged dependent variables, λ_i, also vary across groups $Cov(v_{it}, y_{i,t-1}) \neq 0$ even if the regressors, x_{it}, are serially uncorrelated. For further details see Pesaran and Smith (1995).

[11]

λ and β are derived in Pesaran and Smith (1995). In the case where x_{it} follows a first-order autoregressive process

$$x_{it} = \kappa_i (1 - \rho) + \rho\, x_{i,t-1} + \xi_{it}, \qquad \xi_{it} \sim iid\, (0, \tau_i^2),$$

they show that

$$\plim_{N,T\to\infty} \hat{\lambda}_{FE} = \lambda + \frac{\rho\,(1 - \lambda\rho)\,(1 - \lambda^2)\sigma_\eta^2}{\Psi},$$

and

$$\plim_{N,T\to\infty} \hat{\beta}_{FE} = \beta - \frac{\beta\,\rho^2\,(1 - \lambda^2)\sigma_\eta^2}{\Psi},$$

where $\hat{\lambda}_{FE}$ and $\hat{\beta}_{FE}$ denote the fixed-effects estimators of λ and β in (2.16) and[9]

$$\Psi = (\sigma^2/\tau^2)\,(1 - \rho^2)\,(1 - \lambda\rho)^2 + (1 - \lambda^2\rho^2)\,\sigma_\eta^2 + (1 - \rho^2)\,\beta^2 > 0.$$

The size of the asymptotic bias depends on β, λ, ρ, σ_η^2, and the means of σ_i^2 and τ_i^2, denoted by σ^2 and τ^2. It is easy to check that $\hat{\beta}_{FE}$ always underestimates β, whereas $\hat{\lambda}_{FE}$ over- or underestimates λ depending on whether ρ is positive or negative, respectively.

Pesaran and Smith also show that in the general case where both λ_i and β_i are allowed to vary across i, it is not possible to obtain consistent estimates of $\lambda = E(\lambda_i)$ and $\beta = E(\beta_i)$ by the application of the traditional instrumental variables or the Generalized Method of Moments (GMM) estimators proposed, for example, by Anderson and Hsiao (1981, 1982), Arellano (1989), Arellano and Bover (1995), Keane and Runkle (1992) and Ahn and Schmidt (1995), to a first-differenced form of (2.16). Instead, they propose the mean group (MG) estimator, obtained by estimating the coefficients of each cross-sectional unit separately by the ordinary least squares (OLS) and then by taking an arithmetic average of them. The MG estimators of λ and β (which we denote by $\hat{\lambda}_{MG}$ and $\hat{\beta}_{MG}$) are given by

$$\hat{\lambda}_{MG} = \frac{1}{N} \sum_{i=1}^{N} \hat{\lambda}_i, \quad \hat{\beta}_{MG} = \frac{1}{N} \sum_{i=1}^{N} \hat{\beta}_i, \tag{2.18}$$

where $\hat{\lambda}_i$ and $\hat{\beta}_i$ are the OLS estimates from individual group (country) regressions. The variance-covariance matrix of the MG estimators can then be consistently estimated by

$$\widehat{Cov}(\hat{\boldsymbol{\delta}}_{MG}) = \frac{\sum_{i=1}^{N}(\hat{\boldsymbol{\delta}}_i - \hat{\boldsymbol{\delta}}_{MG})(\hat{\boldsymbol{\delta}}_i - \hat{\boldsymbol{\delta}}_{MG})'}{N(N-1)},$$

---

[9] The expression for Ψ given here corrects a typo in equation (2.9) of Pesaran and Smith (1995), where $(1 - \rho^2)\,\beta^2$ in that equation was incorrectly entered as $(1 - \rho^2)\,\beta$.

where $\hat{\boldsymbol{\delta}}_{MG} = (\hat{\lambda}_{MG}, \hat{\beta}_{MG})'$.[10] MG estimation is possible when both T and N are sufficiently large, and is applicable irrespective of whether the slope coefficients are random (in Swamy's sense) or fixed in the sense that the diversity in the slope coefficients across countries can not be captured by means of a finite parameter probability distribution.[11] When the slopes differ randomly across countries the Swamy estimator (also known as the "empirical Bayes" estimator) computed as weighted averages of $\hat{\lambda}_i$ and $\hat{\beta}_i$, or the Bayes estimator proposed by Hsiao, Pesaran and Tahmiscioglu (1999) can be used. It is shown that all the three estimators; namely the mean group, the Swamy and the Bayes estimators are asymptotically equivalent and have a standard asymptotic normal distribution for large N and large T so long as $\sqrt{N}/T \to 0$, as both N and $T \to \infty$.

Although the MG and the Bayes estimators provide an asymptotically valid distribution theory for making inferences about the average of the slope coefficients, nevertheless they do not take account of the fact that certain parameters may be the same across groups. The traditional pooled estimators, such as the fixed- and random-effects estimators, stand at the other extreme and (with the possible exception of intercepts) assume that the coefficients and error variances are the same across the groups. An intermediate estimator which allows the intercepts, short-run coefficients and error variances to differ across groups, but imposes the equality of one or more of the long-run coefficients across the groups has been recently proposed by Pesaran, Shin and Smith (1999). This estimator is called the Pooled Mean Group (PMG) estimator because it involves both pooling and averaging. Such an intermediate estimation technique is particularly relevant to the cross-country analysis of consumption functions where it is reasonable to assume that aggregate consumption responds very similarly across countries to aggregate disposable income in the long-run, although due to historical and institutional reasons it may not be plausible to assume that aggregate consumption in all countries will adjust similarly and with the same speed to short term changes in income.

3. Empirical Evidence

In this section we examine the empirical evidence on the determinants of cross country private savings rates using the data from MBS (1995, 1998). We have chosen to focus on this particular study because of the relatively comprehensive

---

[10] For further details see, for example, Im, Pesaran and Smith (1996).

[11] In situations where N is small and T is large, the Seemingly Unrelated Regression Estimation (SURE) procedure due to Zellner (1962) can be used. In the cross-country savings literature N is often quite large which precludes the use of the SURE procedure. For example, the full data set used by MBS (1995) covers 61 countries.

list of variables that it considers, and also because it arrives at a number of policy conclusions whose robustness to neglected slope heterogeneity and dynamics are worth checking. As noted earlier, MBS divide their sample into an industrial country group and a developing country group. They present estimates obtained for all the 61 countries taken together as well as those from separate analyses of the two country groups. Homogeneity of the slope coefficients across the two country groupings is strongly rejected possibly also due to important differences between the two groups in terms of data definitions and data reliability. The data for the two groups also cover different time periods. Available time series for the developing countries are substantially shorter than for the industrial countries. Since our primary purpose here is to illustrate the importance of dynamics and slope heterogeneity in cross-country analysis of savings behaviour we confine our analysis to the industrial countries where conceivably slope heterogeneity could be less of an issue. For the industrial countries MBS run fixed-effects regressions of

$PSAV$: the private savings rate, defined as the ratio of aggregate private savings to GDP,

on

SUR : the ratio of general government budget surplus to GDP,
$GCUR$: the ratio of the general government current expenditure to GDP,
GI : the ratio of the general government investment to GDP,
GR : GDP growth rate,
$RINT$: real interest rate,
INF : inflation rate,
$PCTT$: percentage change in terms of trade,
$YRUS$: per capita GDP relative to the U.S.,
DEP : dependency ratio, defined as the ratio of those under 20 and 65 and over to those aged 20-64,
W : ratio of private wealth (measured as the cumulative sum of past nominal private savings) to GDP.

The fixed-effects estimates including all the 10 regressors are reported in the first column of table 2 in MBS (1995). This table also reports least squares and IV estimates for a restricted version of their general model where the regressors GR and W are dropped from the model and the coefficients of $GCUR$ and GI are imposed to be the same. In a more recent version of their paper, MBS (1998)

[14]

report estimates for a slightly more general regression model that also includes the square of the relative income variable (see table 3 in MBS, 1998). They argue that the effect of the income variable on the savings rate is likely to be nonlinear; being positive at low levels of income and then becoming negative at higher income levels. Presumably, similar arguments can also be made with respect to some of the other variables, notably the wealth variable; but such possibilities are not pursued by MBS. Fixed-effects estimates corrected for residual serial correlation, using Cochrane-Orcutt type procedures, are also reported "in order to avoid biasing estimates of standard errors", in the belief that the regression model being estimated is static and does not contain lagged dependent variables. However, their specification implicitly *does* contain lagged dependent variables through the wealth variable. Hence, evidence of significant residual serial correlation can be a problem, and as we suggest below, a likely symptom of dynamic misspecification and neglected slope heterogeneity.[12]

3.1. Slope Heterogeneity in "Static" Cross-Country Savings Regressions

For convenience we begin our investigation by reproducing the MBS's (1998) fixed-effects regression for the industrial countries in table 1. We refer to this specification as model M_0. The estimates under "model M_0" in table 1 are identical to those reported in column 1 of table 3 in MBS (1998); except for a couple of typos. Apart from the coefficient of the GDP growth rate (GR), all the estimated coefficients are statistically (some very highly) significant, and in particular suggest a strong quadratic relationship between saving and per capita income. However, as pointed out earlier, the validity of these estimates and the inferences based on them critically depend on the extent to which slope coefficients differ across countries, and in the case of static models whether these differences are systematic. Whether on balance it is reasonable to ignore slope heterogeneity clearly depends on the particular application under consideration, and it is therefore important that the robustness of the fixed-effects estimates to slope heterogeneity is checked, as far as possible, on a case-by-case basis.

As it was shown in section 2.1, one important implication of neglected slope heterogeneity is the possibility of obtaining spurious nonlinear effects. We explored this possibility by adding quadratic terms in W, INF, $PCTT$, and DEP to the regressors already included in model M_0. While there may be some *a priori* argument for a nonlinear wealth effect in the savings equation, the rationale for

---

[12] In general, the Cochrane-Orcutt type corrections do not adequately deal with dynamic misspecification. See Hendry (1993, Ch. 7), and the discussion in Section 2 on testing for common factor restrictions.

nonlinear effects in the case of the other three variables seems less clear. Nevertheless, as can be seen from the estimates reported under "model M_1" in table 1, all these quadratic terms are statistically highly significant. The quadratic relationships between the private savings rate and the variables W, $PCTT$, and DEP are in fact much stronger than the quadratic relationship between savings and per capita income that MBS focus on. The $\bar{R}^2$ of the augmented model, 0.801, is also appreciably larger than that obtained for model M_0, 0.766. A similar conclusion is reached using other model selection criteria such as the Akaike Information Criterion (AIC) and the Schwarz Bayesian Criterion (SBC) also reported in table 1.

As an alternative to the quadratic specifications used in model M_1, we investigated the possibility that the slope coefficients in each country are *fixed over time*, but are allowed to vary across countries linearly with the sample means of their wealth to GDP ratio or their income per capita. More specifically, denote the vector of slope coefficients for country i by $\boldsymbol{\beta}_i$, and define

$$\bar{W}_i = T^{-1} \sum_{t=1}^{T} W_{it}, \text{ and } \overline{YRUS}_i = T^{-1} \sum_{t=1}^{T} YRUS_{it}.$$

Then, slope heterogeneity in the present application could be modelled by:

$$\boldsymbol{\beta}_i = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_{01}\bar{W}_i + \boldsymbol{\beta}_{02}\overline{YRUS}_i. \tag{3.1}$$

Substituting the above expression for $\boldsymbol{\beta}_i$ in (2.5) we have[13]

$$y_{it} = \mu_i + \boldsymbol{\beta}_0'\mathbf{x}_{it} + \boldsymbol{\beta}_{01}'(\mathbf{x}_{it}\bar{W}_i) + \boldsymbol{\beta}_{02}'(\mathbf{x}_{it}\overline{YRUS}_i) + u_{it}, \tag{3.2}$$

where $y_{it} = PSAV_{it}$,

$\mathbf{x}_{it} = (SUR_{it}, GCUR_{it}, GI_{it}, GR_{it}, RINT_{it}, W_{it}, INF_{it}, PCTT_{it}, YRUS_{it}, DEP_{it})'$.

Initially we estimated all the elements of $(\boldsymbol{\beta}_0, \boldsymbol{\beta}_{01}, \boldsymbol{\beta}_{02})$ freely, but found that in the case of almost all the regressors $\bar{W}_i$ was by far a better proxy than $\overline{YRUS}_i$ for capturing slope heterogeneity in this application. The exception was the coefficient of the real interest variable, $RINT$. The freely estimated elements of $\boldsymbol{\beta}_0$, $\boldsymbol{\beta}_{01}$, and $\boldsymbol{\beta}_{02}$, together with their t-ratios are given in table 2. Apart from the coefficient of the SUR variable, all the other coefficients show systematic variation across countries. The coefficient of the SUR variable seems to be least affected by

---

[13] It is also reasonable to expect that in addition to the systematic effects characterized through the dependence of $\boldsymbol{\beta}_i$ on $\bar{W}_i$ and $\overline{YRUS}_i$, the slope coefficients also differ randomly across i. Such a formulation, however, moves us towards the mean group or weighted mean group estimation procedures discussed in Section 2.2, which we shall consider later in this section.

[16]

slope heterogeneity, and the hypothesis of slope homogeneity can not be rejected in the case of this variable. However, none of the other estimates are directly comparable to the fixed-effects estimates given in table 1. In particular, the co-efficients of output growth variables (GR_{it} and $GR_{it} * \bar{W}_i$) are both statistically significant, while this was not so in the case of the fixed-effects estimates in table 1. Care must also be exercised when interpreting these estimates. For example, the results suggest that the effect of real output growth on the savings rate is likely to be higher in a country with a high wealth-GDP ratio. Similarly, inflation effects on the savings rate are estimated to be higher in countries with higher wealth to GDP ratios. However, these results do not predict, for instance, that an individual country's savings rate will necessarily rise with output growth. table 3 gives the country-specific estimates together with their cross-sectional averages. Except for the coefficient of the SUR variable all the other estimates show con-siderable variation across countries. The coefficient of the output growth rate, for example, ranges from -0.211 (for Greece) to 0.172 (for Portugal). The effect of inflation on savings varies across countries even more widely; ranging from -0.142 to 0.374. The same also applies to the relative income variable, $YRUS$. Only the coefficients of the government expenditure variables, $GCUR$ and GI , do not change sign and remain negative across all the 21 countries.

The country-specific estimates in table 3 are obtained on the assumption that the slope coefficients across countries are *exact* linear functions of $\bar{W}_i$ and/or $\overline{YRUS}_i$. The underlying fixed-effects regression in table 2 also assumes that the error variances, $Var(u_{it}) = \sigma_i^2$, are the same across countries. Clearly, these are rather restrictive assumptions, and the consequences of incorrectly imposing them on the parameters of interest need to be examined. Under the alternative assumption of unrestricted slope and error variance heterogeneity country-specific estimates can be obtained from individual country regressions. Mean group (MG) estimates computed as simple averages of country-specific estimates from country specific regressions can then be used to make inferences about $E(\beta_i) = \beta$ along the lines reviewed in section 2.2.

We first estimated the basic MBS specification containing the relative income variable ($YRUS$) and its square ($YRUS^2$) for 20 out of the 21 OECD economies in the panel. We had to exclude the U.S. equation from this analysis since its parameters are not identified.[14] The quadratic terms $YRUS_{it}^2$ were statistically significant only in the case of four countries: Austria, Belgium, Ireland, and Norway. In view of this and bearing in mind that the statistical significance

---

[14]Notice that $YRUS_{it}$ is defined as the ratio of country i's GDP per capita to that of the U.S., and hence it is indistinguishable from an intercept in the U.S. saving equation. A more appropriate measure would be to measure income per capita relative to the industrial countries average and thus avoid the identification problem in the case of the U.S. saving equation.

of the nonlinear effects in fixed-effects regressions could be spurious, possibly reflecting systematic slope heterogeneity, we decided to confine our analysis to linear individual country regressions. The results are summarized in table 4. The estimated slope coefficients differ considerably across countries, both in terms of their magnitude and their statistical significance. Some of the coefficients are statistically significant only in the case of 3 or 4 countries and in general are very poorly estimated.[15] This is true of the coefficients of GI, GR, W, $PCTT$ and $YRUS$. Also the sign of these estimated coefficients vary quite widely across countries. The coefficients of $RINT$ and INF are better estimated, but still differ significantly both in magnitude and in sign across the countries. Only the estimated coefficients of the SUR and $GCUR$ variables are estimated precisely and tend to be similar across countries. The coefficient of SUR is estimated to be negative in 19 out of the 20 countries. The positive estimate (obtained for New Zealand) is very small and not statistically significant. Thirteen of these estimated coefficients are negative and statistically significant. Similarly, 17 out of 20 coefficients estimated for the $GCUR$ variable have a negative sign. None of the three positive coefficients estimated for $GCUR$ are statistically significant, and 7 of the 17 negative coefficients are statistically significant.

The mean group estimates based on the individual country regressions in table 4 support these general conclusions. Only the MGE of the SUR and the $GCUR$ variables are statistically significant (see the last two rows of table 4). At -0.671, the MGE of the SUR variable is only marginally higher than the corresponding fixed-effects estimate in table 2 that allows for some slope heterogeneity. The MGE of the $GCUR$ variable (-0.401) is also similar to the average estimate (-0.450) reported in table 3, computed using the fixed-effects, varying-slope regression of table 2. These results are quite encouraging; but before reaching a conclusion we also need to consider the possible effects of mis-specified dynamics on the estimates.

3.2. Slope Heterogeneity in Dynamic Cross-Country Savings Regressions

The fixed-effects "static" private savings regressions reported in tables 1 and 2 are subject to a substantial degree of residual serial correlation, and as pointed out earlier can lead to inconsistent estimates even under slope homogeneity since the wealth variable, W, is in fact constructed from accumulation of past savings. Presence of residual serial correlation could be due to a host of factors: omitted variables, neglected slope heterogeneity in the case of serially correlated regressors, and of course neglected dynamics. From the diagnostic statistics provided

---

[15] In Table 4 the statistically significant estimates (at the 95 per cent level) are shown in bold.

in the second part of table 4 it is clear that even when the slope coefficients are allowed to be estimated freely across countries the residual serial correlation still continues to be a problem, at least in the case of some, if not all the countries.[16] The diagnostic statistics computed for the individual country regressions show statistically significant evidence of residual serial correlation in the case of 8 of the 20 countries: Belgium, Finland, Germany, Greece, New Zealand, Norway, Spain and Sweden.[17] The usual time series technique for dealing with dynamic misspecification is to estimate error correction models based on autoregressive distributed lag (ARDL) models of the type discussed in section 2. ARDL models have the advantage that they are robust to integration and cointegration properties of the regressors, and for sufficiently high lag orders could be immune to the endogeneity problem, at least as far as the long-run properties of the model are concerned.[18] In the present application observations for each individual country are available for too short a period to estimate even a first-order ARDL model including all the 10 regressors for each country seperately.[19] Pooling in the form of fixed-effects estimation can compensate for lack of time series observations but as we have shown this can have its own set of problems.

To check the robustness of the "static" fixed-effects estimates presented in table 2 to dynamic misspecification we estimated the following first-order dynamic panel data model:

$$y_{it} = \mu_i + \lambda y_{i,t-1} + \boldsymbol{\beta}_0' \mathbf{x}_{it} + \boldsymbol{\beta}_{01}' (\mathbf{x}_{it} \bar{W}_i) + \boldsymbol{\beta}_1' \mathbf{x}_{i,t-1} + u_{it}. \qquad (3.3)$$

The country-specific long-run coefficients are given by

$$\theta_i = (\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 + \boldsymbol{\beta}_{01} \bar{W}_i)/(1 - \lambda). \qquad (3.4)$$

We also considered the inclusion of the weighted lagged variables $\bar{W}_i \mathbf{x}_{i,t-1}$ and $\bar{W}_i y_{i,t-1}$ to the model but did not find them to be statistically significant. The fixed-effects estimates computed using all the 21 countries over the period 1972-1993 are given in table 5.[20] Clearly, there are significant dynamics, particularly in the relationship between changes in the government surplus and expenditure

---

[16] The diagnostic statistics are computed using the Lagrange multiplier procedure described, for example, in Pesaran and Pesaran (1997, Section 18.5.2), and are valid irrespective of whether the regressions contain lagged dependent variables, implicitly or explicitly.

[17] Under slope homogeneity restrictions, residual serial correlation is a problem for *all* the countries in the panel.

[18] For a detailed discussion see Pesaran and Shin (1999).

[19] A first-order ARDL model in the private savings rate for each country that contains *all* 10 regressors would involve estimating 22 unknown parameters with only 22 time series observations available per country!

[20] For relatively simple dynamic models where T (= 22) is reasonably large and of the same order of magnitude as N (= 21), the application of the IV type estimators to a first differenced

[19]

variables (SUR, $GCUR$, and GI) and the private savings rate. There is also important evidence of cross-sectional variations in the coefficients of the wealth, income and demographic variables (W, $YRUS$ and DEP). However, unlike the static estimates in table 2, the coefficients of GDP growth and the real interest rate are no longer statistically significant. Overall, this equation presents a substantial improvement over the static fixed-effects estimates. In fact the estimated standard error of this dynamic regression is 62 percent lower than the standard error of the fixed-effects estimates favoured by MBS (1998), and reproduced in the first column of table 1.

Using the formula (3.4) we obtained the following estimates of the long-run coefficients:

$$
\begin{array}{lll}
SUR & \underset{(-3.11)}{-0.432} & \\
GCUR & \underset{(-4.65)}{-0.398} & \\
GI & \underset{(-0.91)}{-0.202} & \\
GR & \underset{(-0.03)}{-0.004} & \\
RINT & \underset{(1.64)}{0.154} & \\
W & \underset{(4.58)}{0.224} & \underset{(-3.77)}{-0.00057}\, \bar{W}_i \\
INF & \underset{(3.10)}{0.248} & \\
PCTT & \underset{(4.11)}{0.136} & \\
YRUS & \underset{(2.58)}{1.384} & \underset{(-2.92)}{-0.0047}\, \bar{W}_i \\
DEP & \underset{(2.19)}{0.708} & \underset{(-2.64)}{-0.0027}\, \bar{W}_i
\end{array}
$$

According to these estimates the long-run coefficients of the SUR and $GCUR$ variables are still statistically significant, although the coefficient of the SUR variable is now estimated to be much lower than the estimate based on the static regressions. The long-run coefficients of the GI, GR and $RINT$ variables are no longer statistically significant. It appears that in contrast to government consumption expenditures, the effect of changes in government investment expenditures on private savings is temporary and tends to zero in the long run. The inflation and the terms of trade variables (INF and $PCTT$) have the expected signs and are also statistically significant. The long-run coefficients of the remaining variables

---

version of (3.3) does not seem necessary, and can lead to considerable loss of efficiency. The estimation of fixed-effects dynamic panel data models when the time dimension is short, say in the range of 3-10, and the cross-sectional dimension is relatively large, say 50 or more, is discussed in Anderson and Hsiao (1981,1982), Arellano and Bond (1991), and more recently in Hsiao, Pesaran, and Tahmiscioglu (1998).

vary with country-specific average wealth-GDP ratio and when averaged across countries yield the values of 0.043 [0.026], -0.118 [0.219] and -0.148 [0.125] for W, $YRUS$, and DEP variables respectively. The cross-sectional standard errors of these estimates are given in square brackets. The average estimate of the coefficient of the relative income variable has the wrong sign, but it is not statistically significant. The average estimates of the other two coefficients have the expected signs, but are not statistically significant either. It is clear that the effects of many of the regressors considered in the MBS study are not robust to dynamic misspecifications.

But before reaching a final conclusion we also need to examine the consequences of *jointly* allowing for unrestricted short-run slope heterogeneity and dynamics.[21] To this end we estimated individual country regressions containing first-order lagged values of the savings rates, $PSAV_{i,t-1}$. The MG estimates of the long run coefficients based on these dynamic individual country regressions are given in table 6.[22] For ease of comparison we have also included MGE based on a static version of these regressions, as well as the corresponding fixed-effects estimates. Unlike the fixed-effects estimates the consequences of allowing for dynamics on the MG estimates are rather limited. Once again only the coefficients of the SUR and the $GCUR$ variables are statistically significant; although the dynamic MG estimates suggest the coefficient of the $PCTT$ variable also to be marginally significant.

Finally, in the last column of table 6 we give the pooled mean group estimates of the long-run coefficients recently proposed by Pesaran, Shin and Smith (1999), where the short-run dynamics are allowed to differ freely across countries but equality restrictions are imposed on one or more of the long-run coefficients. The rationale being that due to differences in factors such as adjustment costs or the institutional set-up across countries slope homogeneity is more likely to be valid in the long run. The pooled mean group estimates (PMGE) in table 6 impose the slope homogeneity restrictions only on the long-run coefficients of the SUR variable. We did try similar restrictions on the other long-run coefficients, but they were rejected using a Hausman-type test that compares the MG and the PMG estimates.[23] Only the homogeneity restrictions on the long-run coefficients of the SUR variable could not be rejected. As expected, the PMGE are generally more precisely estimated and confirm that amongst the various determinants of private savings considered by MBS only the effects of the SUR and the $GCUR$ variables

---

[21] The estimates in Table 5 only allow for a limited degree of slope heterogeneity.

[22] Individual country regressions that underlie the estimation of the MGE are available from the authors on request.

[23] For details of the Hausman test applied to MG and PMG estimates see Pesaran, Shin and Smith (1999).

seem to be reasonably robust to the presence of slope heterogeneity and yield plausible estimates for the off-setting effects of government budget surpluses and government consumption expenditures on private savings across OECD countries.

4. Concluding Remarks

Given the paucity of long time series on individual countries, examination of many macroeconomic hypotheses in the literature relies on cross-country panel data. Much of this literature is based on simple fixed-effects estimators and does not adequately address the possibility that there may be considerable heterogeneity in short-run and long-run parameters across countries. This paper has shown that neglecting heterogeneity and dynamics in cross-country savings regressions can lead to misleading inferences about the key determinants of savings behavior. If differences across countries are ignored, one can overestimate the influence of certain factors on the private savings rates and at the same time obtain highly significant, but spurious, nonlinear effects for some of the potential determinants.

Our results indicate that fiscal variables – the general government surplus as a proportion of GDP and the ratio of government consumption to GDP– are the key determinants of private savings rates in the industrial countries in the post world war II period. Keeping government consumption constant, the estimated parameters imply that the Ricardian offset in private savings of a decrease in the government surplus is quite large (about 0.9 in table 6). The estimated parameters also imply that changes in government consumption have a different impact on the private savings rate than changes in government investment. In fact, once slope heterogeneity is adequately taken into account the impact of government investment expenditures on private saving is not statistically significant. The results in the paper also provide some support for the Harberger-Laursen-Metzler effect with improving terms of trade having a small positive effect on the private savings rate. For the OECD economies over the period under consideration, we do not find any evidence that the rate of output growth, the inflation rate, the real interest rate, the ratio of wealth to GDP, or the dependency ratio have had a statistically significant (long-term) effect on the private savings rate.

A. Appendix

Suppose the dependent variables y_{it}, $i = 1, 2, ..., N$, are generated according to the following simple, static panel data model:

$$y_{it} = \mu_i + \beta_i x_{it} + u_{it}, \ i = 1, 2, ..., N; \ t = 1, 2, ..., T. \tag{A.1}$$

The intercepts, μ_i, as well as the slopes, β_i, are allowed to vary across groups. Intercepts will be treated as fixed-effects. For the slopes we use the random-coefficients model

$$\beta_i = \beta + \eta_i, \ i = 1, 2, ..., N, \tag{A.2}$$

but allow for the possibility of systematic dependence between η_i and the regressors x_{it} and z_{it}. In this appendix we derive the inconsistency of the fixed-effects estimators of the slope coefficients δ_x and δ_z in the more general panel data model

$$y_{it} = \alpha_i + \delta_x x_{it} + \delta_z z_{it} + v_{it} = \alpha_i + \boldsymbol{\delta}' \mathbf{w}_{it} + v_{it}, \ i = 1, 2, ..., N; \ t = 1, 2, ..., T, \tag{A.3}$$

when in fact the y_{it}'s are generated according to (A.1). To simplify the derivations we make the following assumptions:

Assumption A.1: u_{it} is serially uncorrelated and distributed independently of u_{jt} for all $i \neq j$.

Assumption A.2: $\mathbf{w}_{it} = (x_{it}, z_{it})'$ is distributed independently of $u_{it'}$, for all i, t and t'.[24]

Assumption A.3: $\mathbf{w}_{it}$ follows a covariance stationary process with the variance-covariance matrix, $\boldsymbol{\Omega}_i$, defined by (2.9), such that

$$E(\boldsymbol{\Omega}_i) = \lim_{N \to \infty} \left(\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\Omega}_i \right), \tag{A.4}$$

is a positive definite matrix.

Assumption A.4: For each t, $\mathbf{w}_{it}$ is distributed independently across i.

Although, not all the above assumptions are necessary when both N and T are sufficiently large. For example, assumption A.4 is not needed when T is sufficiently large. Assumption A.3 can be relaxed when T is small. It is also worth noting that assumption A.3 does not require the correlation matrix of the regressors for all i to be non-singular, only that the "pooled" covariance matrix, $E(\boldsymbol{\Omega}_i)$, defined by (A.4), should be non-singular.

---

[24]These assumptions require $\mathbf{w}_{it}$ to be strictly exogenous, and therefore rule out the case where the model contains lagged dependent variables. The problem of inconsistency of fixed- or random-effects estimators of dynamic panel data models under slope heterogeneity is addressed in Pesaran and Smith (1995).

In matrix notations (A.1) and (A.3) can be written as:

$$\mathbf{y}_i = \mu_i \boldsymbol{\iota} + \beta_i \mathbf{x}_i + \mathbf{u}_i, \tag{A.5}$$

and

$$\mathbf{y}_i = \alpha_i \boldsymbol{\iota} + \mathbf{W}_i \boldsymbol{\delta} + \mathbf{v}_i, \tag{A.6}$$

respectively, where

$$\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})', \boldsymbol{\iota} = (1, 1, ..., 1)', \mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{iT})',$$

$$\mathbf{u}_i = (u_{i1}, u_{i2}, ..., u_{iT})', \boldsymbol{\delta} = (\delta_x, \delta_z)',$$

and

$$\mathbf{W}_i = \begin{bmatrix} x_{i1} & z_{i1} \\ x_{i2} & z_{i2} \\ \vdots & \vdots \\ x_{iT} & z_{iT} \end{bmatrix}, \ \mathbf{v}_i = \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{iT} \end{bmatrix}.$$

The fixed-effects estimators of the slope coefficients in (A.6) can be written as:

$$\hat{\boldsymbol{\delta}}_{FE} = \begin{pmatrix} \hat{\boldsymbol{\delta}}_{x,FE} \\ \hat{\boldsymbol{\delta}}_{z,FE} \end{pmatrix} = \left(\sum_{i=1}^{N} \mathbf{W}_i' \mathbf{H} \mathbf{W}_i \right)^{-1} \left(\sum_{i=1}^{N} \mathbf{W}_i' \mathbf{H} \mathbf{y}_i \right), \tag{A.7}$$

where $\mathbf{H} = \mathbf{I}_T - \boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{\iota})^{-1}\boldsymbol{\iota}'$.[25] Under (A.5) we have

$$\hat{\boldsymbol{\delta}}_{FE} = \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{W}_i' \mathbf{H} \mathbf{W}_i \right)^{-1} \left\{ \frac{1}{NT} \sum_{i=1}^{N} (\mathbf{W}_i' \mathbf{H} \mathbf{x}_i) \beta_i + \frac{1}{NT} \sum_{i=1}^{N} \mathbf{W}_i' \mathbf{H} \mathbf{u}_i \right\}. \tag{A.8}$$

It is now easily seen that under assumptions A.1-A.4 and for N and/or T sufficiently large

$$\frac{1}{NT} \sum_{i=1}^{N} \mathbf{W}_i' \mathbf{H} \mathbf{u}_i \overset{p}{\to} 0, \tag{A.9}$$

where $\overset{p}{\to}$ denotes convergence in probability. Therefore

$$\mathrm{plim}(\hat{\boldsymbol{\delta}}_{FE}) = \left\{ p\lim \left(\sum_{i=1}^{N} \frac{\mathbf{W}_i' \mathbf{H} \mathbf{W}_i}{NT} \right) \right\}^{-1} \mathrm{plim} \left(\sum_{i=1}^{N} \left(\frac{\mathbf{W}_i' \mathbf{H} \mathbf{x}_i}{NT} \right) \beta_i \right). \tag{A.10}$$

---

[25] The fixed effects estimator in (A.7) assumes a balanced panel. But the results readily extend to unbalanced panels.

[24]

In the case where the slopes are homogenous, namely $\beta_i = \beta$, we have

$$\text{plim}(\hat{\boldsymbol{\delta}}_{FE}) = \begin{pmatrix} \beta \\ 0 \end{pmatrix}. \tag{A.11}$$

Consider now the case where the slopes are heterogenous. Using (A.2) in (A.10), it is easily seen that the consistency result in (A.11) will follow if and only if

$$\sum_{i=1}^{N} \left(\frac{\mathbf{W}_i' \mathbf{H} \mathbf{x}_i}{NT} \right) \eta_i = \begin{pmatrix} \frac{1}{NT} \sum_{i=1}^{N} \mathbf{x}_i' \mathbf{H} \mathbf{x}_i \eta_i \\ \frac{1}{NT} \sum_{i=1}^{N} \mathbf{z}_i' \mathbf{H} \mathbf{x}_i \eta_i \end{pmatrix} \overset{p}{\to} 0. \tag{A.12}$$

This condition holds under the random coefficient specification where it is assumed that η_i's are distributed independently of $\mathbf{w}_{it}$ for all i and t. (See also Swamy (1970)). Under assumption A.3 and as $T \to \infty$ we have

$$\frac{1}{NT} \sum_{i=1}^{N} \mathbf{x}_i' \mathbf{H} \mathbf{x}_i \eta_i - \frac{1}{N} \sum_{i=1}^{N} \omega_{ixx} \eta_i \overset{p}{\to} 0,$$

and

$$\frac{1}{NT} \sum_{i=1}^{N} \mathbf{z}_i' \mathbf{H} \mathbf{x}_i \eta_i - \frac{1}{N} \sum_{i=1}^{N} \omega_{izx} \eta_i \overset{p}{\to} 0.$$

Therefore, for the fixed-effects estimators, (A.7), to be consistent we must have

$$\frac{1}{N} \sum_{i=1}^{N} \omega_{ixx} \eta_i \overset{p}{\to} 0, \quad \text{and} \quad \frac{1}{N} \sum_{i=1}^{N} \omega_{izx} \eta_i \overset{p}{\to} 0, \tag{A.13}$$

as $N \to \infty$. Namely, any systematic dependence between β_i and the second order moments of the steady state distribution of the regressors $\mathbf{w}_{it}$ must also be ruled out. When the conditions in (A.13) are not satisfied the inconsistencies of the fixed-effects estimators (for T and N sufficiently large) are given by[26]

$$\text{plim}(\hat{\boldsymbol{\delta}}_{x,FE} - \beta) = \frac{Cov(\omega_{ixx}, \eta_i) E(\omega_{izz}) - E(\omega_{ixz}) Cov(\omega_{ixz}, \eta_i)}{E(\omega_{ixx}) E(\omega_{izz}) - (E(\omega_{ixz}))^2} \tag{A.14}$$

$$\text{plim}(\hat{\boldsymbol{\delta}}_{z,FE}) = \frac{Cov(\omega_{ixz}, \eta_i) E(\omega_{ixx}) - E(\omega_{ixz}) Cov(\omega_{ixx}, \eta_i)}{E(\omega_{ixx}) E(\omega_{izz}) - (E(\omega_{ixz}))^2} \tag{A.15}$$

---

[26]Notice that under slope heterogeneity the fixed effects estimators are inconsistent when N is finite and only $T \to \infty$.

[25]

where

$$Cov(\omega_{ixx}, \eta_i) = \operatorname*{p\,lim}_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{ixx}\eta_i \right), \quad Cov(\omega_{ixz}, \eta_i) = \operatorname*{p\,lim}_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{ixz}\eta_i \right),$$

$$E(\omega_{ixx}) = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{ixx} \right), E(\omega_{izz}) = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{izz} \right), \qquad (A.16)$$

$$E(\omega_{ixz}) = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \omega_{ixz} \right).$$

References

[1] Anderson, T.W. and C. Hsiao, (1981), "Estimation of Dynamic Models with Error Components", *Journal of the American Statistical Association*, Vol. 76, pp. 598–606.

[2] Anderson, T.W. and C. Hsiao, (1982), "Formulation and Estimation of Dynamic Models Using Panel Data", *Journal of Econometrics*, Vol. 18, pp. 47–82.

[3] Arellano, M. and S. Bond, (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, Vol. 58, pp. 277–297.

[4] Attanasio, O.P., L. Picci, and A. Scorcu, (1998), "Saving, Growth and Investment: A Macroeconomic Analysis Using a Panel of Countries", Unpublished Manuscript, University College London and University of Bologna.

[5] Callen, T. and C. Thimann, (1997), "Empirical Determinants of Household Saving: Evidence from OECD Countries", *IMF Working Paper*, WP/97/181, (Washington: International Monetary Fund).

[6] Carroll, C.D. and D.N. Weil, (1994), "Saving and Growth: A Reinterpretation", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 40, pp. 133–192.

[7] Dayal-Gulati, A. and C. Thimann, (1997), "Saving in South East Asia and Latin America Compared: Searching for Policy Lessons", *IMF Working Paper*, WP/97/110, (Washington: International Monetary Fund).

[8] Edwards, S., (1996), "Why Are Latin America's Savings Rates so Low? An International Comparative Analysis", *Journal of Development Economics*, Vol. 51, pp. 5–44.

[9] Gilbert, C.L., (1986), "Professor Hendry's Econometric Methodology", *Oxford Bulletin of Economics and Statistics*, Vol. 48, pp. 283–307.

[10] Hendry, D.F., (1993), *Econometrics: Alchemy or Science?*, Blackwell, Oxford.

[11] Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge University Press, Cambridge.

[12] Hsiao, C., M.H. Pesaran, and A.K. Tahmiscioglu, (1998), "Maximum Likelihood Estimation of Fixed-Effects Dynamic Panel Data Models Covering Short Time Periods", Unpublished Manuscript, University of Southern California and Cambridge University. Available via the Internet: http://www.econ.cam.ac.uk/faculty/pesaran/.

[13] Hsiao, C., M.H. Pesaran, and A.K. Tahmiscioglu, (1999), "Bayes Estimation of Short-Run Coefficients in Dynamic Panel Data Models", in C. Hsiao, K. Lahiri, L.-F. Lee and M.H. Pesaran (eds.), *Analysis of Panels and Limited Dependent Variables: A Volume in Honour of G. S. Maddala*, Cambridge University Press, Cambridge.

[14] Keane, M.P. and D.E. Runkle, (1992), "On the Estimation of Panel Data Models with Serial Correlation when Instruments are not Strictly Exogenous", *Journal of Business and Economic Statistics*, Vol. 10, pp. 1–29.

[15] Masson, P.R., T. Bayoumi and H. Samiei, (1995), "International Evidence on the Determinants of Private Saving", *IMF Working Paper*, WP/95/51, (Washington: International Monetary Fund).

[16] Masson, P.R., T. Bayoumi and H. Samiei, (1998), "International Evidence on the Determinants of Private Saving", the *World Bank Economic Review,* (forthcoming).

[17] Ogaki, M., J. Ostry and C. Reinhart, (1996), "Saving Behavior in Low- and Middle-Income Developing Countries", *IMF Staff Papers*, Vol. 43, pp.38–97.

[18] Pesaran, M.H. and B. Pesaran, (1997), *Working with Microfit 4.0: Interactive Econometric Analysis*, Oxford University Press, Oxford.

[19] Pesaran, M.H. and R. Smith (1995), "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels", *Journal of Econometrics*, Vol. 68, pp. 79–113.

[20] Pesaran, M.H. and Y. Shin, (1999), "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis", in S. Strom, A. Holly, and P. Diamond (eds.), *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, Cambridge University Press, Cambridge.

[21] Pesaran, M.H., Yongcheol Shin, and R. Smith (1999), "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels", *Journal of the American Statistical Association* (forthcoming).

[22] Pesaran, M.H., R. Smith, and K.-S. Im, (1996), "Dynamic Linear Models for Heterogeneous Panels", in L. Matyas and P. Sevestre (eds.), *The Econometrics of Panel Data*, Kluwer Academic Publishers.

[23] Schmidt-Hebbel, K., S.B. Webb, and G. Corsetti, (1992), "Household Saving in Developing Countries: First Cross-Country Evidence", *World Bank Economic Review*, Vol. 6, pp. 529–547.

[24] Sims, C.A., (1980) "Macroeconomics and Reality", *Econometrica*, Vol. 48, pp. 1–48.

[25] Swamy, P.A.V.B. (1970), "Efficient Inference in a Random Coefficient Regression Model," *Econometrica*, Vol. 38, pp. 311–323.

[26] Zellner, A., (1962), "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias", *Journal of the American Statistical Association*, Vol. 57, pp. 348–368.