**DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS**

A PROJECTION PURSUIT APPROACH TO CROSS COUNTRY GROWTH DATA

Andros Kourtellos

Discussion Paper 2002-13

A Projection Pursuit Approach to Cross Country Growth Data

Andros Kourtellos

Department of Economics

University of Cyprus

andros@ucy.ac.cy

June 18, 2002

Abstract

The empirical modeling of the cross-country differences in growth behavior is undoubtedly one of the most predominant research topics in applied macro-econometrics. However, despite the vast research effort it seems that there are only a few firm conclusions on the sources of cross-country differences. Unlike the bulk of the literature which focuses on linear parametric models this paper studies a semi-parametric way of modelling. In particular, it employs projection pursuit regression (PPR) to model the mean regression function of the growth process by a sum of unknown ridge functions (functions of linear combinations of covariates). PPR model was proposed by Friedman and Stuetzle (1981) to approximate high dimensional functions by simpler functions that operate in low dimensional spaces-typically one-dimensional. My findings identify non-linear relationships among the basic Solow-type variables. In particular, initial income and human capital affect growth in a very nonlinear way. Furthermore, there is evidence of interaction effects between human capital and initial income as well as between initial income and population growth rates. These findings suggest the presence of two steady-state equilibria that classify countries into two groups with different convergence characteristics.

1   Introduction

One of the most important set of questions economists face today is why differences in standards of living across countries are huge. Why is the typical person in the United States is twenty to thirty times richer that the typical person in Nigeria? Why do we observe negative growth rates? Do clusters or convergence clubs of economies

1

emerge from the interaction of countries (through trade or technological transfer), or does convergence to equality result? Researchers have tried to empirically answer this kind of questions by looking at cross-country growth rates.

In doing so, researchers have been using the linear Solow growth model as illustrated by Barro (1991, 1997), Barro and Sala-I-Martin (1995), and Mankiw, Romer and Weil (1992), Evans (1993), Islam (1995). This takes the form of a linear regression model of growth rates on population growth rates, saving rates for physical and human capital accumulation and initial income. The coefficient on initial income allows the assessment of the conditional convergence hypothesis; that is the proposition that countries with identical structural characteristics converge to one another in the long run independently of their initial conditions. This baseline model is usually augmented by additional country specific covariates suggested by a range of new growth theories. For instance, Durlauf and Quah (1999) registered more than 80 variables - ranging from market distortions, geographical regions, source endowments, climate, institutions, politics, war etc. - used (and found to be significant) by various researchers. Based on this model researchers make strong predictions about the differences in growth rates across countries.

> "...Growth differences between countries depend first on each country's existing level of output. If a country's current output is below its steady-state level of output, there is a catching-up process, which occurs mainly through technological transfer. Each year's growth eliminates some 2.5 percent of the gap between actual and steady-state output", Barro (1997), pp.vii.

Ironically, Robert Solow in his Economica, 53, S23-34, 1986, is skeptical[1] about his own model which has been taken to explain the cross-country growth differences for every country in the world:

> "...One model is supposed to apply everywhere and always. Each country is just a point on a cross-section regression, or one among several essentially identical regressions, leaving only grumblers to worry about what is exogenous and what is endogenous, and whether simple parameterizations do justice to real differences in the way the economic mechanism functions in one place or another."

Solow's scepticism is embraced by many other researchers. Typical examples include Pack (1994), Durlauf and Quah (1999), and Brock and Durlauf (2000). These studies identify and describe a number of problems which can explain the widespread mistrust with cross-country growth regression.

---

[1]In Solow's original work, Solow (1958, 1958), the model was developed based on stylized facts from the most advanced countries and these facts were never interpreted as universal properties for every country in the world.

This paper focuses on the problem of the linear specification of the cross-country growth regression and proposes a flexible semiparametric specification to uncover nonlinearities in the cross-country growth process. In contrast, the current literature imposes strong homogeneity assumptions on the cross-country growth process by assuming that each country is described by an identical Cobb-Douglas aggregate production function. This is odd given that modern growth theory suggests that different countries should be described by different aggregate production functions. What is more, under many endogenous growth models, the cross-country growth process is profoundly nonlinear. Hence, using the linear Solow model to test among competing growth models is of limited use; see Bernard and Durlauf (1996). Furthermore, the assumption of Cobb-Douglas production functions may not be valid. This is important given that Cobb-Douglas production function is a necessary condition for the linearity of the Solow growth model. To this end, Duffy and Papageorgiou (1999) find evidence in favor of a CES production function rather than the standard Cobb-Douglas specification.

Additionally, several empirical studies have developed evidence of nonlinearities and cross-country heterogeneity. Examples include Durlauf and Johnson (1995), Desdoigts (1999), Liu and Stengos (1999), Canova (1999), Hansen (2000), Durlauf, Kourtellos, and Minkin (2000) and Kourtellos (2001). Durlauf and Johnson employ a tree-regression approach to uncover multiple regimes in the data while Hansen proposes a threshold regression model that leads to a formal test for the presence of a regime change. Liu and Stengos employ a semi-parametric specification test and an additive semiparametric partially linear model to identify nonlinear growth patterns. Canova (1999) uses a predictive density approach, Desdoigts employs an exploratory projection pursuit (density estimation).

An interesting special form of nonlinearity is coefficient heterogeneity: the coefficients of the cross-country growth regression vary across countries. Kourtellos (2001) assumes a local Solow growth model in the sense that a Cobb-Douglas aggregate production function applies to each country but the parameters of this function vary across countries. In particular, he exploits an interesting form of interactions effects by employing a varying coefficient model with a conditional linear structure where the coefficients of the Solow-type variables vary across countries according to initial conditions or other country specific characteristics.

This paper takes an even more gerenal approach by assuming no specific aggregate production function and hence no specific functional form for the cross-country growth process. In particular, it models the cross-country growth process using projection pursuit regression (PPR). The PPR model is a semi-parametric way of modeling nonlinearities in particular interaction effects. Interactions are important in understanding how the parameters of the aggregate production function vary across countries. The projection pursuit model has an advantage over other semiparametric techniques, for a number of reasons. First of all, the smoothing is always one-dimensional and thus the PPR model overcomes the problem of the 'curse of

dimensionality'. Therefore, projection pursuit regressions can be easily, quickly and accurately estimated with relatively low variance. Further, the PPR models are quite general. They are capable to approximate a much richer class of functions than Generalized Additive Models (GAM); see Hastie and Tibshirani (1990). Unlike GAM, PPR model allows for interactions by modeling the regression surface as a sum of general smooth functions of linear combinations of the predictors. Furthermore, under certain identifying restrictions, several parametric and non-parametric models can be viewed as special cases of the PPR model: the Additive Model (with or without interactions), the Partially Linear Additive Model (with or without interactions), the Single Index Model (Horowitz (1999)), the Partial Linear Single Index Model (Xia et al (1999), the Linear Regression Model, etc.

My results show that there exist nonlinearities that characterize the effects of basic Solow-type variables on growth. In particular, my results emphasize the existence of interaction effects human capital and initial income as well as between initial income and population growth rates. These findings seem to suggest the existence of two steady state equilibria in the growth process.

2 Projection Pursuit Regression (PPR) Model

Projection Pursuit Regression (PPR) model was proposed to approximate high dimensional functions, $m(\mathbf{x}_i) = E(g_i/\mathbf{X}_i = \mathbf{x}_i)$, $\mathbf{x}_i \in \mathbb{R}^p$ by simpler functions that operate in low ($d << p$) dimensional spaces - typically for $d = 1$. The model of regression surface is based on projections of the data on planes spanned by the endogenous variable g and a linear combination $\mathbf{a}'\mathbf{x}$ of the explanatory variables. The idea of PPR is to approximate the mean regression function by a sum of unknown ridge functions $m(\mathbf{x}_i) = \sum_{j=1}^{M} \beta_j f_j(\boldsymbol{\alpha}'_j \mathbf{x}_i)$; Diaconis and Shahshahani (1984) provide an approximation theory justification as the number of ridge functions (M) goes to infinity.

In particular, given a random vector $(\mathbf{X}, g) \in \mathbb{R}^{p+1}$ from a random sample, PPR model is defined in form of:

$$g_i = \beta_0 + \sum_{j=1}^{M} \beta_j f_j(\boldsymbol{\alpha}'_j \mathbf{x}_i) + u_i, \ i \in \mathbb{I} \tag{1}$$

n,where $f_j(\cdot) \in C^2$ for $j = 1...M$, $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_M)'$ are the regression coefficients, and $\boldsymbol{\alpha}_j = (\alpha_{1j}, \alpha_{2j}, ..., \alpha_{pj})'$ are the $p \times 1$ direction vectors. For identification reasons we need to assume that $E\{f_j(\boldsymbol{\alpha}'_j \mathbf{x}_i)\} = 0^2$, $E\{f_j(\boldsymbol{\alpha}'_j \mathbf{x}_i)\}^2 = 1$ and, $\boldsymbol{\alpha}'_j \boldsymbol{\alpha}_j = 1$ for $j = 1...M$. The parameters of interest are given by the parameter vector $\boldsymbol{\varphi}$

$$\boldsymbol{\varphi} := (\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2) \in \mathbb{R}^{pM} \times \mathbb{R}^{M+1} \times \mathbb{R}_+. \tag{2}$$

---
[2]Notice that assumption implies that $\beta_0 = E(y)$,

with $\boldsymbol{\beta} := (\beta_1, ..., \beta_M)$ reflecting the regression coefficients for the non-linear and linear part, respectively; $\boldsymbol{\alpha} := (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, .., \boldsymbol{\alpha}_M)$ denoting the $p \times M$ direction vectors; M is the number of ridge functions (terms); σ^2 the conditional variance of the model. The $Rank(\mathbf{x}) = p$, where $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)' : n \times p$. The probability model is specified in terms of the first and second conditional moments

$$E(g_i \mid \mathbf{X}_i = \mathbf{x}_i) = \beta_0 + \sum_{j=1}^{M} \beta_j f_j(\boldsymbol{\alpha}_j' \mathbf{x}_i) \tag{3}$$

$$Var(g_i \mid \mathbf{X}_i = \mathbf{x}_i) = \sigma^2 \tag{4}$$

as well as some conditions regarding the nature of the unknown functions. In particular, the functions $f_1, ..., f_M$ are elements of the Sobolev space, where $\mathcal{H}_j^2(\mathbb{R}) = \{f_j \in C^1(\mathbb{R}) \mid f_j' \text{ absolutely continuous and } f_j'' \in L_2(\mathbb{R})\}$ for $j = 1, .., M$ and the reproducing kernel $(\mathcal{R}_j)$ associated with the space $\mathcal{H}_j^2(\mathbb{R})$, $\boldsymbol{\Sigma}_j = \{\mathcal{R}_j(v_s, v_t)\}$, is a positive definite matrix. Finally, the sampling model is assumed to be an independent identically distributed random sample $(\mathbf{x}, g)$ with $g = (g_1, ..., g_n)'$ drawn from a well behaved distribution $F(g, \mathbf{X})$.

The PPR model is quite general in the sense that several semipametric or parametric models can be viewed as nested hypotheses to the PPR. Specifically, with suitable choices of β, α, M, and $f_j's$ we can get the single index model , the additive model , the partial linear model , the partial linear additive model, simple linear model, etc.

3 Estimation

PPR estimation procedure was introduced by Friedman and Stuetzle (1981) and refined by Friedman (1984b). For given number of terms, M, the model parameters $\boldsymbol{\varphi}$ and the unknown functions $f_1, ..., f_M$ are estimated by minimizing the criterion L_2

$$L_2 = E\left[g_i - \beta_0 - \sum_{j=1}^{M} \beta_j f_j(\boldsymbol{\alpha}_j' \mathbf{x}_i)\right]^2 \tag{5}$$

using the alternating optimization strategy.

Following Friedman (1984b), for a particular term k, the above criterion L_2 can be written as

$$L_2 \equiv L_2^k = E\left[r_i^k - \beta_k f_k(\boldsymbol{\alpha}_k' \mathbf{x}_i)\right]^2 \tag{6}$$

with

$$r_i^k = g_i - \beta_0 - \sum_{j \neq k}^{M} \beta_j f_j(\boldsymbol{\alpha}_j' \mathbf{x}_i) \tag{7}$$

The criterion L_2^k is minimized with respect to the parameters $(\beta_k, f_k, \boldsymbol{\alpha}_k)$ to give the kth projective approximation to $m(\mathbf{x})$, defined by $\beta_k f_{\boldsymbol{\alpha}_k}(\boldsymbol{\alpha}_k' \mathbf{x}_i)$. These parameters

define new residuals $r_i^{k'}, k \neq k'$, to obtain new solutions for the parameters of other terms. This method is iterated over all terms until convergence.

Let us focus on the estimation of the parameters of the k_{th} term given r_i^k. Notice that, given f_k and $\boldsymbol{\alpha}_k$, the solution for β_k is readily given by least squares theory

$$\beta_k^* = \frac{E\left(r_i^k f_k(\boldsymbol{\alpha}_k'\mathbf{x}_i)\right)}{E(f_k(\boldsymbol{\alpha}_k'\mathbf{x}_i))^2} \tag{8}$$

Hence one needs to obtain first the smooth function f_k and the projection $\boldsymbol{\alpha}_k$.

Using the law of iterated expectations (6) can be reexpressed as

$$L_2^k = E_{\boldsymbol{\alpha}_k'\mathbf{x}_i} E\left(\left[r_i^k - \beta_k f_k(\boldsymbol{\alpha}_k'\mathbf{x}_i)\right]^2 \mid \boldsymbol{\alpha}_k'\mathbf{x}_i\right) \tag{9}$$

L_2 is then minimized if f_k is chosen to minimize the conditional loss criterion in (9) for each value of $\boldsymbol{\alpha}_k$. By projection theorem the solution of this problem is given by k^{th} projective approximation

$$f_{\boldsymbol{\alpha}_k}^*(\boldsymbol{\alpha}_k'\mathbf{x}_i) = \frac{1}{\beta_k^2} E\left(r_i^k \mid \boldsymbol{\alpha}_k'\mathbf{x}_i\right) \tag{10}$$

for all $k = 1, 2, ..., M$. Notice that the assumptions that $E(f_k) = 0$ and $E(f_k^2) = 1$ make the regression coefficient β_k irrelevant. Equation (10) leads to the formal representation

$$\begin{pmatrix} \mathbf{I} & \mathbf{P}_1 & \mathbf{P}_1 & \cdots & \mathbf{P}_1 \\ \mathbf{P}_2 & \mathbf{I} & \mathbf{P}_2 & \cdots & \mathbf{P}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_M & \mathbf{P}_M & \cdots & \mathbf{P}_M & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{f}_{\boldsymbol{\alpha}_1}^* \\ \mathbf{f}_{\boldsymbol{\alpha}_2}^* \\ \vdots \\ \mathbf{f}_{\boldsymbol{\alpha}_M}^* \end{pmatrix} = \begin{pmatrix} \mathbf{P}_1\mathbf{g} \\ \mathbf{P}_2\mathbf{g} \\ \vdots \\ \mathbf{P}_M\mathbf{g} \end{pmatrix} \tag{11}$$

where $\mathbf{P}_j = E(\cdot \mid \boldsymbol{\alpha}_k'\mathbf{x})$, $g = (g_1, g_2, \ldots, g_n)'$ and $f_{\boldsymbol{\alpha}_k}^* = (f_{\boldsymbol{\alpha}_k}^*(\boldsymbol{\alpha}_k'\mathbf{x}_1), \ldots, f_{\boldsymbol{\alpha}_k}^*(\boldsymbol{\alpha}_k'\mathbf{x}_n))'$. This system of equations can in principle be solved exactly but it can be very computationally expensive; it requires $O(nM)$ operations to solve the $nM \times nM$ system. In order to cope with this problem we employ the backfitting (Gauss-Seidel) algorithm: given initial values $\widehat{f}_j^{[0]}$ and $\widehat{f}_j^{[0]}$, the respective l^{th} step backfitting estimators are

$$\widehat{\mathbf{f}}_k^{*[l]} = \mathbf{P}_k^{[l]}\left(\mathbf{g} - \sum_{k \neq j}^M \widehat{\mathbf{f}}_j^{[l-1]}\right) \tag{12}$$

,where $\mathbf{P}_k^{[l]}$ is the $n \times n$ positive-definite univariate cubic spline smoother matrix. The algorithm is repeated until some prespecified tolerance is reached. Finally, it remains to minimize L_2^k in (6) with respect to the projection direction $\boldsymbol{\alpha}_k$ with β_k and f_k fixed. The latter problem as well as the smoothing problem will be further discussed below.

The theoretical properties of PPR estimation procedure were investigated by Hall (1989), Huber (1985) and Jones (1985), Jones (1992). Hall showed that the estimation of the direction α can be done with $0(n^{-1/2})$ while the nonparametric estimation[3] of each term f_j has an order worse than $n^{-1/2}$. Huber established weak L_2-convergence, Jones (1987), based on a generalized PPR procedure, proved strong convergence of the PPR estimation procedure and Jones (1992) established a $O(1/\sqrt{M})$ non-sampling convergence rate.

3.1 Smoothing Problem

Friedman and Stuetzle (1981) and Friedman (1984b) estimate the smooth functions using the supersmoother. The supersmoother is two-dimensional nonlinear variable span smoother based on local linear fits, in which local cross-validation is used to estimate the optimal span (see Frideman (1984a). Hwang et al. (1994) use orthonormal Hermite polynomial functions to estimate the smooth functions. Their motivation of using these polynomials instead of the supersmoother was mainly based on the desirable properties of polynomials. They noted that polynomials have fast and accurate derivative calculation and a smooth interpolation. Following Hastie and Roosen (1994) this paper employs smoothing splines as they share the same properties as the polynomials and in addition they can be used to choose the smoothing parameters and the number of terms automatically.

In particular, the smooth function $f_{\boldsymbol{\alpha}_k}(v_{ik})$ with $v_{ik} = \boldsymbol{\alpha}'_k \mathbf{x}_i$ is estimated by minimizing the penalized least squares criterion $\mathcal{S}^k_{\boldsymbol{\lambda}}$

$$\mathcal{S}^k_{\boldsymbol{\lambda}} = \sum_{i=1}^n \left(r_i^k - \beta_k f_{\boldsymbol{\alpha}_k}(v_{ik}) \right)^2 + \lambda_k \parallel f''_{\boldsymbol{\alpha}_k} \parallel_2^2 \tag{13}$$

where $\parallel f''_{\boldsymbol{\alpha}_k} \parallel_2^2 = \int \left(f''_{\boldsymbol{\alpha}_k}(t) \right)^2 dt$ and $\lambda_k > 0$ is the smoothing parameter. Notice that as $\lambda_k \to \infty$, we get the linear model with $p+1$ degrees of freedom and with $\lambda_k \to 0$ we interpolate the data with n degrees of freedom. The solution is given by the cubic smoothing spline. The Sobolev space $\mathcal{H}_j^2$ is supposed to have the decomposition $\mathcal{H}_k^2 = \mathcal{H}_0 \oplus \mathcal{H}_k$, where $\mathcal{H}_0$ is spanned by 2-dimensional polynomials ϕ_1 and ϕ_2 containing terms that are not penalized. The solution is known to be in terms of a basis $\{ \phi_v \}$ for $\mathcal{H}_0$ and the r.k.'s $\mathcal{R}_k$ for the $\mathcal{H}_k$. Let $\mathbf{T}$ be the $n \times 2$ matrix of evaluations of ϕ_j, $\mathbf{T} = \{\phi_1(v_i), \phi_2(v_i)\}_{i=1}^n$. Given assumption [8] the minimizer to (13) is known to have the form (see, Kimeldorf and Wahba (1971))

$$\widehat{f}_{\lambda_k, \boldsymbol{\alpha}_k}(v) = d_{j1}\phi_{j1}(v) + d_{j2}\phi_{j2}(v) + \sum_{i=1}^n c_i \mathcal{R}_k(v_i, v) \tag{14}$$

---

[3]Hall used a kernel smoother but any other smoother with random or non-random bandwidth has similar implications.

where $\boldsymbol{\xi}'(v) = (\mathcal{R}_k(v_1, v), \mathcal{R}_k(v_2, v), ..., \mathcal{R}_k(v_n, v))$ and $\boldsymbol{\phi}'(v) = (\phi_1(v), \phi_2(v))$ and $\mathbf{d} = (d_1, d_2)'$ and $\mathbf{c} = (c_1, ..., c_n)'$ are such that

$$
\begin{aligned}
(\boldsymbol{\Sigma}_k + n\lambda_k \mathbf{I})\mathbf{c} + \mathbf{Td} &= \mathbf{g} \\
\mathbf{T}'\mathbf{c} &= 0
\end{aligned}
\tag{15}
$$

To compute $\mathbf{c}$ and $\mathbf{d}$, let the QR decomposition of $\mathbf{V}_k$, where $\mathbf{V}_k = (v_{1k}, v_{2k}, ..., v_{nk})'$ be $\mathbf{V}_k = \begin{pmatrix} \mathbf{Q}_{1k} & \mathbf{Q}_{2k} \end{pmatrix} \begin{pmatrix} \mathbf{R}_k \\ 0 \end{pmatrix}$ where $\begin{pmatrix} \mathbf{Q}_{1k} & \mathbf{Q}_{2k} \end{pmatrix}$ is an orthogonal matrix and $\mathbf{R}$ is upper triangular, with $\mathbf{V}_k'\mathbf{Q}_{2k} = 0$. By letting $\widehat{\mathbf{f}}_{\lambda_k, \alpha_k}$ denote the predicted data $\widehat{\mathbf{f}}_{\lambda_k, \alpha_k} = (f_{\lambda_k, \alpha_k}(v_{1k}), f_{\lambda_k, \alpha_k}(v_{2k}), ..., f_{\lambda_k, \alpha_k}(v_{nk}))'$, then one can rewrite the above solution in a convenient computational form using the influence matrix $\mathbf{A}(\lambda_k)$,

$$
\widehat{\mathbf{f}}_{\lambda_k, \alpha_k} \equiv \mathbf{A}(\lambda_k)g
\tag{16}
$$

with $\mathbf{A}(\lambda_k) = \mathbf{I}_n - n\lambda_k \mathbf{Q}_{2k}(\mathbf{Q}_{2k}'(\boldsymbol{\Sigma}_k + n\lambda_k \mathbf{I}_n)\mathbf{Q}_{2k})^{-1}\mathbf{Q}_{2k}'$; see (Wahba (1990), pp. 13).

3.2 Projection Direction

The kth projection direction vector $\boldsymbol{\alpha}_k$ is estimated by minimizing

$$
S(\boldsymbol{\alpha}_k) = \sum_{i=1}^{n} \left(r_i^k - \widehat{\beta}_k \widehat{f}_{\alpha_k}(\boldsymbol{\alpha}_k' \mathbf{x}_i) \right)^2
\tag{17}
$$

Notice that the criterion L_2^k is not a quadratic function of $\boldsymbol{\alpha}_k$ with β_k and f_k fixed therefore an iterative optimization method must be performed. Here, I use the Gauss-Newton method. Assuming that the optimal projection direction $\widehat{\boldsymbol{\alpha}}_k$ is an increment Δ from the current direction $\boldsymbol{\alpha}_k$ and let $u_i = r_i^k - \widehat{\beta}_k \widehat{f}_{\alpha_k}(\boldsymbol{\alpha}_k' \mathbf{x}_i)$ then Δ can be obtained by regressing residuals u_i, based on the current direction, on the estimated partial derivative $\frac{\partial u_i}{\partial \alpha_k} = -\beta_k \widehat{f}_k'(\boldsymbol{\alpha}_k' \mathbf{x}_i)$, evaluated at the previous update value of projection direction. Cubic smoothing splines produced with piecewise cubic basis functions have an obvious advantage in obtaining derivatives over other linear smoothers. For a detailed presentation of the Gauss-Newton optimization procedure see Friedman (1984b) and Hwang et al. (1994).

3.3 Smoothing Parameters

All the smoothing and model parameters $\boldsymbol{\psi} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})$ can be chosen by a generalized cross validation (GCV) criterion

$$
GCV(\boldsymbol{\psi}) = \frac{\frac{1}{n}\sum_{i=1}^{n}(g_i - \widehat{m}_{\boldsymbol{\psi}}(\mathbf{v}_i))^2}{(1 - \frac{df_{\boldsymbol{\psi}}}{n})^2}
\tag{18}
$$

8

where df_{ψ} denotes the effective number of parameters used in the fit and is defined as $df_{\psi} \equiv tr(\mathbf{S}_{\psi})$ where $\mathbf{S}_{\psi}$ denotes the PPR smooth operator matrix using the parameters ψ. This definition of df_{ψ} is suggested by Hastie and Tibshirani (1990) to generalize the common notion of degrees of freedom of a linear model. However, this could be an extremely complicated non-linear optimization. In order to avoid this problem Hastie and Roosen (1994) proposed to evaluate the trace of $\mathbf{S}_{\psi}$ based on an approximation in a cyclic manner. The trace of the smooth is approximated as $tr(\mathbf{S}_{\psi}) = M \cdot d + \sum_{j=1}^{M} tr(\mathbf{S}_{\lambda_j})$, where $tr(\mathbf{S}_{\lambda_j})$ is the trace of the smoothing matrix for smooth j. For each term, the linear combination $\mathbf{a}_j$ is charged $p-1$ degrees of freedom, and 1 degree of freedom for the scale coefficient β_j. The optimization over directions and smoothing parameters is performed by cycling through each direction, freezing the parameters for all but one term j and minimizing the partial residuals for the remaining term. Let

$$tr(\mathbf{S}_{\psi})_{offset} = M \cdot d + \sum_{k \neq j} tr(\mathbf{S}_{\lambda_k}) \tag{19}$$

denote the degrees of freedom for all model components except the smooth f_k under consideration. Notice that r_i^k and $tr(\mathbf{S}_{\psi})_{offset}$ are constant with respect to the smoothing parameter λ_k used in f_k. The GCV criterion can then be written as

$$GCV(\psi) = \frac{\frac{1}{n} \sum_{i=1}^{n} \left(r_i^k - \widehat{f}_{\boldsymbol{\alpha}_k, \lambda_k} \left(\boldsymbol{\alpha}_k' \mathbf{x}_i \right) \right)^2}{\left(1 - \frac{1}{n} \left[tr(\mathbf{S}_{\boldsymbol{\lambda}}^k) + tr(\mathbf{S}_{\psi})_{offset} \right] \right)^2} \tag{20}$$

,where $\widehat{f}_{\boldsymbol{\alpha}_k, \lambda_k} \left(\boldsymbol{\alpha}_k' \mathbf{x}_i \right) = \mathbf{S}_{\lambda_k} \cdot r_i^k$.

4 Empirical Results

4.1 Data

The data is a pooled cross-country dataset averaged over the periods $1960-69, 1970-79, 1980-89$ based on King and Levine (1988). This dataset was chosen over alternative datasets on the grounds of comparisons with previous studies and available sample size of a balanced panel. The list of countries included in our dataset can be found in Liu and Stengos (1999). The explanatory variables comprise the observable variables suggested by the Solow growth model. They include (i) *gpop*, logarithm of average growth rate of the population plus 0.05 for depreciation; (ii) *inv*, logarithm of average proportion of real investments (including government) to real GDP; (iii) *h*, logarithm of average years of secondary schooling in the total population in each period; (vi) y_0, logarithm of initial per capita income. Figure 1 presents kernel density estimates of the variables used along with the 95% confidence intervals. The

quartic kernel , $k(x) = \frac{15}{16}(1-x^2)^2 I(\mid x \mid \leq 1)$ was used and the bandwidth was based on the Silverman's rule of thumb. Income growth rates, g, reflect residuals from a regression of the change in the log of income per capita over the 1960s, the 1970s, and 1980s on three time dummies.

4.2   Projections

The estimation procedure is implemented by the Automatic Smoothing Spline Projection Pursuit (ASP) proposed by Hastie and Roosen (1994). Appendix 1 summarizes the ASP procedure. GCV criterion picked one non-linear flattened s-shaped projection which seemed to be robust in changes of degrees freedom and in random initial values for the staring direction. Figure 2 shows the projection where for low values of the first direction the projection is apparently u-shaped while for high values the projection is more like an inverted stretched u-shape. Tables 1(a) and 1(b) shows the first direction for the first projection obtained by GCV and the $\sqrt{n} - consistent$ regression coefficients.

A second projection of a quadratic shape was not picked by GCV as significant. This could be a sensitivity for the overall small sample size. If I specify the number of terms in Friedman style we get a second projection of a quadratic shape in addition to the one above. Figure 3(a) and 3(b) show the two projections. It is worth noticing that the first projection is approximately the same with the one picked by GCV. Table 2(a) and 2(b) present the direction vectors and the regression coefficients.

4.3   Interpretation

Having established that there are interesting (non-linear) projections in cross-country growth data I proceed to interpretation. Although the projection directions do not have any direct interpretation, combining the predictive ability of the smooths with the directions one can extract valuable information regarding the influence of the Solow-type variables on growth. Figures 3 (a)-(d) show the relation between predicted growth and the Solow-type variables $y0$, h, inv, $gpop$, respectively along with 95% pointwise confidence intervals. The superimposed horizontal line refers to the corresponding least square constant coefficient from a linear regression. The results are quite revealing. Figure 3a shows the relation between predicted growth and initial income per capita, $y0$. It suggests that the convergence hypothesis is only true for countries with per capita initial income higher than \$1200, which corresponds to Somalia. For countries with lower per capita income than \$1200 the relationship between growth and initial income per capita is positive and therefore the convergence hypothesis is not valid. What is more, the least square line cuts the 95% confidence interval in two occasions suggesting that the linear model is mispecified. This result is similar to that found by Stengos and Liu (1999) based on an additive partial linear model. This results is also suggestive of the presence of two steady state equilibria

10

in the growth process with the twin peaks found by Quah (1997) in the cross-country income distribution. Figure 4b shows the relation between growth and secondary enrollment rates, respectively. For very low and medium values of the sample the schooling influence on growth appears to be a positive while for high enrollment levels it appears to taper off. Figure 3(c) and 3(d) reflect the relationship between growth and investment and growth and population growth, respectively. For these variables there is no strong evidence of the presence of nonlinearities.

As I mentioned earlier in the paper, the power of the projection pursuit regression stems from its ability to capture interactions. These interaction effects are largely ignored by additive models. It is worth noting that Figures 3(a)-(d) are incapable of presenting interaction effects because all the variables but the one in question are held fixed at their mean values. Instead, I use coplots of predicted growth $\widehat{g}$ against pairs of Solow-type variables. Figure 4 shows the coplot of $\widehat{g}$ against y_0, given h. The dependence panels are the 3×3 array, and the given panel is at the top. On each dependence panel, $\widehat{g}$ is graphed against y_0 for those observations whose values of h lie in a given interval. The intervals are shown on the given panel; as we move from left to right through these intervals, we move from left to right and then bottom to top through the dependence panels. For very low levels of human capital the relationship between $\widehat{g}$ and y_0 is rather linear with downward slope. For countries with higher levels of human capital this relationship becomes nonlinear with a quadratic shape similar to Figure 3(a). This behavior suggests the presence of more than one steady-state equilibrium in the growth process with the low initial income and low human capital countries converging to a different equilibrium from the high initial income and high human countries. Furthermore, for low initial income and higher human capital countries there is no evidence of convergence.

Figure 5 shows the coplot of $\widehat{g}$ against h, given y_0. For very low initial income countries and high initial income, the relationship between $\widehat{g}$ and h is linear with positive slope as the one predicted by the Solow growth model. For middle initial income countries, this relationship becomes quadratic. This graph suggests the presence of interaction between initial income and human capital such that the positive effect of human capital on growth is only true for countries with low and high initial income.

Figure 6 shows the coplot of $\widehat{g}$ against y_0, given *gpop*. For countries with low population growth rates - i.e. mostly industrialized countries - the relationship between $\widehat{g}$ and y_0 is linear and downward sloping. This suggests that at least for most of the industrialized countries the relationship between growth and initial income can be taken to be linear as that predicted from the Solow growth model. For countries with higher population growth rates - i.e. mostly developing and undeveloped countries - the relationship between $\widehat{g}$ and y_0 is quadratic similar to Figure 3(a). This suggests the presence of interaction between initial income and population growth that prevents "poor" countries with high population growth rates to converge to the same steady state equilibrium as the "poor" countries with low population growth rates.

5   Conclusion

This paper introduces a new way of modeling cross-country growth data by the means of projection pursuit regression. It demonstrates that projection pursuit is capable of modeling deep non-linearities and in particular, interaction effects in the cross-country growth process which are generally ignored by previous studies. I find that initial income and human capital affect growth in a very nonlinear way. Furthermore I find evidence of interaction effects between human capital and initial income as well as between initial income and population growth rates. These interaction effects provide a richer understanding of cross-country growth differences. Overall my empirical results suggest the presence of two steady state equilibriums in the growth process that classify the countries into two regimes. A full understanding of cross-country growth differences will need to explain the existence of interaction effects and to test the projection pursuit model against other nested models.

References

[1] Bernard, A. B., and Durlauf, S. N. (1996), "Interpreting Tests of the Convergence Hypothesis", Journal of Econometrics, 71, 161-174.

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

[3] Barro, R. (1997), "Determinants of Economic Growth", The MIT Press, Cambridge, Massachusetts.

[4] Barro, R. and Sala-I-Martin, X. (1995), "Economic Growth", New York: McGraw-Hill.

[5] Diaconis, P. and Shahshahani, M. (1984). On Nonlinear Functions of Linear Combinations. SIAM J. Sci and Stat. Comp., 5, 175-191.

[6] Durlauf, S. and Brock, W. (2000), "Growth Economics and Reality", Working Paper, Department of Economics, University of Wisconsin.

[7] Durlauf, S. and Johnson, P. (1995), "Multiple Regimes and Cross-Country Growth Behavior", Journal of Applied Econometrics, 10, 365-384.

[8] Durlauf, S. and Quah, D. (1999), "The New Empirics of Economic Growth", Handbook of Macroeconomics, Vol. 1, Ch.4, pp.235-308.

[9] Friedman, J. H., Stuetzle, W. (1981), "Projection Pursuit Regression", Journal of American Statistical Association, 76, 817-823.

[10] Friedman, J.H. (1984a). A Variable Span Smoother. Tech. Report 5, Laboratory for Computational Statistics, Dept. of Statistics, Stanford University, CA.

[11] Friedman, J.H. (1984b) "SMART User's Guide", Technical Report No.1, Stanford University, Dept. of Statistics.

[12] Friedman, J. H. (1987). "Explaratory Projection Pursuit", Journal of American Statistical Association, 82,, 249-266.

[13] Hall, P. (1989), "On Projection Pursuit Regression", The Annals of Statistics, 17, 573-588.

[14] Hansen, B. (2000), "Sample Splitting and Threshold Estimation", Econometrica, 68, pp. 575-603

[15] Hastie T. J.and Tibshirani R. J. (1990). "Generalized Additive Models", Monographs on Statistics and Applied Probability 43, Chapman and Hall.

[16] Hastie T. J.and Tibshirani R. J. (1993). "Varying Coefficient Models", Journal of Royal Statistical Society B, 55, 757-796.

[17] Hastie, T. and Roosen, C. (1994), "Automatic Smoothing Spline Projection Pursuit, Journal of Computational and Graphical Statistics", 3(3):235-248.

[18] Horowitz J. L. (1999). Semiparametric Methods in Econometrics, Lecture Notes in Statistics, Springer, ch.2.

[19] Hwang, J.N., Lay, S. R., Maechler, M., Martin, D., and Schimert, J. (1994), "Regression Modeling in Back-Propagation and Projection Pursuit Learning", IEEE Transations on Neural Networks, 5, 3,342-353..

[20] Huber, P.J. (1985), " Projection Pursuit" (with discussion), The Annals of Statistics, 13, 435-525.

[21] Ichimura and Lee (1991). Semiparametric Least Squares Estimation of Multiple Index Models: Single Equation Estimation, in Nonparametric and Semiparametric Methods in Econometrics and Statistics, W. A. Barnett, J. Powell, and G. Tauchen, eds., Cambridge CUP, ch.1.

[22] Jones, L. (1985), "On a Conjecture of Huber Concerning the Convergence of Projection Pursuit Regression", The Annals of Statistics, 15, pp. 880-882.

[23] Jones, L. (1992), "A Simple Lemma on Greedy Approximation in Hilbert Space and Convergence Rates for Projection Pursuit Regression and Neural Network Training", The Annals of Statistics, 20, pp. 608-613.

[24] Kimerdorf, G. and Wahba, G. (1971), "Spline Functions on Tchebycheffian Spline Functions", J. Math. Anal. Appl., 33, pp. 82-95.

[25] Levine, D. and Renelt, D. (1992), "A Sensitivity Analysis of Cross-Country Growth Regressions", American Economic Review 82, 942-963.

[26] Lingjerde O. C. and Liestol K. (1998), "Generalized Projection Pursuit Regression", SIAM, J. Sci. Comput., Vol. 20, No.3, pp. 844-857.

[27] Liu, C. and Stengos, T. (1999), "Non-Linearities in Cross Country Growth Regressions: A Semiparametric Approach", Journal of Applied Econometrics, 14, 5, pp.527-538.

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

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

14

[30] Pack, H., (1994), "Endogenous Growth Theory: Intellectual Appeal and Empirical Shortcomings", Journal of Economic Perspectives, 8, 1, 55-72.

[31] Sala-I-Martin, X. (1999), "I Just Ran Four-Million Regressions", NBER Working Paper 6252.

[32] Sala-I-Martin, X. (1997), "I just Ran Two-Million Regressions", American Economic Review, 87 (2), pp.178-183.

[33] Solow, R., (1956), "A Contribution to the Empirics of Economic Growth," Quarterly Journal of Economics, 70, 65-94.

[34] Solow, R., (1958), "A Skeptical Note on the Constancy of Factor Shares," American Economic Review, 48, 618-631.

[35] Wahba, G., "Spline Models for Observational Data", Philadelphia, PA, SIAM.

[36] Xia, Y., Tong, H., and Li, W. K. (1999), "On Extended Partially Linear Single-Index Models", Biometrika, 86, 4, pp. 831-842.

Figure 1a - Density estimate of growth
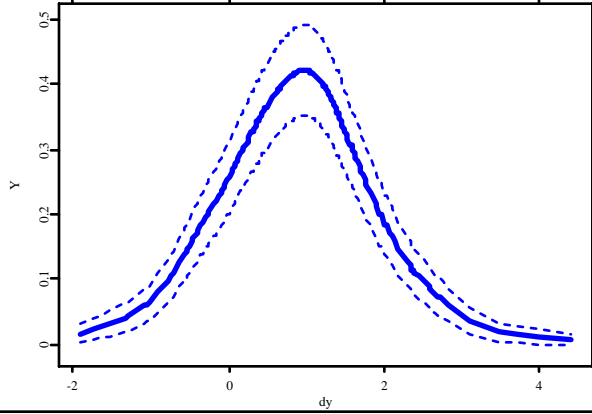
Figure 1b - Density estimate of population growth
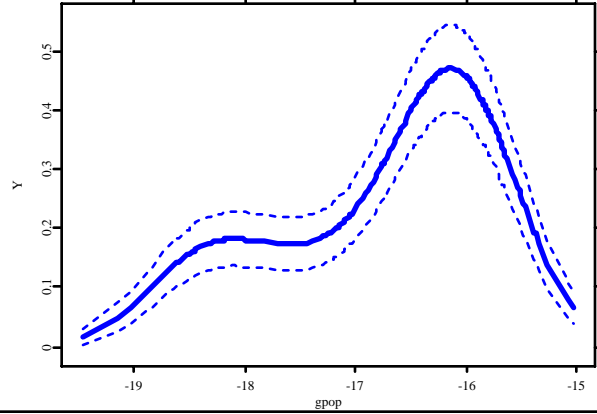
Figure 1c - Density estimate of investments
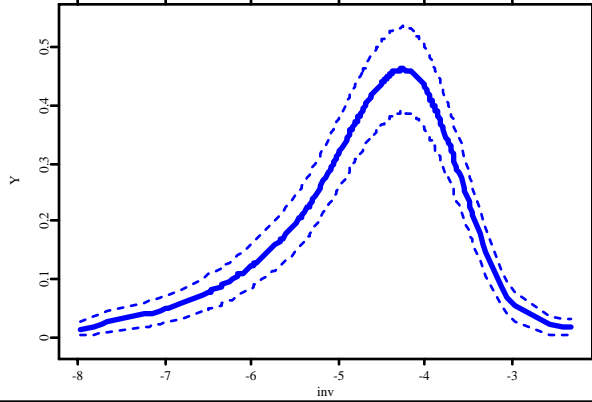
Figure 1d - Density estimate of initial income
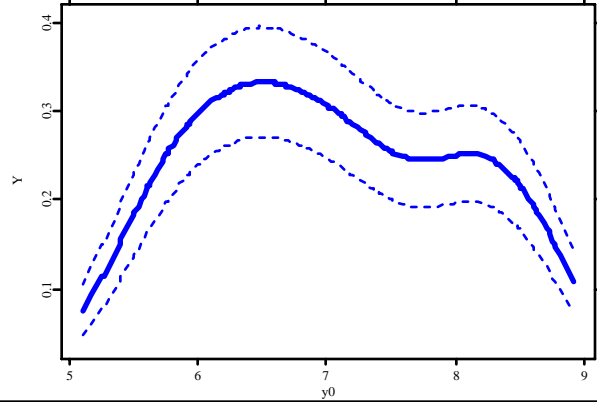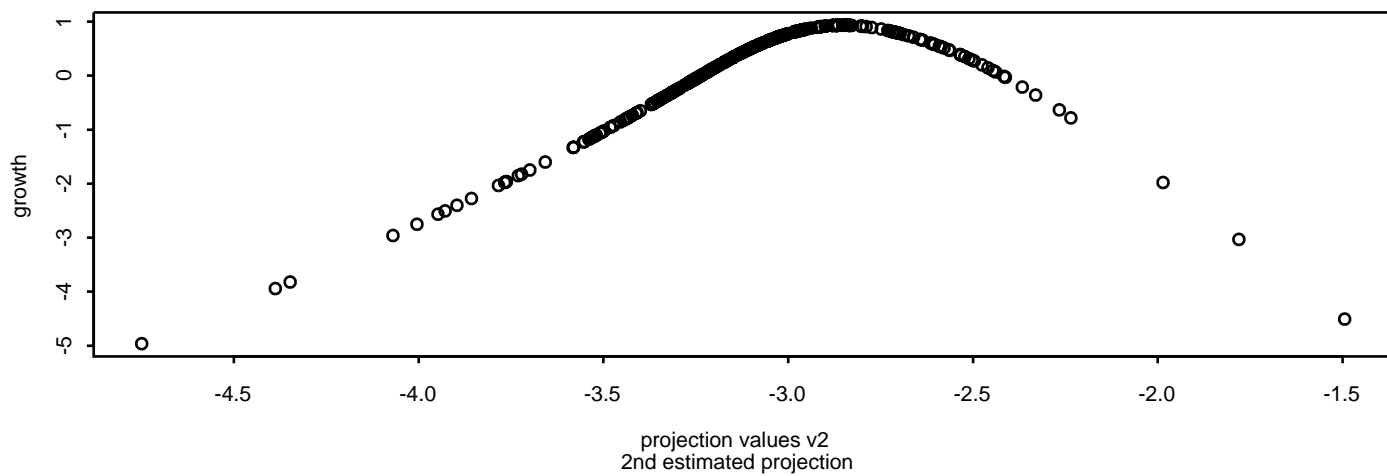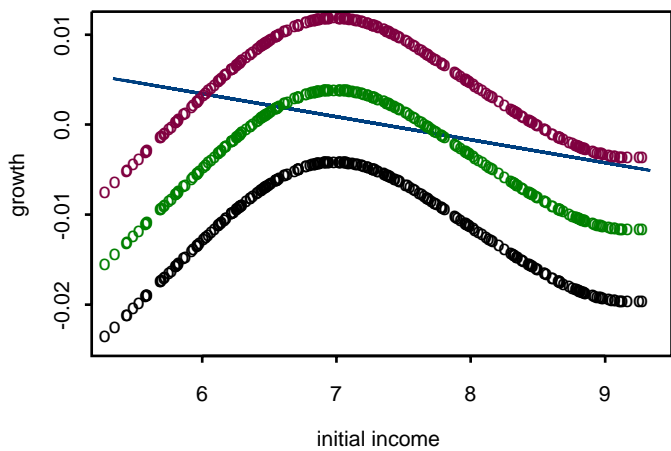
Figure 1e - Density estimate of initial human capital (h)
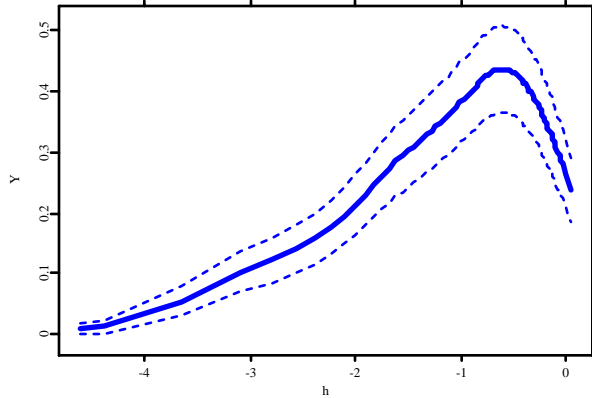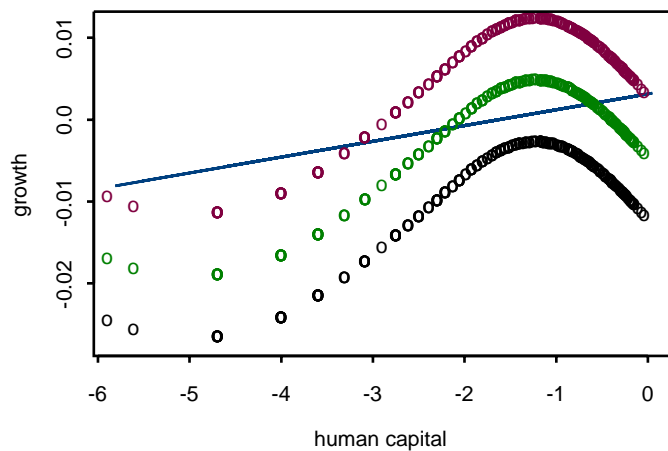
Figure 2a
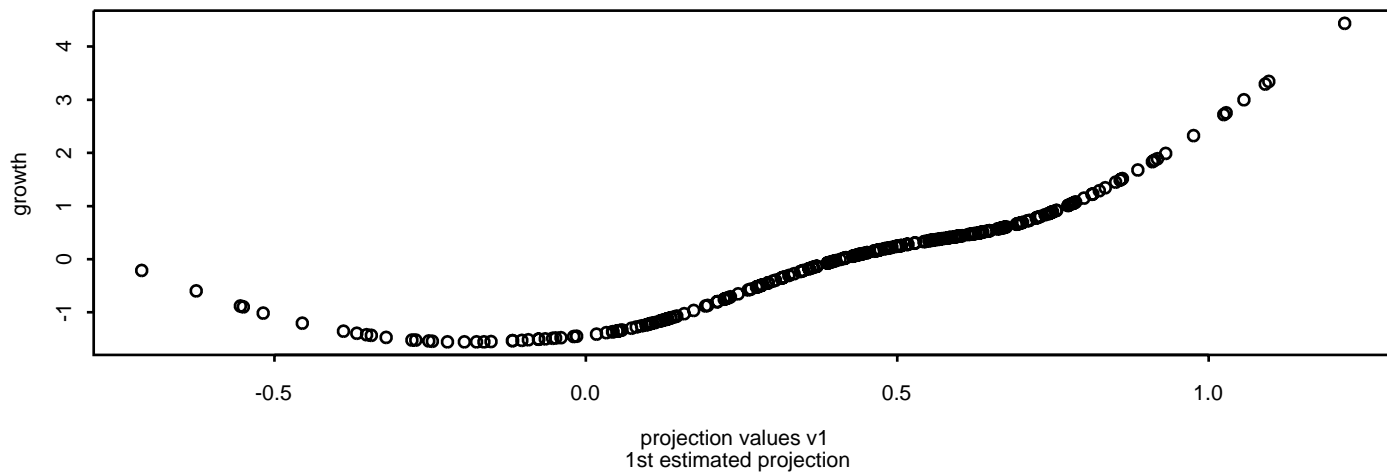
Figure 2b
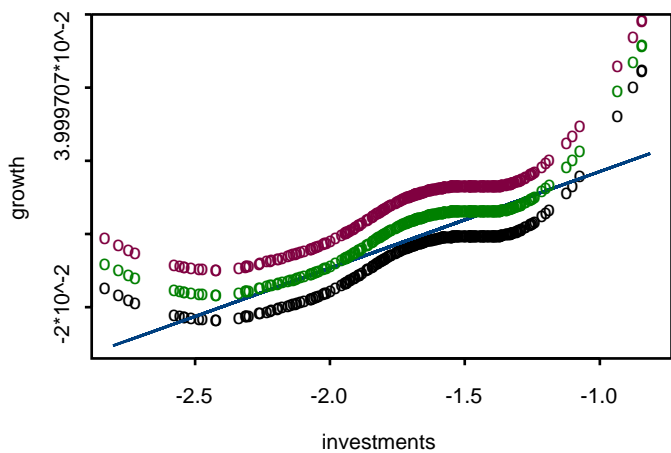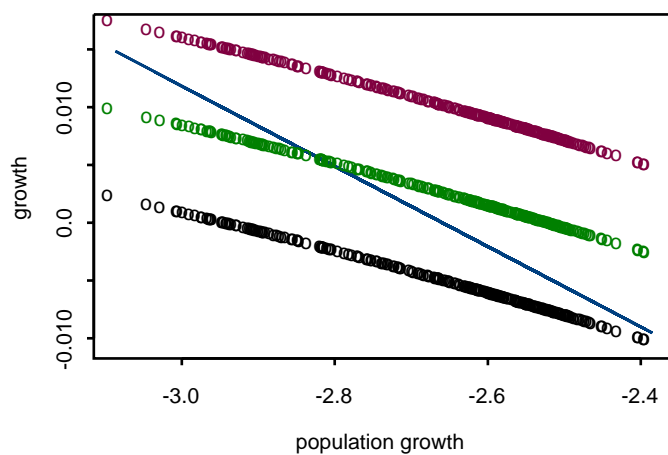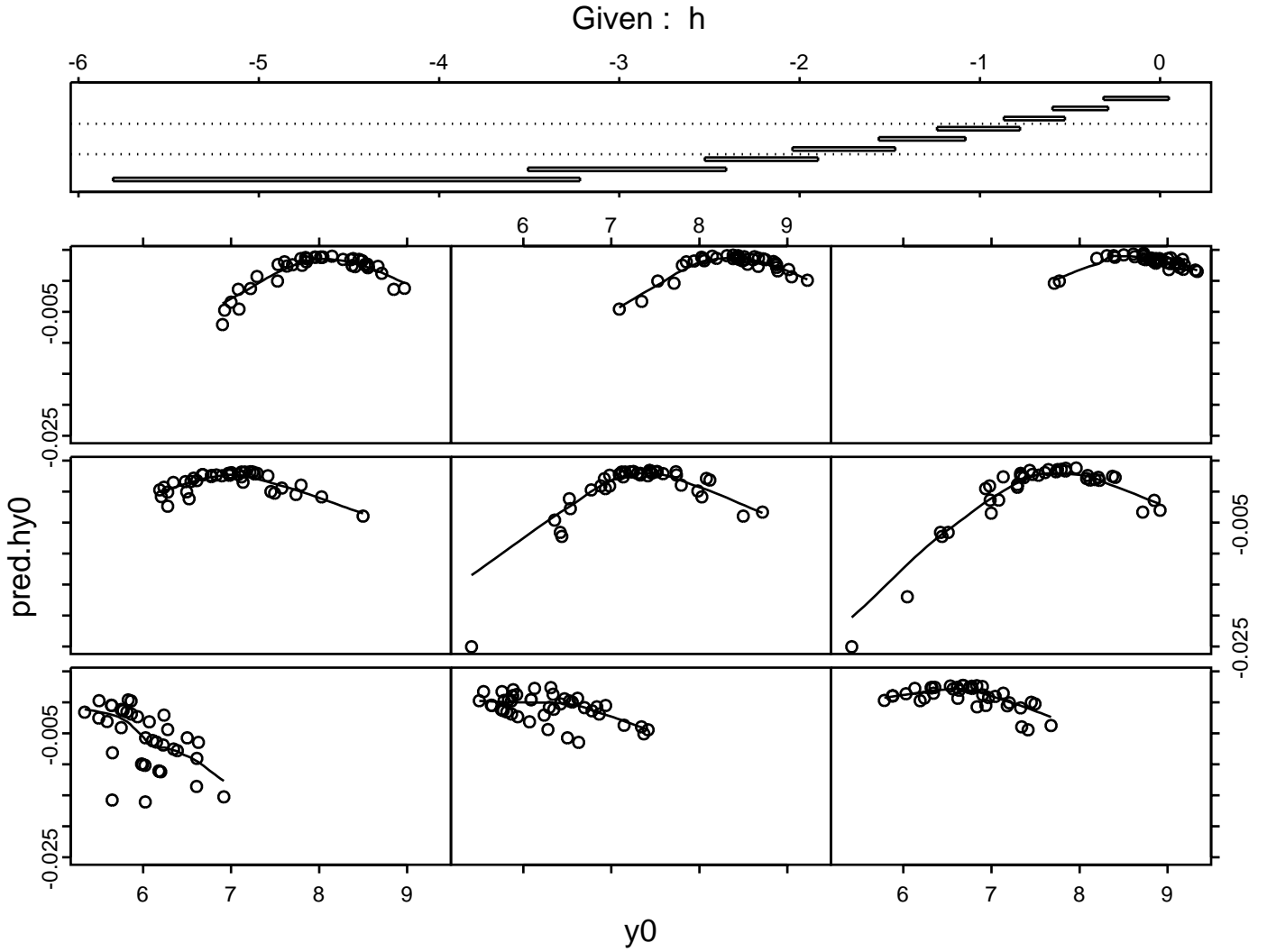
Figure 3a

Figure 3b

Figure 3c

Figure 3d

Figure 4 - coplot of predicted growth against initial income given human capital with scatterplot smoothings
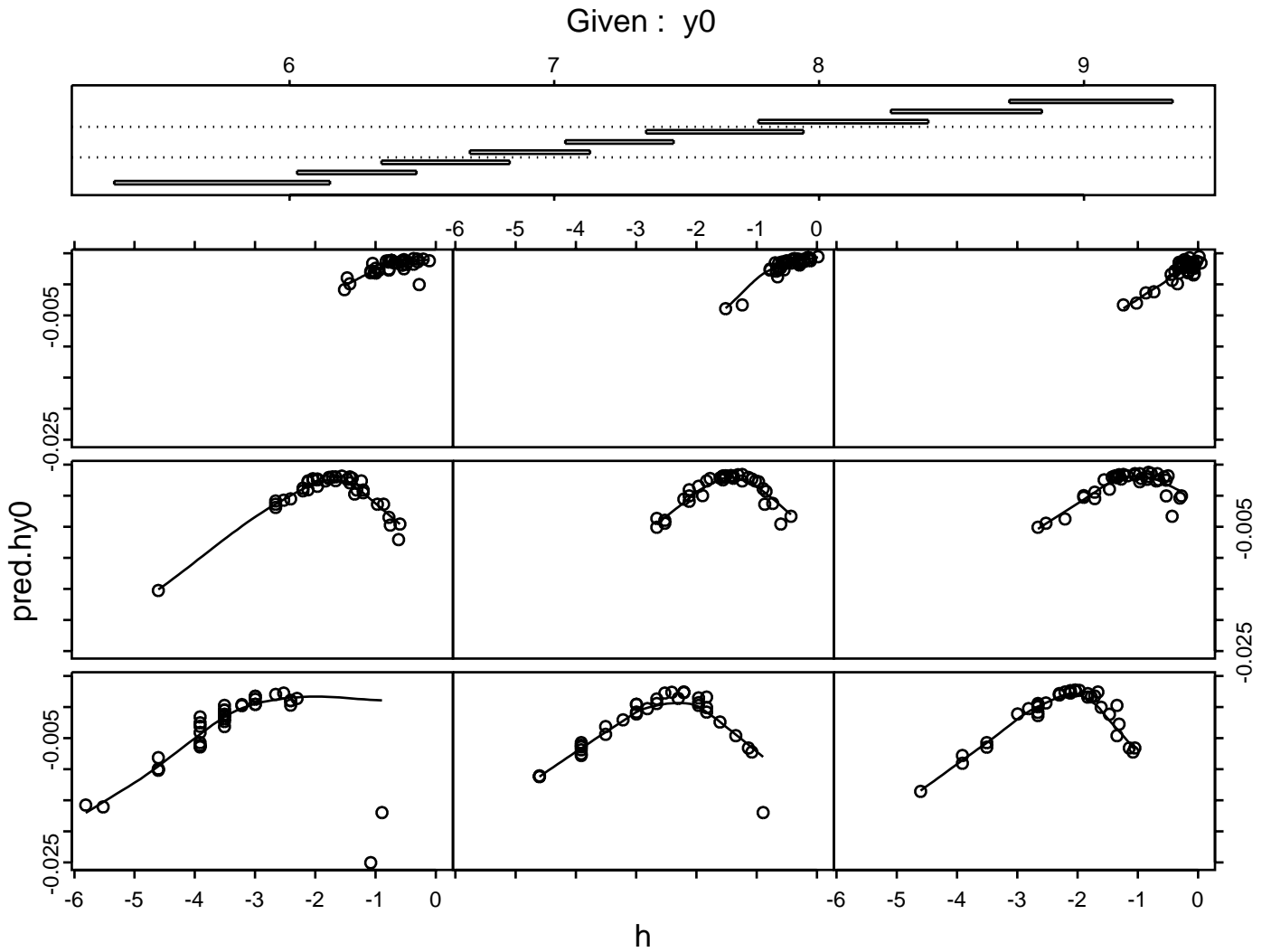
Figure 5 - coplot of predicted growth against human capital given initial income with scatterplot smoothings
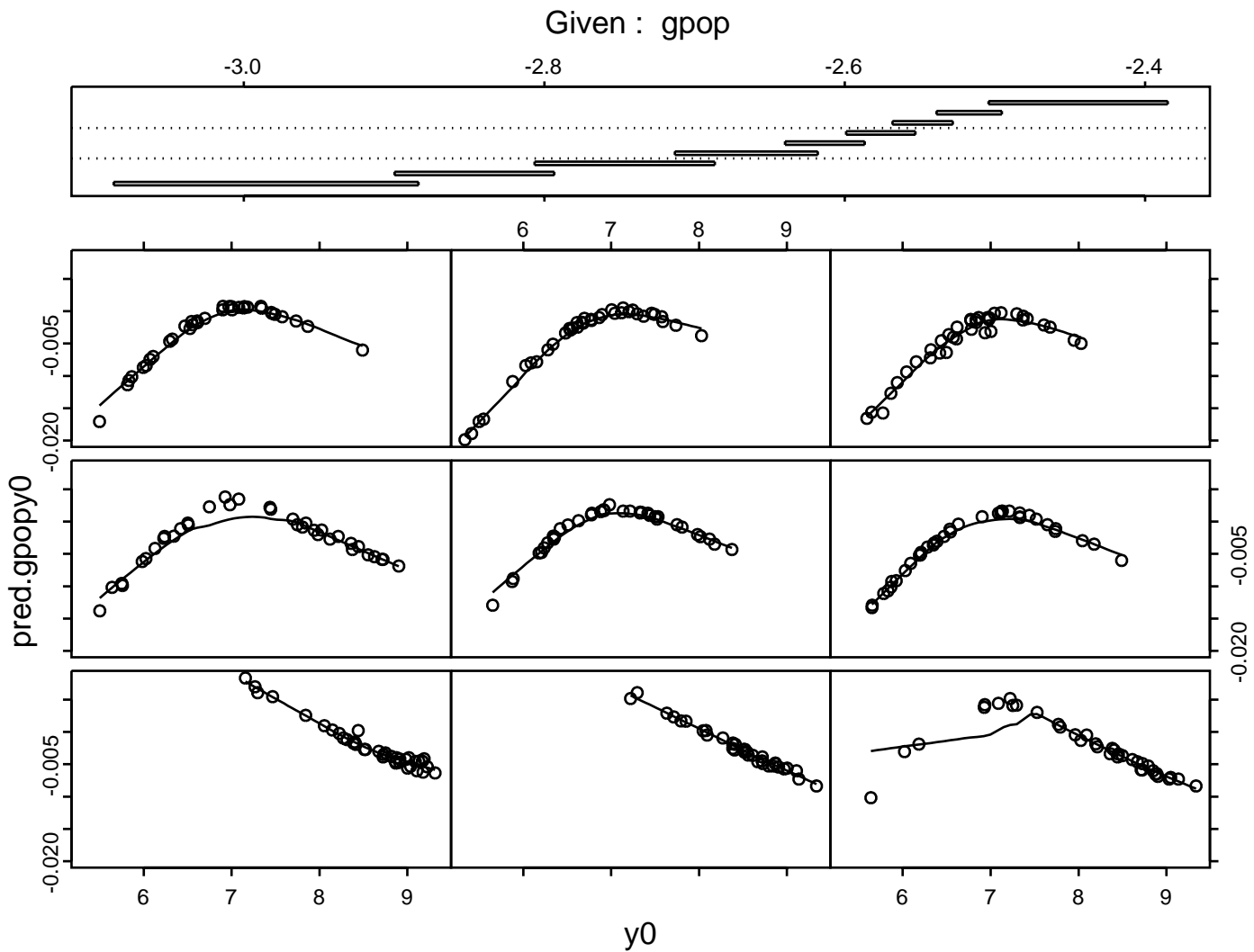
Figure 6 - coplot of predicted growth against initial income given population growth with scatterplot smoothings

SELECTED RECENT PUBLICATIONS

Andreou E. and E. Ghysels, Rolling Volatility Estimators: Some New Theoretical, Simulation and Empirical Results, *Journal of Business and Economic Statistics*, forthcoming 2001.

Andreou E. and A. Spanos, Testing Trend versus Difference Stationarity and Statistical Adequacy, forthcoming *Econometric Reviews*, 2001.

Andreou E., N. Pittis and A. Spanos, Modelling Stock Returns: The Empirical Literature, *Journal of Economic Surveys*, 15, 2, 187-220.

Andreou E., R. Desiano and M. Sensier, The Behaviour of Stock Returns and Interest Rates over the Business Cycle in the US and UK, *Applied Economic Letters*, 8, 233-238, 2001.

Andreou E., D. R. Osborn and M. Sensier, A Comparison of the Statistical Properties of Financial Variables in the USA, UK and Germany over the Business Cycle, *The Manchester School*, 68, 4, 396-418, 2000.

Anil K. Bera and Y. Bilias, Rao´s Score, Neyman´s C (alpha) and Silvey´s LM Tests: An Essay on Historical Developments and Some New Results, *Journal of Statistical Planning and Inference,* 97, 9-44, 2001.

Bertaut C. and M. Haliassos, Precautionary Portfolio Behavior from a Life-Cycle Perspective, *Journal of Economic Dynamics and Control,* 21, 1511-1542, 1997.

Bilias Y., Minggao Gu and Zhiliang Ying, Towards a General Asymptotic Theory for the Cox model with Staggered Entry, *The Annals of Statistics*, 25, 662-682, 1997.

Blundell R., P. Pashardes and G. Weber, What Do We Learn About Consumer Demand Patterns From Micro-Data?, *American Economic Review*, 83, 570-597, 1993.

Bougheas S., P. Demetriades and T. P. Mamouneas, Infrastructure, Specialization and Economic Growth, *Canadian Journal of Economics,* forthcoming.

Caporale W., C. Hassapis and N. Pittis, Unit Roots and Long Run Causality: Investigating the Relationship between Output, Money and Interest Rates, *Economic Modeling*, 15(1), 91-112, January 1998.

Caporale G. and N. Pittis, Efficient estimation of cointegrated vectors and testing for causality in vector autoregressions: A survey of the theoretical literature, *Journal of Economic Surveys*, forthcoming.

Caporale G. and N. Pittis, Unit root testing using covariates: Some theory and evidence, *Oxford Bulletin of Economics and Statistics*, forthcoming.

Caporale G. and N. Pittis, Causality and Forecasting in Incomplete Systems, *Journal of Forecasting*, 16, 6, 425-437, 1997.

Clerides K. S., S. Lach and J.R. Tybout, Is Learning-by-Exporting Important? Micro-Dynamic Evidence from Colombia, Morocco, and Mexico, *Quarterly Journal of Economics* 113(3), 903- 947, August 1998.

Cukierman A., P. Kalaitzidakis, L. Summers and S. Webb, Central Bank Independence, Growth, Investment, and Real Rates", Reprinted in Sylvester Eijffinger (ed), Independent Central Banks and Economic Performance, Edward Elgar, 416-461, 1997.

Dickens R., V. Fry and P. Pashardes, Non-Linearities and Equivalence Scales, *The Economic Journal*, 103, 359-368, 1993.

Demetriades P. and T. P. Mamuneas, Intertemporal Output and Employment Effects of Public Infrastructure Capital: Evidence from 12 OECD Economies, *Economic Journal*, July 2000.

Eicher Th. and P. Kalaitzidakis, The Human Capital Dimension to Foreign Direct Investment: Training, Adverse Selection and Firm Location". In Bjarne Jensen and Kar-yiu Wong (eds), Dynamics,Economic Growth, and International Trade, The University of Michigan Press, 337-364, 1997.

Fry V. and P. Pashardes, Abstention and Aggregation in Consumer Demand, *Oxford Economic Papers*, 46, 502-518, 1994.

Gatsios K., P. Hatzipanayotou and M. S. Michael, International Migration, the Provision of Public Good and Welfare, *Journal of Development Economics*, 60/2, 561-577, 1999.

Guiso, L., M. Haliassos, and T. Jappelli, Household Portfolios: An International Comparison, forthcoming in Guiso, Haliassos, and Jappelli (Eds.), *Household Portfolios*, Cambridge, MA: MIT Press, 2002.

Haliassos M., On Perfect Foresight Models of a Stochastic World, *The Economic Journal*, 104, 477-491, 1994.

Haliassos M. and C. Bertaut, Why Do So Few Hold Stocks?, *The Economic Journal*, 105, 1110- 1129, 1995.

Haliassos M. and C. Hassapis, Non-expected Utility, Saving, and Portfolios, *The Economic Journal*, 111, 69-102, 2001.

Haliassos, M. and A. Michaelides, Portfolio Choice and Liquidity Constraints, *International Economic Review*, forthcoming.

Haliassos, M. and A. Michaelides, Calibration and Computation of Household Portfolio Models, forthcoming in Guiso, Haliassos, and Jappelli (Eds.), *Household Portfolios*, Cambridge, MA: MIT Press, 2002.

Haliassos M. and J. Tobin, The Macroeconomics of Government Finance, reprinted in J. Tobin, *Essays in Economics,* vol. 4, Cambridge: MIT Press, 1996.

Hassapis C., S. Kalyvitis and N. Pittis, Cointegration and Joint Efficiency of International Commodity Markets", *The Quarterly Review of Economics and Finance,* 39, 213-231, 1999.

Hassapis C., N. Pittis and K. Prodromidis, Unit Roots and Granger Causality in the EMS Interest Rates: The German Dominance Hypothesis Revisited, *Journal of International Money and Finance*, 18(1), 47-73, 1999.

Hassapis C., N. Pittis and K. Prodromides, EMS Interest Rates: The German Dominance Hypothesis or Else?" in European Union at the Crossroads: A Critical Analysis of Monetary Union and Enlargement, Aldershot, UK., Chapter 3, 32-54, 1998. Edward Elgar Publishing Limited.

Hatzipanayotou, P. and M.S. Michael, Public Goods, Tax Policies and Unemployment in LDC's, *Southern Economic Journal,* 68/1, 107-119, 2001.

Hatzipanayotou P., and M. S. Michael, General Equilibrium Effects of Import Constraints Under Variable Labor Supply, Public Goods and Income Taxes, *Economica*, 66, 389-401, 1999.

Hatzipanayotou, P. and M. S. Michael, Public Good Production, Nontraded Goods and Trade Restriction, *Southern Economic Journal,* 63, 4, 1100-1107, 1997.

Hatzipanayotou, P. and M. S. Michael, Real Exchange Rate Effects of Fiscal Expansion Under Trade Restrictions, *Canadian Journal of Economics,* 30-1, 42-56, 1997.

Kalaitzidakis P., T. P. Mamuneas and Th. Stengos, A Nonlinear Sensitivity Analysis of Cross-Country Growth Regressions, *Canadian Journal of Economics,* forthcoming.

Kalaitzidakis P., T. P. Mamuneas and Th. Stengos, European Economics: An Analysis Based on Publications in Core Journals, *European Economic Review*, 1999.

Kalaitzidakis P., On-the-job Training Under Firm-Specific Innovations and Worker Heterogeneity, *Industrial Relations,* 36, 371-390, July 1997.

Ludvigson S. and A. Michaelides, Does Buffer Stock Saving Explain the Smoothness and Excess Sensitivity of Consumption?, *American Economic Review*, 631-647, June 2001.

Lyssiotou P., P. Pashardes and Th. Stengos, Age Effects on Consumer Demand: An Additive Partially Linear Regression Model, *Canadian Journal of Economics*, forthcoming 2002.

Lyssiotou P., Dynamic Analysis of British Demand for Tourism Abroad, *Empirical Economics*, 15, 421-436, 2000.

Lyssiotou P., P. Pashardes and Th. Stengos, Testing the Rank of Engel Curves with Endogenous Expenditure, *Economics Letters,* 64, 61-65, 1999.

Lyssiotou P., P. Pashardes and Th. Stengos, Preference Heterogeneity and the Rank of Demand Systems, *Journal of Business and Economic Statistics*, 17 (2), 248-252, April 1999.

Lyssiotou P., Comparison of Alternative Tax and Transfer Treatment of Children using Adult Equivalence Scales, *Review of Income and Wealth*, 43 (1), 105-117, March 1997.

Mamuneas, Theofanis P., Spillovers from Publicly – Financed R&D Capital in High-Tech Industries, *International Journal of Industrial Organization,* 17(2), 215-239, 1999.

Mamuneas, Theofanis P. and M. I. Nadiri, R&D Tax Incentives and Manufacturing-Sector R&D Expenditures, in *Borderline Case: International Tax Policy, Corporate Research and Development, and Investmen*t, James Poterba (ed.), National Academy Press, Washington D.C., 1997. Reprinted in *Chemtec*h, 28(9), 1998.

Mamuneas, Theofanis P. and M. I. Nadiri, Public R&D Policies and Cost Behavior of the US Manufacturing Industries, *Journal of Public Economics*, 63, 57-81, 1996.

Michael S. Michael and Panos Hatzipanayotou, Welfare Effects of Migration in Societies with Indirect Taxes, Income Transfers and Public Good Provision, *Journal of Development Economics*, 64, 1-24, 2001.

Michaelides, A. and S. Ng, Estimating the Rational Expectations Model of Speculative Storage: A Monte Carlo Comparison of three Simulation Estimators, *Journal of Econometrics,* 96(2), 231-266, June 1997.

Pashardes P., Equivalence Scales in a Rank-3 Demand System, *Journal of Public Economic*s, 58, 143-158, 1995.

Pashardes P., Bias in Estimating Equivalence Scales from Grouped Data, *Journal of Income Distributio*n, Special Issue: Symposium on Equivalence Scales, 4, 253-264,1995.

Pashardes P., Bias in Estimation of the Almost Ideal Demand System with the Stone Index Approximation, *Economic Journa*l, 103, 908-916, 1993.

Spanos A., Revisiting Date Mining: ´Hunting´ With or Without a License, *Journal of Methodology,* July 2000.

Spanos A., On Normality and the Linear Regression Model, *Econometric Reviews,* 14,195-203, 1995.

Spanos A., On Theory Testing in Econometrics: Modeling with nonexperimental Data, *Journal of Econometrics,* 67, 189-226, 1995.

Spanos A., On Modeling Heteroscedasticity: The Student's t and Elliptical Linear Regression Models, *Econometric Theor*y, 10, 286-315, 1994.

Zacharias E. and S. R. Williams, Ex Post Efficiency in the Buyer's Bid Double Auction when Demand Can be Arbitrarily Larger than Supply, *Journal of Economic Theory* 97, 175-190, 2001.