Optimal Tests for Reduced Rank Time Variation in Regression Coefficients and Level Variation in the Multivariate Local Level Model

November 2003
(Revised October 16, 2004)

Piotr Eliasz
Department of Economics, Princeton University

James H. Stock
Department of Economics, Harvard University
and the National Bureau of Economic Research

and

Mark W. Watson*
Woodrow Wilson School and Department of Economics, Princeton University
and the National Bureau of Economic Research

* This research was funded in by NSF grant SBR-0214131. We thank Zongwu Cai and Ulrich Müller for useful discussions.

Abstract

This paper constructs tests for martingale time variation in regression coefficients in the regression model $y_t = x_t'\beta_t + u_t$, where β_t is $k\times1$, and Σ_β is the covariance matrix of $\Delta\beta_t$. Under the null there is no time variation, so H_o: $\Sigma_\beta = \mathbf{0}$; under the alternative there is time variation in r linear combinations of the coefficients, so H_a: rank$(\Sigma_\beta) = r$, where r may be less than k. The Gaussian point optimal invariant test for this reduced rank testing problem is derived, and the test's asymptotic behavior is studied under local alternatives. The paper also considers the analogous testing problem in the multivariate local level model $Z_t = \mu_t + a_t$, where Z_t is a $k\times1$ vector, μ_t is a level process that is constant under the null but is subject to reduced rank martingale variation under the alternative, and a_t is an I(0) process. The test is used to investigate possible common trend variation in the growth rate of per-capita GDP in France, Germany and Italy.

Keywords: TVP tests, multivariate local level model, POI tests

JEL Numbers: C12, C22, C32

1. Introduction

A long-standing problem in econometrics involves testing for stability of regression coefficients in the linear regression model. Using standard notation, the model is

$$y_t = x_t'\beta_t + u_t \quad (1.1)$$

where y_t is a scalar and x_t is a $k\times 1$ vector. Under the null hypothesis the regression coefficients are stable, while under the alternative they are time varying. When k is large, standard tests for time variation have low power because they look for time variation in k different dimensions. However, in many empirical applications it is plausible to assume that time variation in the coefficients will be restricted to a relatively small number of linear combinations of the regression coefficients. For example, it might be assumed that any time variation is concentrated in the linear combinations $R\beta_t$ where R is an $r\times k$ matrix. When R is known, the regression model can be transformed to isolate the coefficients $R\beta_t$, and the problem involves testing whether a subset of the regression coefficients are unstable (see Leybourne (1993)).

In many applications a researcher might not know the value of R, and test power will deteriorate if the wrong value is used. This concern leads us to consider the testing problem when R is unknown. That is, we suppose that r linear combinations of the regression coefficients are unstable under the alternative hypothesis, but that these linear combinations are unknown. In our leading case $r = 1$, so there is only one dimension of time variation in the regression coefficients. We are concerned with three related questions. First, what are the power gains that can be attained using this rank information relative to tests that look for time variation in all of the regression coefficients? Second, what are the power losses associated with using only the rank information relative to tests that use the value of R? Finally, what are the power losses from using a pre-specified but incorrect value of R?

We carry out the analysis using an otherwise standard framework. We assume that $\Delta\beta_t$ is an I(0) process with covariance matrix $\Sigma_{\Delta\beta}$. Under the null hypothesis $\Sigma_{\Delta\beta} = 0$, while under the alternative $\Sigma_{\Delta\beta} \neq 0$ with rank($\Sigma_{\Delta\beta}$) = r. As usual, we consider tests that are invariant to transformations $y_t \rightarrow y_t + x_t'b$. However, to capture the notion that R is unknown, we also restrict attention to tests that are invariant to transformations $x_t \rightarrow Ax_t$. As in Shively (1988), Stock and Watson (1998), and Elliott and Muller (2002) we consider versions of Gaussian point optimal invariant tests.

As it turns out, a closely related problem involves testing whether a $k\times1$ vector process Z_t is I(0) against the alternative that is I(1). We write this model as

$$Z_t = \mu_t + a_t \tag{1.2}$$

where a_t is I(0) and μ_t is I(1). This is a version of what Harvey (1989) calls the "local level model," because μ_t represents the local level of the process. If $\Sigma_{\Delta\mu} = 0$, then μ_t is constant and Z_t is I(0); when $\Sigma_{\Delta\mu} \neq 0$, then Z_t is I(1). In many applications, shifts in μ_t are a function of a small number of factors or common trends, so that the elements of Z_t are cointegrated, $\Sigma_{\Delta\mu}$ has reduced rank, and (1.2) is a reduced rank local level model. This leads us to consider testing the null that $\Sigma_{\Delta\mu} = 0$ against the alternative that $\Sigma_{\Delta\mu} \neq 0$, but with rank($\Sigma_{\Delta\mu}$) = r.

This testing problem is also carried out in an otherwise standard framework. Stock (1994) surveys the large literature concerned with testing $\Sigma_{\Delta\mu} = 0$ in (1.2) when k

=1; we utilize multivariate versions of the Gaussian point optimal invariant tests derived from King (1980) that have been used for the univariate testing problem. These tests are invariant to transformations of the form $Z_t \rightarrow Z_t + b$. We further restrict the tests so that they are invariant to transformations of the form $Z_t \rightarrow AZ_t$ to capture the notion that the I(1) linear combinations of Z_t are unknown. Jansson (2002) considers a multivariate problem closely related to ours. He supposes that $r = 1$, but that the I(1) linear combination of Z_t is known. Our analysis can then be viewed as an extension to the case of an unknown linear combination.

The paper is organized as follows. Section 2 considers the reduced rank multivariate local level model (1.2), and presents exact results using a benchmark Gaussian version of the model, and then extends the results to more general stochastic processes using asymptotic approximations. Section 3 shows the asymptotic equivalence of the testing problems for the regression model (1.1) and the multivariate local level model (1.2). This implies that the testing results derived in section 2 carry over to regression model. Section 4 presents asymptotic power results and answers to the questions posed above about the power gains and losses that result from using rank restrictions. Critical values for the test statistics are also tabulated in this section. Section 5 contains an empirical application that tests for common I(1) time variation in the growth rates of GDP for several European countries. Concluding comments are offered in the final section.

2. The Reduced Rank Multivariate Local Level Model

Let Z denote a $k\times 1$ vector of time series. In the multivariate local level model, deviations of Z_t from the local-level process μ_t follow a zero-mean stationary process. The process μ_t evolves smoothly; we characterize μ_t as an I(1) process. In many applications it is natural to think of the $k\times 1$ vector μ_t as evolving in response to a reduced number of common factors, so that $(1\text{-L})\mu_t = \Lambda(1\text{-L})f_t$, where f_t is an $r\times 1$ vector of I(1) variables with $r \leq k$, and Λ is a $k\times r$ matrix of factor loadings. As discussed in Stock and Watson (1988), the elements of f_t can be thought of as common trends leading to low-frequency variability in Z_t, and when $r < k$, Z_t is a cointegrated processes.

Using this common trends formulation, we write the multivariate local level model as

$$Z_t = \mu_t + a_t \tag{2.1}$$

$$\mu_t = \mu_0 + \Lambda f_t. \tag{2.2}$$

$$a_t = \theta_a(\text{L})\varepsilon_t \tag{2.3}$$

$$\Delta f_t = \theta_f(\text{L})\eta_t \tag{2.4}$$

$e_t = (\varepsilon_t'\ \ \eta_t')'$ is a martingale difference sequence and $\theta_a(1)$ and $\theta_f(1)$ are finite. When $r < k$, we refer to this as the reduced rank multivariate local level model.

The model is over-parameterized as written in (2.1)-(2.4), and we use the following normalizations:

(N.1) $\quad$ $\text{E}(e_te_t') = I_{k+r}$

(N.2) $\theta_a(0)$ is lower triangular

(N.3) (i) $\Lambda = \theta_a(1)P$

(ii) $P'P = I_r$

(iii) $\theta_f(1) = \text{diag}(\gamma_i)$, $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_r \geq 0$.

The normalizations (N.1) and (N.2) imply that (2.3) is written in Wold causal form with innovation standard deviations on the diagonal of $\theta_a(0)$. Because of the factor structure in (2.2), Λ and second moments of Δf_t are not separately identified, and (N.3) provides a normalization of the factors that will be convenient for the testing problem considered below. In this normalization the γ_i's are the square roots of the eigenvalues of $\theta_a(1)^{-1}\Omega\theta_a(1)$, where Ω is the spectral density matrix of $\Delta\mu_t$ at frequency zero; the columns of P are the corresponding orthonormal eigenvectors.

We consider testing the null hypothesis that μ_t is constant. Under the alternative μ_t evolves as a reduced-rank process. We simplify the testing problem by assuming that $\gamma_i = \gamma$, for all i, so that the null and alternative can be written in terms of the single parameter γ as H_0: $\gamma = 0$ versus H_1: $\gamma > 0$. This simplification is without loss of generality when $r = 1$, the single common trend model that is the leading case used in empirical analysis. When $r > 1$, the simplication means that the tests described below will be optimal when the normalized factors have a common variance.

We begin by developing tests that are optimal in a Gaussian version of the model. We will then show that the size and local power properties of these tests continue to hold under more general distributional and serial correlation assumptions when the sample size is large. The Gaussian model is characterized by (2.1)-(2.4), (N.1)-(N.3) and the following additional assumptions:

(G.1) $e_t \sim i.i.d.$ N(0, I)

(G.2) $\theta_a(\mathrm{L}) = \theta_a(0)$, where $\theta_a(0)$ is non-singular

(G.3) $\theta_f(\mathrm{L}) = \theta_f(0)$

These assumptions imply that $\{a_t, \Delta f_t\}$ is a sequence of i.i.d. normal random variables.

There are three sets of nuisance parameters that complicate the testing problem: $\theta_a(0)$, P and μ_0. As we show below, $\theta_a(0)$ is not a major problem because it can be consistently estimated under the null and local alternatives, and the resulting sampling error has no effect on sampling distribution of the optimal test statistic in large samples. Thus, we will assume that this parameter is known when developing the optimal Gaussian test. The parameters P and μ_0 are more problematic. Sampling error in μ_0 affects the properties of the test even in large samples; P is unidentified under the null and cannot be consistently estimated under local alternatives. We eliminate μ_0 and P from the testing problem by restricting attention to tests that are invariant to transformations of the form $Z_t \rightarrow AZ_t + b$, where A is an arbitrary orthonormal matrix and b is an arbitrary constant.

Theorem 1 provides an expression for the point optimal invariant test. The theorem uses the following notation: Z denotes a $T\times k$ matrix with t'th row given by Z_t', $V_\gamma = \gamma^2 FF' + I_T$, where F is a lower triangular matrix of 1's, l denotes a $T\times 1$ vector of 1's, and $M_l = I - l(l'l)^{-1}l'$.

Theorem 1: In the Gaussian model, the best invariant test for H_0: $\gamma = 0$ versus H_1: $\gamma = \bar{\gamma}$ rejects the null hypothesis for large values of $B_{\bar{\gamma}}$, where

$$B_{\bar{\gamma}} = \int \exp\{-\tfrac{1}{2} trace[S' D_{\bar{\gamma}} S]\} dU(S), \tag{2.5}$$

$D_{\bar{\gamma}} = \theta_a(0)^{-1}\{Z'[V_{\bar{\gamma}}^{-1} - V_{\bar{\gamma}}^{-1} l(l' V_{\bar{\gamma}}^{-1} l)^{-1} l' V_{\bar{\gamma}}^{-1}]Z - Z' M_l Z\}\theta_a(0)'^{-1}$, S is a $r\times k$ matrix, and $U(S)$ is a weighting distribution that puts uniform weight on the unit sphere, that is $dU(S)$ is constant on $S'S = I_r$ and zero elsewhere.

The proof of this theorem is straightforward. Notice that any test that is invariant to transformations of the form $Z_t \rightarrow AZ_t$ must have constant power for all values of P with $P'P = I_r$. Thus, invariant tests have the property that average power on $P'P = I_r$ coincides with power for any particular value of P. This means that an invariant test with best average power on $P'P = I_r$ is a best invariant test. Following the analysis in King (1980) it is straightforward to derive optimal tests that are invariant to transformations of the form $Z_t \rightarrow Z_t + b$ for known values of P, say $P = S$. The likelihood ratio based on the maximal invariants is given by $\exp\{-\tfrac{1}{2} trace[S' D_{\bar{\gamma}} S]\}$. Inspection of (2.5) then reveals that a test that rejects for large values of $B_{\bar{\gamma}}$ will have the greatest average power, where the average is constructed using U as the weighting function (Andrews and Ploberger (1994)). Because U is uniform, $B_{\bar{\gamma}}$ is invariant to transformations of the form $Z_t \rightarrow AZ_t$. Thus, a test that rejects the null for large values of $B_{\bar{\gamma}}$ is the best invariant test.

The test statistic $B_{\bar{\gamma}}$ has a relatively simple large-sample distribution under the null hypothesis and under local alternatives of the form $\gamma = g/T$, where g is a constant. These distributions hold under assumptions more general (G.1)-(G.3), and we now describe one such generalization.

(F.1) (i) e_t is a stationary and ergodic martingale difference sequence.

(ii) $E(e_t e_t' \mid e_{t-1}, e_{t-2}, \ldots) = I_{k+r}.$

(iii) $\max_{i,j,k,l} E(e_{it}e_{jt}e_{kt}e_{lt}) < \kappa$ for all t.

(F.2) θ_a(L) is 1-summable and $\theta_a(1)$ has full rank.

(F.3) (i) $\gamma = \gamma_T = T^{-1}g$, where g will be held fixed as T grows large.

(ii) $\theta_f(\mathrm{L}) = \theta_{f,T}(\mathrm{L}) = \gamma_T H(\mathrm{L})$, where H(L) is fixed as T grows large, $H(1) = I_r$, and H(L) is 1-summable.

Assumption (F.1) implies that $T^{-1/2}\sum_{t=1}^{[sT]} e_t \Rightarrow W(s)$ a Wiener process and replaces the Gaussian assumption (G.1). Assumption (F.2) allows a_t to be serially correlated. Assumption (F.3i) describes the asymptotic nesting for the local power analysis. Assumption (F.3ii) defines the factorization of θ_f(L) and replaces (G.3) with a limited memory assumption.

Theorem 2 characterizes the asymptotic behavior of the appropriately modified test statistics under (F.1)-(F.3). The theorem uses the following notation: $\overline{P} = [I_r\ 0_{r\times(k-r)}]'$, $W(s)$ is a k+r dimensional Wiener process partitioned as $W(s) = [W_1(s)'\ W_2(s)]'$ where $W_1(s)$ is k×1 and $W_2(s)$ is r×1, $X(s) = g\overline{P}\int_0^s W_2(\tau)d\tau + W_1(s)$, $Y(s) = X(s) - sX(1)$, $q(s) = \int_0^s e^{-\overline{g}(s-\tau)}dY(\tau)$, and $J = \int_0^1 e^{-s\overline{g}}dY(s) - \overline{g}\int_0^1 e^{-s\overline{g}}q(s)ds$.

Theorem 2:

In the model (2.1) -(2.4), suppose that $\hat{\theta}_a(1)$ is a consistent estimator of $\theta_a(1)$, and (F.1)-(F.3) are satisfied. Let $\hat{D}_{\overline{\gamma}} = \hat{\theta}_a(1)^{-1}\{Z'[V_{\overline{\gamma}}^{-1} - V_{\overline{\gamma}}^{-1}l(l'V_{\overline{\gamma}}^{-1}l)^{-1}l'V_{\overline{\gamma}}^{-1}]Z - Z'M_l Z\}\hat{\theta}_a(1)'^{-1}$ and $\hat{B}_{\overline{\gamma}} = \int \exp\{-\tfrac{1}{2}trace[S'\hat{D}_{\overline{\gamma}}S]\}dU(S)$. Then $\hat{D}_{\overline{g}/T} \Rightarrow F_{\overline{g}}$ and

$$\hat{B}_{\bar{g}/T} \Rightarrow B = \int \exp\{-½ trace[S' F_{\bar{g}} S]\} dU(S),$$

where $F_{\bar{g}} = -\bar{g}\,[q(1)q(1)' + \bar{g}\int_0^1 q(s)q(s)'ds + 2(1-e^{-2\bar{g}})^{-1}JJ']$

Proof: See Appendix.

The theorem requires a consistent estimator of $\theta_a(1)$, and these are readily obtained. Under (F.1) - (F.3), $T^{-1}\sum(Z_t-\bar{Z})(Z_{t-k}-\bar{Z})' = T^{-1}\sum(a_t-\bar{a})(a_{t-k}-\bar{a})' + O_p(T^{-1})$, so that consistent estimators of the long-run covariance matrix of a can be constructed from standard kernel-based or parametric estimators with Z_t used in place of a_t.

3. The Linear Regression Model with Reduced Rank Time Varying Coefficients

As discussed in the introduction, the regression model with stationary regressors and time varying regression coefficients behaves much like the local level model. In this section we show how the testing results for the reduced rank multivariate local level model can be applied to the reduced rank time varying coefficient regression model. We write the regression model as

$$y_t = x_t'\beta_t + u_t \tag{3.1}$$

$$\beta_t = \beta_0 + \Gamma f_t \tag{3.2}$$

$$\Delta f_t = \theta_f(\mathrm{L})\eta_t \tag{3.3}$$

where y_t is a scalar, x_t is $k\times 1$, β_t is a vector of potentially time varying regression coefficients, and u_t is the regression error. From (3.2), the $k\times 1$ vector β_t evolves as a function of the $r\times 1$ vector f_t, where $r \le k$, so that $k-r$ distinct linear combinations of the regression coefficients are stable. The factors follow the process (3.3). The regression errors u_t may be serially correlated; we discuss this below.

To see the relation between this model and the multivariate local level of the last section, let $\tilde{Z}_t = x_t(y_t - x_t'\beta_0)$, $\mu_t = \Sigma_{xx}(\beta_t - \beta_0)$, $a_t = x_t u_t$, and $w_t = (x_t x_t' - \Sigma_{xx})(\beta_t - \beta_0)$, where $\Sigma_{xx} = E(x_t x_t')$. Then $\tilde{Z}_t = \mu_t + a_t + w_t$. If the regressors are well-behaved, w_t will be negligible in large samples, so that $\tilde{Z}_t \approx \mu_t + a_t$, and any time variation in β_t will be reflected as time variation in μ_t, which can be detected as in the last section.

To keep the analysis of the regression model parallel to the analysis used in the last section time, variation in β_t is parameterized as in (N.3), but with (N.3i) replaced by

(N.3′) $$\Gamma = \Sigma_{xx}\Omega_{aa}^{1/2}P,$$

where $\Omega_{aa}^{1/2}$ is the long-run covariance matrix of $a_t = x_t u_t$. As in the last section we consider the null and alternative hypotheses H_0: $\gamma = 0$ versus H_1: $\gamma = \bar{\gamma}$, where $\theta_f(1) = \gamma I_r$. Invariance in the regression testing problem involves transformations $y_t \rightarrow y_t + x_t'b$ and $x_t \rightarrow Ax_t$, where b is an arbitrary vector of constants and A is an orthonormal matrix.

As the heuristic suggests, tests in the regression model will be well-behaved in large samples when $w_t = (x_t x_t' - \Sigma_{xx})(\beta_t - \beta_0)$ is asymptotically negligible, which happens when the regressors are sufficiently well-behaved. We show this result using an argument from Stock and Watson (1998), so it is convenient to use their assumption about the regressors. The assumption uses the following notation. For a stationary vector process b_t, let $c_{i_1 \ldots i_n}(r_1,\ldots,r_n)$ denote the nth joint cummulant of $b_{i_1 t_1},\ldots,b_{i_n t_n}$, where $r_j = t_j - t_n$, $j = 1, \ldots, n-1$, and let $C(r_1, \ldots, r_{n-1}) = \sup_{i_1,\ldots,i_n} c_{i_1 \ldots i_n}(r_1,\ldots,r_n)$. The assumption is

(R.1) (i) X_t is stationary with eighth-order cummulants that satisfy

$$\sum_{r_1,\ldots,r_7}^{\infty} | C(r_1,\ldots r_7) | < \infty .$$

(ii) $\{X_t\}$ is independent of $\{\eta_t\}$.

Part (i) of the assumption restricts attention to stationary regressors with eight moments and limited temporal dependence. Importantly, it rules out trending and integrated regressors. From part (ii), the regressors are assumed to be independent of time variation in the regression coefficients.

Optimality results analogous to those presented in Theorem 1 are derived under the following Gaussian assumptions:

(G.4) (i) $u_t = \sigma_u \varepsilon_t$ and $e_t = (\varepsilon_t \ \eta_t')'$ is i.i.d N(0, I_{r+1}).

(ii) $\{x_t\}$ and $\{u_t\}$ are independent.

(iii) $\theta_f(\mathrm{L}) = \theta_f(0) = T^{-1}gI_r$

Parts (i) and (iii) imply that u_t and $\Delta\beta_t$ are i.i.d. normal random variables as in the Gaussian model of the last section; part (ii) makes the regressors exogenous. Because the results for the regression model are asymptotic, part (iii) includes the same asymptotic nesting used in Theorem 2. As in that theorem, the asymptotic size and local power of the Gaussian test can be obtained under more general assumptions. Letting $a_t = x_t u_t$ and $\mu_t = \Sigma_{xx}\beta_t$, the assumptions summarized (F.1)-(F.3) of the last section will suffice.

Theorem 3 summarizes the results for the regression model.

Theorem 3

Consider the regression model (3.1)-(3.3), with regressors satisfying (R.1) and let $Z_t = x_t \hat{u}_t$, where $\hat{u}_t$ is the OLS residual from the regression of y_t onto x_t.

(a) If (G.4) holds, then the best invariant asymptotic test can be constructed as in Theorem 1 with $\theta_a(0) = \sigma_u \Sigma_{XX}^{1/2}$.

(b) If $a_t = x_t u_t$ and $\mu_t = \Sigma_{xx}\beta_t$, satisfy (F.1)-(F.3) then theorem 2 holds for the regression model.

The proof of part (a) parallels the proof to Theorem 1 after replacing King's (1980) result on the likelihood ratio for the maximal invariants for known P with an asymptotic approximation from Elliott and Müller (2002). Part (b) follows from a straightforward calculation. Details are provided in the appendix.

4. Local Power Comparisons

This section has four purposes. First, it compares the performance of the point-optimal invariant tests derived in Sections 2 and 3 to other invariant tests. Second, it verifies that appropriately chosen point-optimal tests work well for a wide range of values of γ under the alternative. Third, it compares the performance of the point-optimal invariant tests to a point optimal non-invariant test that uses a pre-specified, but potentially incorrect value of Λ. Finally, it tabulates large sample critical value for an easy-to-compute test statistic with performance that is indistinguishable from the optimal invariant test.

In addition to the point-optimal invariant test, we consider two alternative invariant tests. We introduce these tests using the notation introduced for the Gaussian model studied in Theorem 1; in empirical applications they would be computed using the modifications described in Theorem 2 and in section 3. The first statistic is

$\xi_{sup} = \sup_{S'S=I} -S' D_{\bar{\gamma}} S = \sum_{i=1}^{r} eig_i(-D_{\bar{\gamma}})$ where $eig_i(-D_{\bar{\gamma}})$ is the i'th eigenvalue of $-D_{\bar{\gamma}}$ ordered from largest to smallest. This statistic uses the information that $r < k$, but in a way that may not be optimal. The second statistic that we consider is $\xi_{trace} = trace(-D_{\bar{\gamma}})$. This is the optimal test when $r = k$ and therefore it does not use the information that Λ has reduced rank. The statistics are invariant because they depend on $D_{\bar{\gamma}}$ only through its eigenvalues.

Figure 1 shows the asymptotic local power envelope based on B_{γ}, the optimal test from Theorem 1 constructed using the true value of $\gamma = g/T$, for a test with size 5%. Results are shown for $r = 1$ and $k = 1$, 5, 10 and 20. Also shown are the corresponding powers of the ξ_{sup} and ξ_{trace} tests constructed using the same values of γ. When $k = 1$, the three tests coincide, so that only one function appears in panel (a) of the figure. For other values of k, the tests differ, so that three power functions are plotted in the remaining panels. Two results stand out from the figure. First, the power function for ξ_{sup} essentially coincides with the power envelope. This is an important result because ξ_{sup} is much easier to compute than the point optimal test because it does not require the numerical integration of (2.5). Second, the optimal tests and ξ_{sup} outperform ξ_{trace}, but the gains are not large. At power 50%, the Pitman efficiency of ξ_{trace} relative to the optimal test is 96% ($k = 5$), 93% ($k = 10$), and 92% ($k = 20$). Thus the optimal reduced rank test has only moderately more power over the corresponding test that does not use the information about rank. We have also computed analogous power functions for $r > 1$. As in the $r = 1$ case, ξ_{sup} is essentially optimal. The power gains of the optimal test relative to ξ_{trace} gets smaller as r grows. For example, when $r = 2$, the (50% power) Pitman efficiency of the trace test is greater than 98% for all values of $k \leq 20$.

Figure 2 shows the asymptotic power envelope as a function of g along with the power function of the point optimal, ξ_{sup}, and ξ_{trace} tests constructed using $\bar{\gamma} = 10/T$. This value of $\bar{\gamma}$ is chosen because it corresponds with approximately 50% power for k = 5. Results are shown for $r = 1$ and for k = 5, 10 and 20. Apparently there is little loss of power using these point optimal tests, at least for the range of values of g considered in the figures.

Asymptotic critical values of for ξ_{sup}, and ξ_{trace} using $\bar{\gamma} = 10/T$ are summarized in Table 1. Results are shown for $r = 1$ and $0 \leq k \leq 20$. (Because ξ_{sup} is easier to compute than $B_{\bar{\gamma}}$ and shares its power properties, critical values for $B_{\bar{\gamma}}$ are not reported.)

The results in figure 1 show that power could be improved substantially if P, the matrix determining the factor loading matrix, was known. The optimal invariant test uses the information that P has rank r, but does not use any information about the likely value of P under the alternative. For example, if it was known that $P = \bar{S}$, then the optimal test would be based on the statistic $\xi_{\bar{S}} = -\bar{S}' D_{\bar{\gamma}} \bar{S}$. When $r = 1$, the asymptotic power envelope would be given by panel (a) of figure 1, regardless of the value of k. This suggests that $\xi_{\bar{S}}$ might work well when $\bar{S}$ is close to but not necessarily equal to P. This is indeed the case. Let $\beta(g, k)$ denote the asymptotic power of the point optimal test as a function of g and k. For $r = 1$, it is straightforward to show that the power function for $\xi_{\bar{S}}$ is given by $\beta(g\,|\bar{S}'P|, 1)$. This is the power of the point optimal test for $k = 1$, but with g scaled by the factor $|\bar{S}'P|$. Because $\bar{S}$ and P are orthonormal, $|\bar{S}'P| \leq 1$. Figure

3 compares the power functions $\beta(g\,|\bar{S}\,'P|, 1)$ and $\beta(g, k)$ for k=5 and k=10, and for $|\bar{S}\,'P| = 0.5, 0.70, 0.90$ and 1.0. Roughly speaking, when $k = 5$, $\xi_{\bar{S}}$ dominates the point-optimal invariant test when $|\bar{S}\,'P| > 0.70$. When $k = 10$, $\xi_{\bar{S}}$ dominates the invariant test for somewhat smaller values of $|\bar{S}\,'P|$.

5. Empirical Results

Figure 4 plots quarterly values of per capita real GDP growth rates for France, Germany and Italy from 1960-2002. (See the Data Appendix for a description of the data.) The figure suggests a common decline in growth rates over the sample period, and this is reinforced by the estimated value of μ_t also plotted in each figure. (These estimates are computed using a Kalman smoother from a models that is described below.) Table 2 presents mean growth rates over various subsamples and these too suggest a common decline in the average growth rates. This informal analysis suggests that the data are consistent with the multivariate local level model (2.1) - (2.4) with time varying level process μ_t driven by a single common factor that experienced a persistent decline over 1960-2002.

Table 3 presents tests of the null hypothesis of no time variation in μ_t. Panel (a) shows p-values for the point optimal invariant test B and the corresponding ξ_{sup} test for $r = 1$; also shown is the p-value for the ξ_{trace} test statistic. These statistics were computed using a parametric VAR(4) estimator for $\theta_a(1)$ and $\hat{D}_{\bar{\gamma}}$ was computed using $\bar{\gamma} = 10/T$. Panel (b) shows the results for tests constructed using pre-specified values of the factor loadings. For this purpose the stochastic component of the local level process μ_t is expressed as $\bar{\Lambda} f_t$, where the value of $\bar{\Lambda}$ is given in the shown in the first column of the table. The test statistic is then computed as $P'\hat{D}_{\bar{\gamma}} P$ where $P = \hat{\theta}_a(1)^{-1}\bar{\Lambda}\omega$, where $\omega =$ $(\bar{\Lambda}'\hat{\theta}_a(1)^{-1}{}'\hat{\theta}_a(1)^{-1}\bar{\Lambda})^{1/2}$ (see the normalization (N.3)). The statistic $P'\hat{D}_{\bar{\gamma}} P$ can also be used to construct a point estimate and confidence interval for the standard deviation of Δf_t using the procedure outlined in Stock and Watson (1998); these are shown in the last two columns of panel (b).

The results in panel (a) show some evidence of time variation. The p-values for the point optimal test B and ξ_{sup} are roughly 10% and the p-value for ξ_{trace} is 16%. The first row of panel (b) shows a more significant rejection (p-value = .03) when $\bar{\Lambda}$ is restricted to be a column of 1's (so that the f_t affects each of the countries equally). The final three rows of panel (b) show that there is no significance time variation when attention is focused on a single country. For example, testing for time variation in μ_{France} while maintaining stability in $\mu_{Germany}$ and μ_{Italy} (so that $\bar{\Lambda}' = (1,0,0)$) yields a p-value of 0.52 as shown in the second row of the table.

The third column panel (b) shows a point estimate of $\sigma_{\Delta f}$ of 0.20% per quarter for the specification with equal factor loadings. Over the 43 year span of the data, this implies a standard deviation for the change in μ of approximately 2.6%. Using this point estimate, the estimated VAR(4) parameters, and $\bar{\Lambda} = \iota$ (a vector of 1's), the realization of f_t can be estimated using a Kalman Smoother, and this can be used to estimate the local level process μ_t for each country. This estimate is shown as the smooth series plotted in figure 4.

6. Discussion and Conclusions

This paper has investigated the problem of testing for time varying coefficients in a regression model when only a few linear combinations of the coefficients are potentially time varying. Optimal tests in this reduced rank time varying coefficient model were developed; analogous tests were also constructed for a version of the multivariate local level model in which a small number of common factors drive the variation in the vector level process.

Analysis of the optimal tests led to three main conclusions. First, the power performance of the easy-to-compute "sup-test" is essentially identical to the optimal test, making this test an attractive alternative to tests currently in use. Second, while the restricted rank information leads to power increases over optimal tests that do not use this information, the power gains are not large. Finally, power gains can be obtained using information about which linear coefficients of regression coefficients are unstable in the regression model or about the values of the factor loadings in the local level model. This "direction" information can lead to large power gains even when it is only approximately correct.

There are several issues related to the testing problem that have not been addressed here. The analysis used a pre-specified value of r, the number of linear combinations of potentially time varying coefficients, and it would be useful to consider the problem of estimating r. Because tests using the true value of r have local power only slightly greater than tests using the wrong value of r (compare the "sup" and "trace" tests, for example), it seems likely that r cannot be precisely estimated when there is only a small amount of time variation. It would also be useful to develop methods for estimating the unknown parameters in Λ or R when the amount of time variation is small, perhaps by generalizing the procedures developed by Stock and Watson (1998) in the model with $k = 1$.

Appendix

This appendix contains proofs of theorems that are not included in the text. We start with a set of standard results that will be used in the proofs. For any sequence $\{c_t\}$ define $c_T(s) = T^{-1/2}\sum_{t=1}^{[sT]} c_t$. Let $G(\mathrm{L})$ denote a matrix of 1-summable lag polynomials ($\sum_{i=0}^{\infty} i\left|G_i\right| < \infty$) and $D(\mathrm{L})$ to denote a matrix of absolutely summable lag polynomials ($\sum_{i=0}^{\infty}\left|D_i\right| < \infty$). Let $g_t = G(\mathrm{L})e_t$ and $d_t = D(\mathrm{L})e_t$. Then

$$T^{-1/2}\sum_{t=1}^{[sT]} e_t \Rightarrow W(s) \quad \text{(AP.1)}$$

follows from F.1.

$$g_T(s) = T^{-1/2}\sum_{t=1}^{[sT]} G(L)e_t \Rightarrow G(1)W(s) \quad \text{(AP.2)}$$

follows from Stock (1994, equation 2.9); see also Phillips and Solo (1992, Theorem 3.4).

$$T^{-1}\sum_{t=1}^{T} d_t d_t' \xrightarrow{p} \sum_{i=0}^{\infty} D_i D_i' \quad \text{(AP.3)}$$

follows from Fuller (1976, Lemma 6.5.1) extended to martingale differences in a straightforward manner.

A.1 Proof of Theorem 2

Without loss of generality because of invariance, we assume $\mu_0 = 0$ and $P = \bar{P} = [I_r\ 0_{r\times(k-r)}]'$.

A.1.1 Part (a)

We begin the proof by considering the limiting behavior of several random variables that characterize the test statistic. Let $X_t = \theta_a(1)^{-1}Z_t = \theta_a(1)^{-1}\Lambda f_t + \theta_a(1)^{-1}a_t$,

$Y_t = X_t - T^{-1}\sum_{i=1}^{T} X_i$, $q_t = T^{-1/2}\sum_{i=0}^{t-1}(1-T^{-1}\bar{g})^i Y_{t-i}$, $Q_t = \bar{g}\, T^{-1/2} q_{t-1}$, $S_t = Y_t - Q_t$, $J_T =$ $T^{-1/2}\sum_{i=1}^{T}(1-T^{-1}\bar{g})^{i-1} S_i$, $\upsilon_t = \theta_a(1)^{-1}\Lambda f_t$, and $b_t = \theta_a(1)^{-1}a_t$. Then

$$X_T(s) = g\,\theta_a(1)^{-1}T^{-3/2}\sum_{t=1}^{[sT]}\sum_{j=1}^{t}\Lambda H(L)\eta_j + b_T(s)$$

$$\Rightarrow X(s) = g\bar{P}\int_0^s W_2(\tau)d\tau + W_1(s). \tag{AP.4}$$

follows from (AP.2) and the the normalization in (N.3i) (because H(L) and θ_a(L) are 1-summable by (F.2) and (F.3)). Also

$$Y_T(s) \Rightarrow Y(s) = X(s) - sX(1) \tag{AP.5}$$

$$q_{[sT]} \Rightarrow q(s) = \int_0^s e^{-\bar{g}(s-\tau)}dY(\tau) \tag{AP.6}$$

$$Q_T(s) \Rightarrow Q(s) = \bar{g}\int_0^s q(\tau)d\tau \tag{AP.7}$$

$$S_T(s) \Rightarrow S(s) = Y(s) - Q(s) \tag{AP.8}$$

$$J_T \Rightarrow \int_0^1 e^{-\bar{g}\tau}dS(\tau) = J \tag{AP.9}$$

$$T^{1/2}\upsilon_{[sT]} \Rightarrow \upsilon(s) = g\bar{P}\,W_2(s) \tag{AP.10}$$

$$\sum_{t=1}^{T}\upsilon_t b_t{}' \Rightarrow g\bar{P}\int_0^1 W_2(s)dW_1(s)' \tag{AP.11}$$

follow from (AP.1), (AP.4), the continuous mapping theorem, and the 1-summability of H(L) and θ_a(L).

$$T^{-1}\sum_{t=1}^{T} X_t \xrightarrow{p} 0 \tag{AP.12}$$

$$T^{-1}\sum_{t=1}^{T} b_t b_t' \xrightarrow{p} \Sigma_{bb} = \theta_a(1)^{-1}[\sum_{i=0}^{\infty}\theta_{a,i}\theta_{a,i}']\theta_a(1)^{-1}' \tag{AP.13}$$

$$T^{-1}\sum_{t=1}^{T} \upsilon_t \upsilon_t' \xrightarrow{p} 0 \tag{AP.14}$$

$$T^{-1}\sum_{t=1}^{T} \upsilon_t b_t' \xrightarrow{p} 0 \tag{AP.15}$$

where (AP.12) follows from (AP.4), (AP.13) follows from (F.1), (F.2) and (AP.3), (AP.14) follows from (AP.10), and (AP.15) follows from(AP.11). In addition,

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T} Y_t Y_t' &= T^{-1}\sum_{t=1}^{T} X_t X_t' - (T^{-1}\sum_{t=1}^{T} X_t)(T^{-1}\sum_{t=1}^{T} X_t)' \\
&= T^{-1}\sum_{t=1}^{T} X_t X_t' + o_p(1) \\
&= T^{-1}\sum_{t=1}^{T} b_t b_t' + T^{-1}\sum_{t=1}^{T} \upsilon_t \upsilon_t' \\
&\quad + T^{-1}\sum_{t=1}^{T} b_t \upsilon_t' + T^{-1}\sum_{t=1}^{T} \upsilon_t b_t' + o_p(1) \\
&\xrightarrow{p} \Sigma_{bb}
\end{aligned} \tag{AP.16}$$

follows from (AP.12)-(AP.15);

$$\begin{aligned}
\sum_{t=1}^{T} Q_t Q_t' &= \bar{g}^2 T^{-1}\sum_{t=1}^{T-1} q_t q_t' \\
&\Rightarrow \bar{g}^2 \int_0^1 q(s)q(s)'ds
\end{aligned} \tag{AP.17}$$

follows from the definition of Q_t and (AP.6);

$$\begin{aligned}
\sum_{t=1}^{T} Q_t Y_t' + \sum_{t=1}^{T} Y_t Q_t' &= \bar{g}[T^{-1/2}\sum_{t=1}^{T} q_{t-1} Y_t' + T^{-1/2}\sum_{t=1}^{T} Y_t q_{t-1}'] \\
&= (1-\bar{g}T^{-1})^{-1}\bar{g}[q_T q_T' + (2\bar{g}T^{-1} - \bar{g}^2 T^{-2})\sum_{t=1}^{T} q_{t-1}q_{t-1}' - T^{-1}\sum_{t=1}^{T} Y_t Y_t'] \\
&\Rightarrow \bar{g}[q(1)q(1)' + 2\bar{g}\int_0^1 q(s)q(s)'ds - \Sigma_{bb}]
\end{aligned}$$

(AP.18)

where the first line uses the definition of Q_t, the second line follows from squaring both sides of $q_t = (1-\bar{g}\,T^{1})q_{t-1} + T^{1/2}Y_t$, and the convergence follows from (AP.6) and (AP.16)

Let $r = 1 - T^{-1}\bar{g} + o(T^{-1})$, $\tilde{q}_t = T^{-1/2}\sum_{i=0}^{t-1} r^i Y_{t-i}$, $\tilde{Q}_t = T^{1/2}(1-r)\tilde{q}_{t-1}$, $\tilde{S}_t = Y_t - \tilde{Q}_t$,

$\tilde{J}_T = T^{-1/2}\sum_{i=1}^{T} r^{i-1}\tilde{S}_i$. Then $\sum_{t=1}^{T}\tilde{Q}_t\tilde{Q}_t' - \sum_{t=1}^{T}Q_tQ_t' \xrightarrow{p} 0$, $\sum_{t=1}^{T}\tilde{Q}_tY_t' - \sum_{t=1}^{T}Q_tY_t' \xrightarrow{p} 0$, and

$\tilde{J}_T - J_T \xrightarrow{p} 0$ follows from the results above and $\sup_{0\le s\le 1} r^{[sT]} - (1 - T^{-1}\bar{g})^{[sT\}} \to 0$.

With these preliminary results in hand, we now prove the theorem. Let $D^I_{\bar{g}/T}$ denote the infeasible version of $\hat{D}_{\bar{g}/T}$ constructed using $\theta_a(1)$ in place of $\hat{\theta}_a(1)$. Using the Moore-Penrose this can be written as :

$$D^I_{\bar{g}/T} = X'M_l[M_l\, V_{\bar{g}/T}\, M_e]^+\, M_l\, X - X'\, M_l\, X. \qquad \text{(AP.19)}$$

From Elliott and Müller (2002, lemma 5)

$$[M_l\, V_{\bar{g}/T}\;\; M_l]^+ = rG - rGl(l'Gl)^{-1}l'G, \qquad \text{(AP.20)}$$

where $G = G^{1/2\prime}G^{1/2}$ with

$$G^{1/2} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -(1-r) & 1 & 0 & \cdots & 0 \\ -r(1-r) & -(1-r) & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -r^{T-2}(1-r) & -r^{T-3}(1-r) & \cdots & -(1-r) & 1 \end{bmatrix}$$

and

$$r = (1 - \bar{g}\,T^{-1}) + o(T^{-1}) \qquad \text{(AP.21)}$$

Let $R = (1\; r\; r^2\; .\;.\;.\; r^{T-1})'$, Y denote a $T\times k$ matrix with rows given by Y_t', and S, Q, $\tilde{S}$ and $\tilde{Q}$ be defined analogously. Then a straightforward calculation shows that $Y{=}M_lX$, $\tilde{S} = G^{1/2}Y$, $R = G^{1/2}l$ and $R'R = (1 - r^{2T})/(1 - r^2)$ for $r \neq 1$. Thus

$$\begin{aligned} D^I_{\bar{g}/T} &= X'M_l[M_l V_{\bar{g}/T}\, M_l]^+ M_lX - X'M_lX \\ &= X'M_l[rG - rGl(l'Gl)^{-1}l'G]M_lX - X'M_lX \end{aligned}$$

$$= r\tilde{S}'\tilde{S} - \frac{r(1-r^2)}{1-r^{2T}}(\tilde{S}'RR'\tilde{S}) - Y'Y$$

$$= (r-1)Y'Y - r(Y'\tilde{Q} + \tilde{Q}'Y) + r\tilde{Q}'\tilde{Q} - \frac{r(1-r^2)}{1-r^{2T}}(\tilde{S}'RR'\tilde{S}) \qquad \text{(AP.22)}$$

Now,

$$(r-1)Y'Y = -\bar{g}\,T^{-1}\sum_{t=1}^{T} Y_t Y_t' + o_p(1)$$

$$\xrightarrow{p} -\bar{g}\,\Sigma_{bb} \qquad \text{(AP.23)}$$

from (AP.16),

$$r(Y'\tilde{Q} + \tilde{Q}'Y) \Rightarrow \bar{g}[q(1)q(1)' + 2\bar{g}\int_0^1 q(s)q(s)'ds - \Sigma_{bb}] \qquad \text{(AP.24)}$$

from (AP.18),

$$r\tilde{Q}'\tilde{Q} \Rightarrow \bar{g}^2\int_0^1 q(s)q(s)'ds \qquad \text{(AP.25)}$$

from (AP.17), and

$$\frac{r(1-r^2)}{1-r^{2T}}(\tilde{S}'RR'\tilde{S}) = T\frac{r(1-r^2)}{1-r^{2T}}\tilde{J}_T\tilde{J}_T' \Rightarrow \frac{2\bar{g}}{1-e^{-2\bar{g}}}JJ'. \qquad \text{(AP.26)}$$

Substituting (AP.23), (AP.24), (AP.25), and (AP.26) into (AP.22) yields the result in the theorem.

To complete the proof, note that $\hat{D}_{\bar{\gamma}} = \hat{\theta}_a(1)^{-1}\theta_a(1)D^I_{\bar{\gamma}}(\hat{\theta}_a(1)^{-1}\theta_a(1))' \xrightarrow{p} D^I_{\bar{\gamma}}$ by Slutksy's theorem.

A.2 Proof of Theorem 3

Without loss of generality we set $\beta_0 = 0$, and we use notation that parallels the proof of theorem 2. Part (a) follows from the same logic as Theorem 1 if

$\exp\{-\frac{1}{2} trace[S' D_{\bar{\gamma}} S]\}$ can be shown to be asymptotically equivalent to the best invariant test with $S = P$ known, where invariance is with respect to transformations of the form $y_t \to y_t + x_t' b$. Inspection of Theorem 2 of Elliot and Müller (2002) shows that this result follows if

$$\sup_s T^{-1/2} \sum_{t=1}^{[sT]} (x_t x_t' - \Sigma_{xx}) \beta_t \xrightarrow{p} 0 \tag{AP.27}$$

and

$$\sup_s T^{-1} \sum_{t=1}^{[sT]} (x_t x_t' - \Sigma_{xx}) \beta_t \beta_t' \xrightarrow{p} 0 . \tag{AP.28}$$

From (G.4), the elements of $T^{1/2}\beta_t$ and $\beta_t\beta_t'$ have bounded fourth moments, so that (AP.27) and (AP.28) follow from (R.1) and Lemma 2 of Stock and Watson (1998).

Part (b) follows from an argument like that used to prove theorem 2. Write $Z_t = x_t \hat{u}_t = a_t + \mu_t + \hat{w}_t$, where $\hat{w}_t = (x_t x_t' - \Sigma_{xx})\beta_t - x_t x_t' \hat{\beta}$. Inspection of that proof of the theorem reveals that part (b) of theorem 3 will hold if $Y_T(s)$ behaves as in (AP.5) and $T^{-1} \sum Y_t Y_t$ behaves as in (AP.16). Given the assumption on μ_t and a_t, these follow from $\sup_s (\hat{w}_T(s) - s\hat{w}_T(1)) \xrightarrow{p} 0$. To show this write

$$\hat{w}_T(s) - s\hat{w}_T(1) = T^{-1/2} \sum_{t=1}^{[sT]} (x_t x_t' - \Sigma_{xx}) \beta_t - sT^{-1/2} \sum_{t=1}^{T} (x_t x_t' - \Sigma_{xx}) \beta_t$$
$$- T^{-1} \sum_{t=1}^{[sT]} (x_t x_t' - S_{xx})(\sqrt{T} \hat{\beta})$$

(AP.29)

where $S_{xx} = T^{-1} \sum_{t=1}^{T} x_t x_t$. The first two terms are uniformly negligible by (AP.27). For the last term, from Lemma 2 of Stock and Watson (1998),

$$\sup_s T^{-1}\sum_{t=1}^{[sT]}(x_t x_t' - \Sigma_{xx}) \xrightarrow{p} 0. \qquad \text{(AP.30)}$$

Also,

$$\sqrt{T}\hat{\beta} = S_{xx}^{-1}T^{-1/2}\sum a_t + o_p(1) \qquad \text{(AP.31)}$$

so that $\sqrt{T}\hat{\beta}$ is $O_p(1)$. Thus $\sup_s(\hat{w}_T(s) - s\hat{w}_T(1)) \xrightarrow{p} 0$ as required.

Data Appendix

Real GDP series were used for the sample period 1960:1–2002:4. In the cases of France and Italy, series from two sources were spliced. The table below gives the data sources and sample periods for each data series used. Abbreviations used in the source column are (DS) DataStream, and (E) for the OECD Analytic Data Base series from Dalsgaard, Elmeskov, and Park (2002), provided to us by Jorgen Elmeskov via Brian Doyle and Jon Faust. Three outlying data points associated with a general strike in France and German reunification were eliminated from dataset. The dates are 1968:2-1968.3 for France and 1991.1 for Germany.

Country	Series Name	Source	Sample period
France	frona017g	OECD (DS)	1960:1 1977:4
	frgdp…d	I.N.S.E.E. (DS)	1978:1 2002:4
Germany	bdgdp…d	DEUTSCHE BUNDESBANK (DS)	1960:1 2002:4
Italy		OECD (E)	1960:1 1969:4
	itgdp…d	ISTITUTO NAZIONALE DI STATISTICA (DS)	1970:1 2002:4

References

Andrews, W.W.K. and W. Ploberger (1994), "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62, 1383-1414.

Elliott, Graham and Ulrich K. Müller (2002), "Optimally Testing General Breaking Processes in Linear Time Series Models," manuscript (revised April 2003).

Dalsgaard, T., J. Elmeskov, and C-Y Park (2002), "Ongoing Changes in the Business Cycle – Evidence and Causes," OECD Economics Department Working Paper no. 315.

Fuller, W.A. (1976) *Introduction to Statistical Time Series*, New York: Wiley.

Harvey, A.C. (1989), *Forecasting, structural time series models and the Kalman Filter*, Cambridge: Cambridge University Press.

Jansson, M. (2002), "Point Optimal Tests of the Null Hyptothesis of Cointegration," manuscript, Department of Economics, UC Berkeley.

King, M.L. (1980), "Robust Tests for Spherical Symmetry and Their Application to Least Squares Regression," *The Annals of Statistics*, Vol. 8, No. 6, 1265-1271.

Leybourne, S. J. "Estimation and Testing of Time-varying Coefficient Regression Models in the Presence of Linear Restrictions," *Journal of Forecasting*, Vol. 12, 49-62.

Phillips, P.C.B. and V. Solo (1992), "Asymptotics for Linear Processes," *Annals of Statistics*, Vol. 20, No. 2, pp. 971–1001.

Shively, T.S. (1988), "An Exact Test of a Stochastic Coefficient in a Time Series Regresison Model," *Journal of Time Series Analysis*, 9, 81-88.

Stock, J.H. (1994) , "Unit Roots, Structural Breaks, and Trends," in R.F. Engle and D. McFadden (eds), *Handbook of Econometrics*, Volumne IV, Amsterdam: Elsevier.

Stock, J.H. and M.W. Watson (1988), "Testing for Common Trends," *Journal of the American Statistical Association,* December, 83, pp. 1097-1107.

Stock, J.H. and M.W. Watson (1998), "Asymptotically Median Unbiased Estimation of Coefficient Variance in a Time Varying Parameter Model," *Journal of the American Statistical Association*, Vol. 93, No. 441, pp 349-358.

Table 1
Asymptotic Critical Values for the *Sup* and *Trace* Tests

k	ξ_{sup}				ξ_{trace}		
	10%	5%	1%		10%	5%	1%
1	7.12	8.32	10.85		7.12	8.32	10.85
2	9.27	10.60	13.59		12.79	14.28	17.55
3	11.03	12.54	15.74		18.13	19.91	23.50
4	12.50	13.99	17.52		23.31	25.26	29.48
5	13.99	15.58	19.11		28.30	30.41	34.68
6	15.36	17.08	20.76		33.40	35.70	40.46
7	16.70	18.47	22.45		38.40	40.81	45.67
8	18.06	19.90	23.80		43.45	45.94	51.59
9	19.33	21.32	25.47		48.30	51.00	56.49
10	20.57	22.51	26.97		53.23	56.07	61.73
11	21.79	23.81	28.23		57.93	60.82	66.81
12	23.01	25.22	29.60		62.94	65.91	71.79
13	24.18	26.33	31.10		67.73	70.95	77.02
14	25.43	27.62	32.32		72.47	75.78	82.52
15	26.56	28.79	33.43		77.33	80.49	87.00
16	27.69	30.02	35.09		82.25	85.64	92.28
17	28.87	31.28	36.31		87.15	90.58	97.75
18	29.91	32.21	37.37		91.78	95.22	102.40
19	31.04	33.61	38.70		96.41	100.25	107.61
20	32.24	34.73	40.21		101.37	105.07	112.53

Notes: ξ_{sup} is the largest eigenvalue of $-D_{\bar{\gamma}}$ and ξ_{trace} is the trace of $-D_{\bar{\gamma}}$, where $\bar{\gamma}$ =10/T and $D_{\bar{\gamma}}$ is defined in the text. The asymptotic critical values were estimated computed using 30,000 draws from the Gaussian model with T = 500.

Table 2
Average Per Capita Real GDP Growth Rates (PAAR)

Sample Period	France	Germany	Italy
1960-1969	4.07	3.83	4.55
1970-1979	2.97	2.59	3.61
1980-1989	1.78	1.57	2.17
1990-2002	1.34	1.31	1.32

Table 3
Tests for Time Varying Mean Growth Rates

a. Invariant Tests

Test	P-value
B	0.12
ξ_{sup}	0.11
ξ_{trace}	0.16

b. NonInvariant Tests Based on a Prespecified Factor Loading Matrix (Λ)

$\bar{\Lambda}$	P-value	Median Unbiased Estimate of $\sigma_{\Delta f}$	90 % Confidence Interval for $\sigma_{\Delta f}$
1, 1, 1	0.03	0.20	0.04 - 0.55
1, 0, 0	0.52	0.00	0.00 – 0.18
0, 1, 0	0.90	0.00	0.00 – 0.08
0, 0, 1	0.17	0.09	0.00 – 0.30

Notes: All of the tests are based on $\hat{D}_{\bar{\gamma}}$ as described in theorem 2, using a VAR(4) estimator of θ_a(L) and $\bar{\gamma} = 10/T$. In panel (a) B denotes the asymptotically point optimal invariant test, ξ_{sup} is the largest eigenvalue of $-\hat{D}_{\bar{\gamma}}$ and ξ_{trace} is the trace of $-\hat{D}_{\bar{\gamma}}$. The tests in panel (b) were computed for the fixed values of the factor loadings $\bar{\Lambda}$, summarized in the first column of the table. $\hat{\sigma}^{mub}_{\Delta f}$ is the median unbiased estimator. Estimates of $\sigma_{\Delta f}$ and 90% confidence interval were computed by inverting the percentiles of the test statistic as described in Stock and Watson (1998).

Figure 1
Power of Invariant Tests

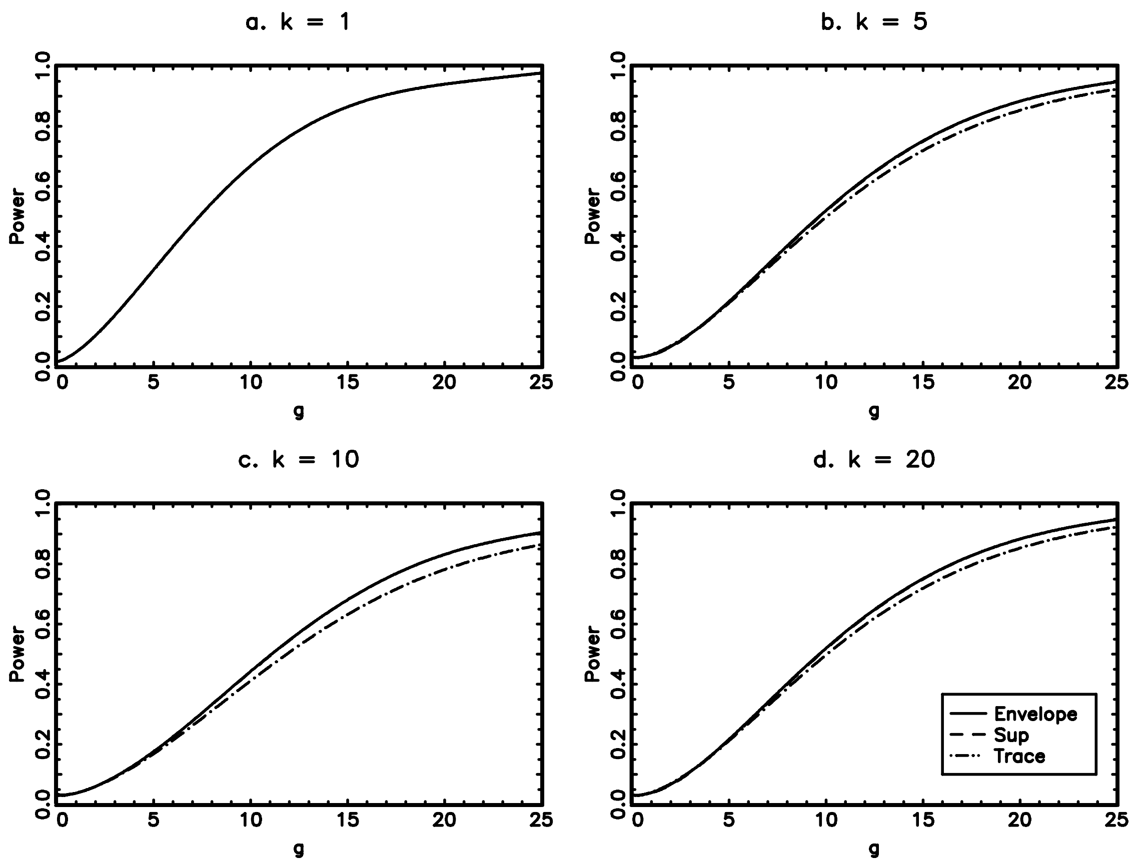

Figure 2
Power of Point-Optimal Invariant Tests

‘

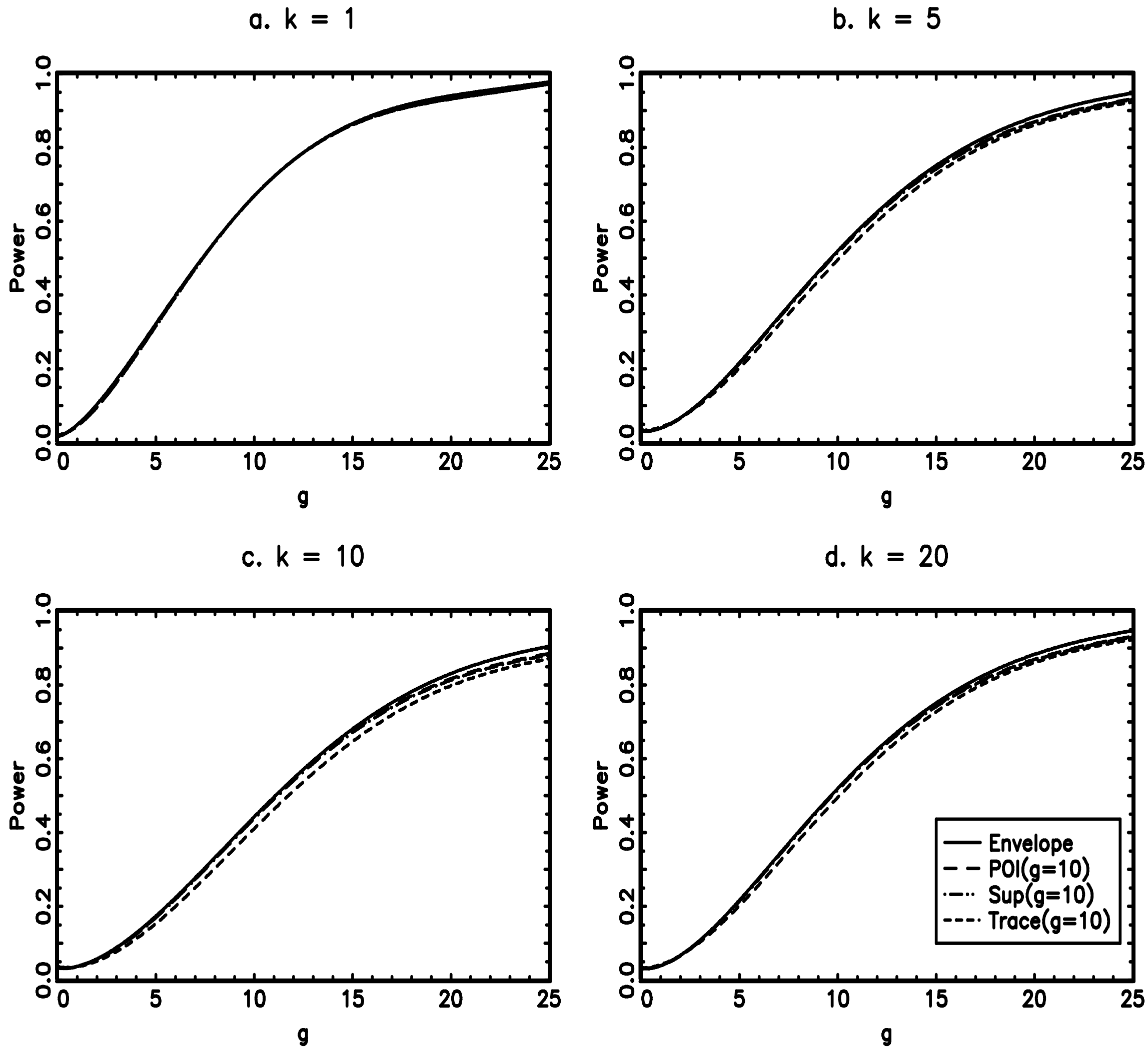

Figure 3
Power of Noninvariant Tests

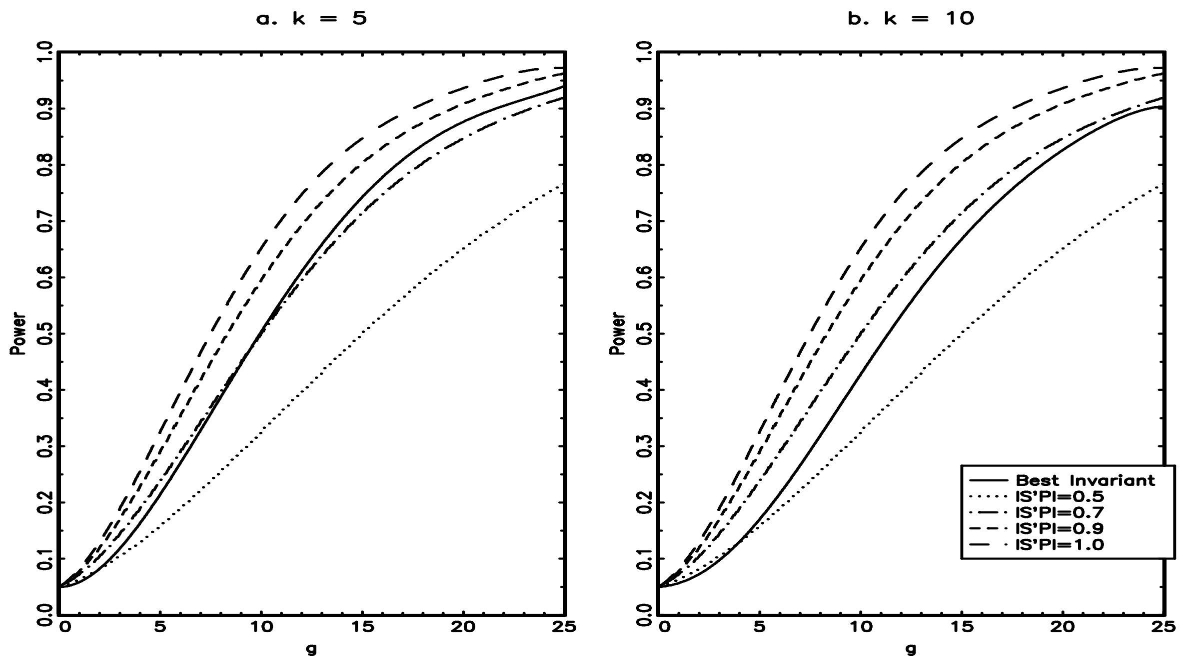

Figure 4
Real GDP Per Capita Growth Rates (thick line) and Estimated Trend (thin line)