

UNIVERSIDADE FEDERAL DO CEARÁ - UFC

**FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, ATUÁRIA E
CONTABILIDADE**

CURSO DE PÓS-GRADUAÇÃO EM ECONOMIA - CAEN

**Estimando o Impacto do Estoque de Capital Público sobre o
PIB per capita na Presença de Mudança Estrutural**

JIMMY LIMA DE OLIVEIRA

**Fortaleza
2006**

JIMMY LIMA DE OLIVEIRA

**Estimando o Impacto do Estoque de Capital Público sobre o
PIB per capita na Presença de Mudança Estrutural**

Dissertação de Mestrado apresentada
ao curso de Pós-Graduação em
Economia da Universidade Federal do
Ceará – CAEN/UFC, como requisito
parcial para obtenção do título de
Mestre em Economia.

Orientador: Prof. Dr. Sérgio Aquino de Souza

**Fortaleza
2006**

Esta dissertação foi submetida como parte dos requisitos necessários à obtenção do título de Mestre em Economia, outorgado pela Universidade Federal do Ceará, e encontra-se à disposição dos interessados na Biblioteca do Curso de Pós-Graduação em Economia - CAEN da referida Universidade.

A citação de qualquer trecho desta dissertação é permitida, desde que seja feita em conformidade com as normas científicas.

Jimmy Lima de Oliveira

Dissertação aprovada em _____ de _____ de _____

Prof. Sérgio Aquino de Souza
(Orientador)

Prof. Marcelo Lettieri Siqueira
(Membro da Banca Examinadora)

Prof. Samuel Façanha Câmara
(Membro da Banca Examinadora)

AGRADECIMENTOS

Aos meus pais, Aldenir e Cleonice, que me ensinaram a não temer desafios e a superar os obstáculos com humildade.

A minha avó materna, Cleia, pelos conselhos e ensinamentos recebidos.

As minhas irmãs, Jamille e Rayana e aos meus sobrinhos, Maria Eduarda e João Pedro.

A minha namorada, Raquel Peixoto, pela compreensão e apoio nos momentos difíceis.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, pela concessão de financiamento, através da bolsa de estudo durante o curso.

A todos os professores e funcionários do CAEN que sempre se dispuseram a ajudar nos momentos de necessidade. Em especial, aos professores Flávio Ataliba e Edinaldo Tebaldi.

Ao professor Sérgio Aquino, pela atenção dispensada na orientação desta pesquisa.

Aos professores Marcelo Lettieri e Samuel Façanha, por fazerem parte da banca examinadora.

Aos amigos da turma de 2004 do CAEN, André Oliveira, Antônio Germano, Bruno Wichmann, Daniel de Pádua, Débora Gaspar, Leandro Costa, Nicolino Trompieri, Paulo Saraiva e Victor Hugo.

Aos amigos da turma de 2005 do CAEN, Cristiano Pena (Panqueca), Guilherme Irff (Guigui), Helio Berni, Luciano Bauermann (Gaúcho), Vitor Borges (Vitão).

Aos amigos do futebol e do CAEN, Carlos Manso, Dílson Sena, Felipe Oliveira, Marcelo Furtado e Rodolfo Alves.

Ao Cléber, o gordo mais mentiroso que já conheci, pelo seu grande conhecimento em todas as áreas da economia e ciências afins.

E a todos os demais, que de forma direta ou indireta contribuíram na elaboração deste trabalho.

RESUMO

O presente trabalho estima a elasticidade produto-gasto público para economia brasileira, no período de 1950 a 2003, utilizando um modelo vetorial de correção de erro (VECM) para controlar possíveis mudanças estruturais nas séries. Quando existem mudanças estruturais, os vários testes estatísticos de Dickey-Fuller são viesados em direção da não rejeição de uma raiz unitária. Este viés significa que o teste de Dickey-Fuller é viesado em direção da hipótese nula de uma raiz unitária, mesmo se a série é estacionária dentro de cada subperíodo. Sem controlar para mudanças estruturais, os testes de cointegração podem apresentar resultados enganosos, e as estimativas obtidas podem ser viesadas.

Palavras Chaves: Elasticidade Produto-Gasto Público; Modelo Vetorial de Correção de Erros; Mudanças Estruturais.

ABSTRACT

Aiming to estimate the elasticity product-public expenditure to the Brazilian economy, during the period 1950-2003, it was used a vector error correction model (VECM) to control for possible structural changes in the series. When structural changes were observed, many of the Dickey-Fuller statistic tests are biased towards the non-rejection of the existence of a unit root. This bias means that the Dickey-Fuller test is biased towards the null hypothesis of unit root, even if the series is stationary within each sub period. Without controlling for structural changes, the cointegration tests may present deceiving results and the estimates obtained may be biased.

Keywords: Elasticity Product-Public Expenditure; Vector Error Correction Model; Structural Changes.

SUMÁRIO

1. Introdução	1
2. Revisão de Literatura	3
3. Modelo Econômico	6
4. As Trajetórias do PIB e do Investimento (1950-2003)	10
5. Base de Dados	14
6. Metodologia	16
6.1. VARs e VECMs	17
6.1.1. O Modelo	17
6.1.2. Termos Determinísticos	18
6.2. Estimação	19
6.2.1. Estimação de um VAR sem Restrição	19
6.2.2. Estimação de VECMs	21
6.3. Estacionariedade	27
6.3.1. Autocorrelações Amostrais e Autocorrelações Parciais	28
6.4. Testes de Raiz Unitária	31
6.4.1. O Teste ADF	31
6.4.2. Teste para Processos com Mudança de Nível	34
6.5. Especificando o Rank de Cointegração	38
7. Resultados	44
7.1. Estimativas	44
7.2. Teste de Chow	45
7.3. Análise de Impulso-Resposta	48
8. Conclusão	52
Referencias Bibliográficas	53

LISTA DE TABELAS

Tabela 1 – Teste de Raiz Unitária Dickey-Fuller Aumentado (ADF)	32
Tabela 2 – Teste de Raiz Unitária com Mudança de Nível	38
Tabela 3 – Teste de Cointegração sem Considerar a Mudança Estrutural	41
Tabela 4 – Teste de Cointegração Considerando a Mudança Estrutural	43
Tabela 5: Relações de Cointegração Estimadas por Máxima Verossimilhança	44
Tabela 6 Relações de Cointegração Estimadas pelo Método de Dois Estágios	45
Tabela 7: Resultado do Teste de Chow	48

LISTA DE GRÁFICOS E FIGURAS

Figura 1- Funções de Autocorrelações Amostrais e Autocorrelações Parciais	30
Gráfico 1 – Tamanho do Governo	9
Gráfico 2: Logaritmo do PIB per capita	14
Gráfico 3: Logaritmo do Estoque de Capital Público	14
Gráfico 4: Logaritmo do PIB per capita em primeira diferença	33
Gráfico 4: Logaritmo do Estoque de Capital Público em primeira diferença	33
Gráfico 5: Teste de Raiz Unitária com Mudança de nível para o PIB per capita em primeira diferença	37
Gráfico 6: Teste de Raiz Unitária com Mudança de nível para o Estoque de Capital Público em primeira diferença	37
Gráfico 7: Função Impulso-Resposta de um Choque de Uma Unidade de Desvio-Padrão no Estoque de Capital Público	51

1. Introdução

Nos últimos anos a questão do impacto do capital público, principalmente no que diz respeito aos investimentos em infra-estrutura, sobre o produto e a produtividade total dos fatores tem ocupado lugar de destaque no debate econômico. A idéia de que investimentos em infra-estrutura afetam positivamente a produtividade dos fatores privados ganhou relevância a partir dos modelos de crescimento endógeno.

A idéia é que um dado conjunto de fatores privados ao ter acesso a melhores estradas, portos, aeroportos, água, energia e comunicação de qualidade e a baixo custo, obtém um aumento do produto final, com conseqüente aumento de produtividade e redução de custo unitário. A maior produtividade representa, por sua vez, aumento na remuneração dos fatores, o que estimula o investimento e o emprego.

Com a estagnação dos investimentos públicos, especialmente em infra-estrutura, torna-se imperativo investigar a relação entre investimento público, infra-estrutura e crescimento, de modo a possibilitar avaliação dos custos da redução do papel do setor público na economia brasileira.

A taxa de investimento da economia brasileira apresentou uma trajetória crescente no período compreendido entre o imediato pós-guerra até meados dos anos 70. Após atingir um pico de 25% do PIB em 1975/1976, houve uma contínua e progressiva queda dessa taxa. . A sua desaceleração a partir de meados dos anos 70 foi substituída pela redução absoluta e variações intensas ao longo dos anos 80. O patamar de investimentos em infra-estrutura declinou ainda mais ao longo dos anos 90.

A taxa de crescimento do PIB acompanhou de perto a trajetória da taxa de investimento. Durante o período 1956-61 o PIB cresceu à taxa anual de 8,2%, o que significou uma elevação de 5,1% ao ano da renda per capita. Entre 1968 e 1973, o PIB real cresceu à taxa média de 11,2%, alcançando um máximo de 14% em 1973. A quadruplicação dos preços do petróleo no final de 1973 estreitou as opções de crescimento do Brasil, país fortemente dependente do petróleo importado. A manutenção do crescimento ao final da década de 70 ainda foi possível em razão da

existência de financiamento externo. O segundo choque do petróleo e a elevação das taxas de juros internacionais em 1979 marcam o fim do período de crescimento do PIB às taxas históricas observadas durante as décadas anteriores.

O presente trabalho analisa a relação de longo prazo entre as variáveis PIB per capita e estoque de capital público levando em consideração as mudanças estruturais observadas em ambas as séries. As mudanças estruturais são reflexos da acentuada redução das taxas de investimento do setor público e de crescimento do PIB per capita nas últimas décadas.

O objetivo deste trabalho é estimar a elasticidade produto-capital público para economia brasileira, no período de 1950 a 2003, utilizando um modelo vetorial de correção de erro (VECM) levando em consideração a ocorrência de mudanças estruturais nas séries.

2. Revisão de Literatura

Diversos trabalhos empíricos sugerem que os gastos públicos podem elevar o crescimento econômico por meio do aumento da produtividade do setor privado. O artigo de Aschauer (1989) para dados da economia americana foi o primeiro a abordar este tema. Utilizando Mínimos Quadrados Ordinários (MQO), ele estimou que um aumento de 1% no capital público implicaria em um aumento entre 0,36 e 0,39 % no produto. Munnell (1990) obteve estimativas semelhantes para dados regionais americanos. Easterly e Rebelo (1993) estudaram o relacionamento cross-country entre diferentes categorias de investimento público e o crescimento econômico. Seus principais resultados revelam que o investimento público é positivamente correlacionado com o crescimento econômico.

No Brasil, estudos recentes exploram os efeitos do capital público sobre o crescimento e a produtividade. Ferreira (1996) estima o impacto do capital de infraestrutura federal (telecomunicações, energia elétrica, portos, setor marítimo e ferrovias) e o impacto do capital público total (capital das estatais e administrações públicas) sobre o PIB. As estimativas mostraram que um aumento de 1% no capital de infra-estrutura gerava, no longo prazo, um aumento entre 0,34 e 1,12 % no PIB, dependendo da taxa de depreciação utilizada (6, 8 ou 10%). Quanto à série ampla, capital público total, o impacto estimado sobre o PIB situou-se entre 0,71 e 1,05 %. O método de estimação utilizado para a obtenção destas elasticidades foi o de cointegração. Os seus resultados indicavam uma forte relação entre investimentos em infra-estrutura e PIB no longo prazo.

Utilizando dados para o período de 1950-1995, Ferreira e Malliagos (1998) estimam o impacto de longo de prazo do capital de infra-estrutura sobre o PIB e a produtividade total dos fatores (PTF). Além disso, eles investigam também o sentido de causação entre as variáveis. Utilizando uma análise de cointegração baseada na metodologia de Johansen (1991, 1995), os autores encontram estimativas para elasticidade-renda que se situam entre 0,55 e 0,61.

Baseado no modelo de crescimento endógeno desenvolvido por Barro (1990), Assis e Dias (2004) avaliam o impacto da política fiscal e do nível de tecnologia sobre a taxa de crescimento econômico do Brasil, entre os anos de 1951 e 2000, a partir da metodologia de vetores autorregressivos (VAR). Eles verificam que os gastos do setor público com investimentos geram externalidades positivas que podem ampliar a taxa de crescimento econômico do país. Todavia, ressaltam os autores, um aumento dos gastos financiados pela elevação de impostos leva a uma redução na rentabilidade dos investimentos privados, determinando um efeito total negativo da participação do governo na economia. Concluem que a carga tributária, utilizada para financiar os gastos públicos, causa impacto negativo sobre a taxa de crescimento quando ultrapassa o nível de 25% do produto.

Utilizando dados agregados, Candido Jr. (2001), estimou as elasticidades gasto-produto e o diferencial de produtividade em relação ao setor privado. Os efeitos dos gastos públicos no Brasil foram avaliados por duas metodologias. A primeira, utilizando MQO, permitiu estimar o efeito da externalidade dos gastos e o diferencial de produtividade em relação ao setor privado. No conceito que engloba consumo mais transferências, o efeito externalidade foi negativo. Na segunda definição de gasto total (que inclui os investimentos), os resultados indicam uma externalidade positiva, mas o diferencial de produtividade, em relação ao setor privado, apresentou-se negativo, ou seja, a produtividade do setor público representou apenas 60% da produtividade do setor privado. A segunda metodologia utilizada foi um modelo do tipo ADL (Autoregressive Lag Distributed Model), baseada na abordagem de Hendry. Os valores das elasticidades produto-gasto público nos dois conceitos foram negativos. A equação de curto prazo mostrou que os gastos públicos defasados no período de um ano surtem impactos positivos sobre o PIB. No longo prazo, porém, esse efeito se reverte. Esse conjunto de resultados levou o autor a concluir que a proporção de gasto público no Brasil se encontrava acima de seu nível ótimo.

Ferreira e Araújo (2006) utilizam técnicas econométricas de séries de tempo, particularmente VAR, para estimar o impacto sobre o produto de variações na infraestrutura pública. Eles estimam uma equação de cointegração do produto por trabalhador sobre capital público, capital privado e capital humano. São estimadas quatro equações que diferem na maneira como o capital humano é introduzido, em nível

ou em logaritmo, e no período de tempo considerado. Em todas as equações o coeficiente do capital público foi estimado com o sinal correto, mas na maioria dos casos foi estatisticamente não significativa aos níveis usuais. Quando utilizam um período diferente, 1960-1996 ao invés de 1960-2000, as estimações são mais precisas com respeito ao coeficiente do capital público (0,33). Eles utilizam esse sistema para simular a resposta das variáveis econômicas a choques na infra-estrutura. O impacto das variações no capital público sobre o capital privado e o produto per capita é de magnitude considerável, especialmente se consideradas as respostas de longo-prazo. O produto per capita aumenta em 10% no longo prazo e o capital privado aumenta em quase 20%. As respostas acumuladas são muito altas e bem acima de exercícios similares que utilizam dados para os U.S. e outros países da OCDE.

A partir das séries temporais de produto, investimento público e Produtividade Total dos Fatores (PTF), Candido Jr. (2006) investiga as relações dinâmicas de curto e longo prazo entre investimento público e produto e investimento público e PTF para três países da América do Sul: Argentina, Brasil e Chile, no período de 1970-2000. Os resultados encontrados foram: uma relação de longo prazo positiva entre investimento público e produto em todos os países, e uma relação de longo prazo entre investimento público e PTF positiva somente para a economia chilena. O autor conclui que esse resultado poderia corroborar a maior magnitude da elasticidade produto-investimento público do Chile com relação à Argentina e Brasil, dado que o investimento público na economia do Chile eleva o produto indiretamente, por meio da PTF. Os valores das elasticidades foram bastante próximos na Argentina e no Brasil, da ordem de 0,56 e 0,58, respectivamente. Já no caso do Chile, o valor da elasticidade no período foi aproximadamente sete vezes maior do que na Argentina e Brasil.

3. Modelo Econômico

O modelo teórico aqui apresentado foi desenvolvido por Barro (1990), que destaca a política fiscal como um fator que pode explicar as diferenças de renda per capita e as taxas de crescimento entre os países. Nele, a estrutura tributária e a provisão eficiente de bens e serviços públicos influenciam a produtividade do setor privado e a acumulação de capital.

Barro (1990) considera que a quantidade de bens e serviços públicos entra como insumo na função de produção juntamente com os insumos privados, trabalho e capital.

$$Y = F(K, L, G),$$

onde K é capital privado, L é trabalho e G é bens e serviços públicos. A incorporação dos bens e serviços públicos na função de produção torna possível o crescimento endógeno. Para um nível de bens e serviços públicos constante a função de produção exhibe retornos decrescentes para a acumulação de capital. Contudo, se ambos, bens públicos e insumos privados aumentam, a função de produção apresenta retornos constantes para G e K. Isto deriva das características dos bens públicos. Bens públicos são não rivais e não excludentes.

Note que, a forma da função de produção implica que os bens públicos e os insumos privados são complementares, no sentido que um aumento dos bens públicos aumenta o produto marginal do capital. Logo ser necessário guardar determinada proporção na combinação dos insumos públicos e privados. A função de produção pode ser escrita em termos per capita como.

$$y = f(k.g) = k\varphi(g/k), \varphi' > 0 \text{ e } \varphi'' < 0 \quad (2.1)$$

O governo financia a provisão de bens e serviços públicos retirando uma parcela do produto privado por meio de tributação. Supondo que a tributação seja proporcional à renda. E que a cada período o orçamento seja equilibrado, isto é, que não haja endividamento público.

$$g = \tau \cdot y = \tau \cdot k\varphi(g/k)$$

Assumindo que τ seja constante e, por conseguinte, a razão g/k seja constante ao longo do tempo, o lucro da firma i pós-tributação é dado por

$$L_i[(1-\tau)y - w - (r + \delta)k]$$

Maximização do lucro e a condição de lucro zero implicam que a taxa de aluguel do capital é igual ao produto marginal do capital pós-tributação. A taxa de aluguel do capital é dada por

$$\begin{aligned} r + \delta &= \frac{\partial y}{\partial k} = (1-\tau)\varphi(g/k) \left(1 - \varphi' \cdot \frac{g}{k}\right) \\ &= (1-\tau)\varphi(g/k)(1 - \varepsilon_{yg}) \end{aligned} \quad (2.2)$$

A produtividade marginal do capital depende da relação (g/k) e da elasticidade produto-gasto público, ε_{yg} .

A taxa de crescimento de estado estacionário é encontrada resolvendo o problema de maximização da utilidade intertemporal dos agentes. Supondo um modelo de interação entre gerações na qual a geração corrente maximiza utilidade considerando uma restrição orçamentária sobre um horizonte infinito. O problema de maximização da utilidade pode ser escrito como

$$\text{Max} \int_0^{\infty} u(c) e^{-\rho t} dt \quad (2.3)$$

$$\text{Sujeito a } \dot{k} = (1-\tau)y - c$$

em que c é o consumo per capita e $\rho > 0$ é a taxa de desconto temporal. O Hamiltoniano para este problema é

$$H(c, k, t) = u(c)e^{-\rho t} + \lambda[(1 - \tau)y - c] \quad (2.4)$$

Assumindo uma função utilidade com elasticidade de substituição intertemporal constante, $u(c) = \frac{c^{1-\theta} - 1}{1-\theta}$, em que θ é o inverso da elasticidade de substituição intertemporal, as condições de primeira ordem para este problema são

$$\frac{\partial H}{\partial c} = 0 \Rightarrow c^{-\theta} = \lambda \quad (2.5)$$

$$\dot{\lambda} = \lambda\rho - \frac{\partial H}{\partial k} \Rightarrow \frac{\dot{\lambda}}{\lambda} = \rho - (1 - \tau)y \frac{\partial y}{\partial k} \quad (2.6)$$

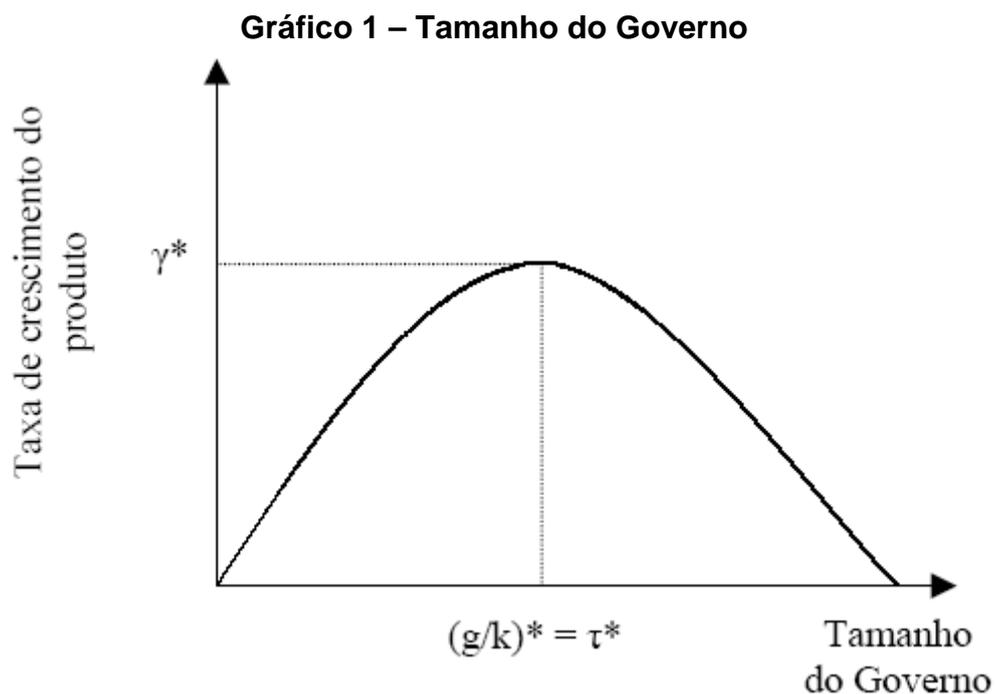
Derivando a equação (2.5) em relação ao tempo e usando a equação (2.6), temos

$$\frac{\dot{c}}{c} = \frac{1}{\theta} \left[(1 - \tau) \frac{\partial y}{\partial k} - \rho \right] \quad (2.7)$$

Substituindo a equação (2.2) na expressão acima, chegamos à taxa de crescimento de estado estacionário:

$$\gamma = \frac{\dot{c}}{c} = \left[(1 - \tau) \varphi(g/k)(1 - \varepsilon_{yg}) - \rho \right] \quad (2.8)$$

A taxa de crescimento do consumo per capita γ é a mesma para o produto per capita e o estoque de capital per capita. O impacto da política fiscal sobre γ se dá por meio de dois canais de transmissão. O primeiro se refere ao efeito positivo dos gastos públicos sobre o produto; e o segundo diz respeito ao efeito negativo dos impostos que reduzem os recursos disponíveis para o setor privado. O efeito líquido depende do tamanho do setor público em termos da relação (g/y) e da correspondente alíquota de tributação necessária para financiar o orçamento público. Assim, para uma relação (g/y) relativamente pequena o efeito da participação do setor público sobre o crescimento econômico é positivo; para um nível elevado de gastos públicos, o efeito da participação do setor público sobre o crescimento econômico é negativo.



Fica evidente, entretanto, que o ponto ótimo τ^* depende decisivamente do grau de eficiência dos gastos públicos. Nesse sentido, caso os gastos impliquem em alto grau de externalidades positivas sobre a lucratividade do setor público, o ponto ótimo τ^* desloca-se para direita.

Com base neste modelo teórico, o objetivo do presente trabalho é estimar a elasticidade produto-gasto público ε_{yg} utilizando um modelo de correção de erros que permita estimar o efeito de longo da infra-estrutura pública sobre o PIB per capita para economia brasileira.

4. As Trajetórias do PIB e do Investimento (1950-2003)

O período imediato ao pós-guerra até o final dos anos 70 se caracteriza pela adoção de um modelo de crescimento econômico baseado no Estado desenvolvimentista interventor. No caso da infra-estrutura econômica, os investimentos foram efetuados, em sua grande maioria, pelas empresas estatais.

Todos os planos elaborados durante o período colocavam nas mãos do setor público a tarefa de conduzir as obras de infra-estrutura que objetivavam romper os estrangulamentos que impediam o processo de industrialização.

As obras necessárias à adequação da infra-estrutura econômica do país ao crescimento industrial sustentado desde 1948 representavam uma forte pressão sobre os gastos públicos, sendo o melhor exemplo nesse sentido o surto de investimentos de 1951 e 1952 representados pelos gastos com a ampliação da capacidade de geração de energia elétrica.

Dentro da tradição dos planos de desenvolvimento cabe destacar o Plano de Metas, elaborado de forma a realizar as inversões de capital público em obras de natureza básica ou infra-estrutural e facilitar e estimular as atividades e investimentos privados. Os setores de energia e transporte foram as principais áreas de investimentos, com mais de 70% dos investimentos, a cargo quase integralmente do setor público.

Durante o período 1956-61 a formação de capital do governo central cresceu a taxa anual real de mais de 15%, enquanto as empresas públicas aumentaram seus gastos em investimento em mais de 25% ao ano. O PIB cresceu à taxa anual de 8,2%, o que significou uma elevação de 5,1% ao ano da renda per capita.

A participação da formação bruta de capital fixo no PIB, depois de permanecer no nível médio de 15,2% no período 1964-66, aumentou a partir de 1967, quando alcançou 16,2%. De 1968 a 1970, essa participação passou para 18,9% até alcançar 20,5% no período 1971-73.

Em 1967 foi lançado o Plano Estratégico de Desenvolvimento cujo objetivo básico declarado pelo governo era o desenvolvimento econômico. A geração de energia elétrica e os serviços de transporte e comunicação cresceram de forma sustentada como resultado das políticas setoriais estabelecidas pelo governo.

A partir de 1968, tanto o produto global como os produtos setoriais apresentaram forte crescimento. Entre 1968 e 1973, o PIB real cresceu à taxa média de 11,2%, alcançando um máximo de 14% em 1973, diante de uma média histórica no período do pós-guerra até o início dos anos 60, da ordem de 7%. Os investimentos das estatais cresceram a uma taxa de quase 20% ao ano em termos reais no mesmo período.

No tocante à política regional, o I Plano Nacional de Desenvolvimento (1972-74) definiu uma política nacional de desenvolvimento regional. A ênfase da política regional no Brasil era no sentido de promover o desenvolvimento do Nordeste, através de incentivos fiscais e obras de infra-estrutura.

Além do I PND, foi levado adiante o Programa de Integração Nacional que, além da irrigação de amplas áreas no Nordeste e da criação de corredores de exportação na região, previa a abertura da região amazônica através de rodovias, a Transamazônica e a Cuiabá-Santarém.

A intervenção no processo econômico através de mecanismos fiscais e creditícios representaram uma deterioração da receita pública. Por outro lado, o aumento das responsabilidades do setor público pelos investimentos requereu um maior endividamento.

Essa necessidade de financiamento foi crescente na medida em que o déficit das estatais passou a corresponder a uma parcela crescente do PIB, aumentando de uma média de 1,5% entre 1967 e 1970 para 2,1% entre 1971 e 1973, paralelamente ao aumento da participação do investimento das estatais no PIB.

Com a expansão da liquidez internacional no final dos anos 60 e início dos anos 70, as estatais recorreram, de forma crescente, a empréstimos de bancos não

oficiais. Essa utilização crescente de recursos do exterior explica parte significativa do aumento da dívida externa brasileira no período 1967-73.

A quadruplicação dos preços do petróleo no final de 1973 estreitou as opções de crescimento do Brasil, país fortemente dependente do petróleo importado. O país respondeu positivamente aos desafios da crise do petróleo, mas cujos custos foram o retorno da inflação, a acumulação de dívida externa e a desestruturação do setor público, tanto do ponto de vista de sua capacidade financeira como do seu papel estratégico na superação das restrições ao crescimento da economia brasileira na década seguinte.

O período de crescimento da economia brasileira que se estende de meados da década de 1970 até o início dos anos 80 só pode ser entendido a partir do programa de investimentos liderado pelo estado por meio do II PND. A alternativa de crescimento só foi possível em razão da existência de financiamento externo, pois permitiu manter a taxa de câmbio apreciada, constituindo de modo implícito, um subsídio à energia e as demais matérias-primas importadas.

A elevação dos preços do petróleo e a deterioração dos termos de troca criaram um déficit substantivo na balança comercial, ao mesmo tempo em que o crescimento dos juros elevava os encargos da dívida, ampliando o déficit em transações correntes.

Apesar das condições crescentemente restritivas do ponto de vista do financiamento, o crescimento econômico foi preservado, mesmo diante da mudança de ênfase da política econômica a partir de 1976. Observa-se uma desaceleração dos investimentos a partir de 1977.

A desaceleração do investimento fica evidente pela evolução da taxa de investimento, pois, atingindo o pico de 25% do PIB em 1975/1976, houve uma contínua e progressiva queda dessa taxa. Os dados das Contas Nacionais mostram que a composição do investimento relativo a máquinas e equipamentos mantém a participação de aproximadamente 40% no total do investimento, no triênio 1974-76, declinando progressivamente até alcançar 35% em 1980.

O segundo choque do petróleo e a elevação das taxas de juros internacionais em 1979 marcam o fim do período de crescimento do PIB às taxas históricas observadas durante as décadas anteriores. Ao contrário do que se seguiu ao primeiro choque do petróleo, foram sentidos os primeiros sinais de escassez de financiamento.

A década de 1980, conhecida como a década perdida, ficou marcada pelas reduzidas taxas de crescimento, em alguns anos negativas. A adoção de políticas restritivas a partir de outubro de 1980 foi responsável pelo primeiro declínio do PIB real, no pós-guerra.

Ainda mais importante do que a retração do crescimento do produto foi a trajetória do investimento. A sua desaceleração a partir de meados dos anos 70 foi substituída pela redução absoluta e variações intensas ao longo dos anos 80. O investimento público em obras de infra-estrutura mostrou crescimento desprezível ao longo da década.

O patamar de investimentos em infra-estrutura declinou ainda mais ao longo dos anos 90. Dos três setores mais importantes, o investimento caiu sensivelmente em energia elétrica, manteve-se em transportes e cresceu apenas em telecomunicações.

Nas telecomunicações, o investimento nacional recuperou-se de 1995 em diante, tendo atingido em 1996/1997, patamares semelhantes aos dos anos 70. Ainda sob o controle da estatal Telebrás.

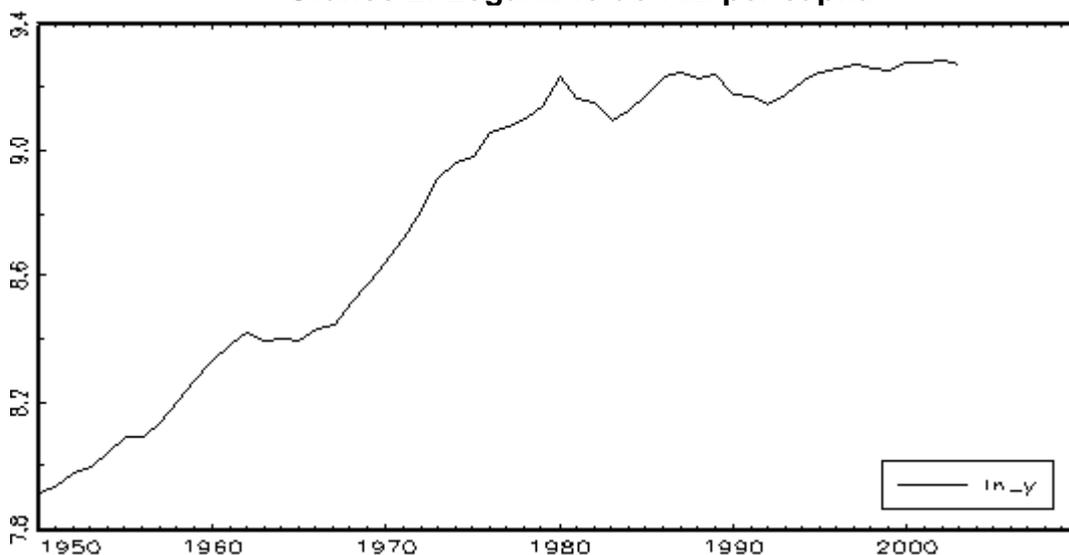
Os investimentos no setor de energia elétrica declinaram significativamente nos anos 90, representando em 1998, como porcentagem do PIB, cerca de um quarto do que foram no início da década. Em termos médios, os investimentos da década caíram para metade dos valores dos anos 80 e um terço daqueles dos anos 70. A insuficiência dos investimentos evidenciou-se em 2001 pela necessidade de estabelecer o racionamento no consumo de energia elétrica.

No que diz respeito ao setor de transportes, a taxa de investimento, embora tenha se recuperado ligeiramente, ainda está num patamar muito abaixo quando comparado aos anos 70 e mesmo aos anos 80.

5. Base de Dados

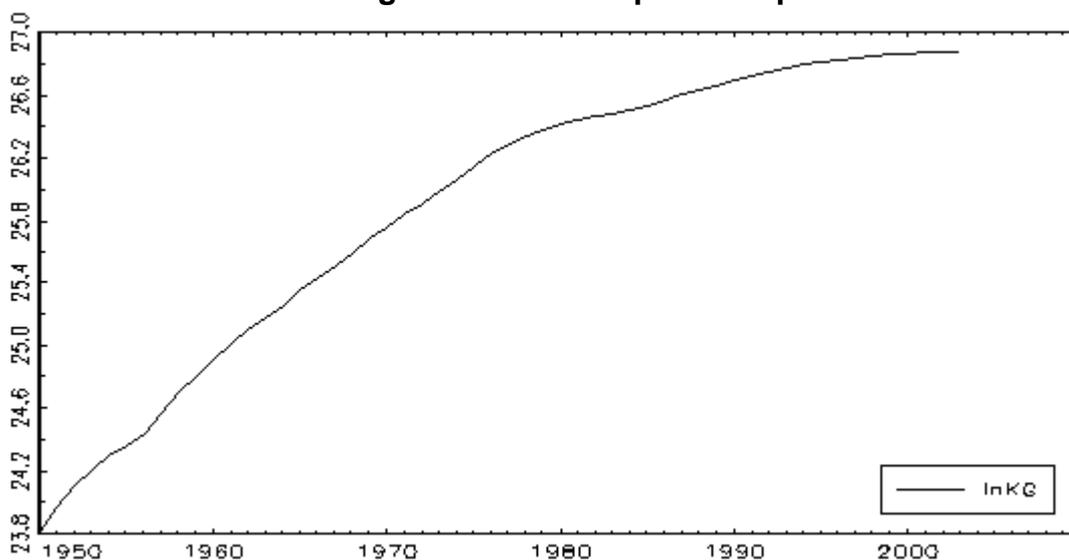
As séries utilizadas neste trabalho foram obtidas no IPEADATA. O PIB per capita é obtido utilizando-se o deflator implícito do PIB nominal e a população residente em 1º de julho.

Gráfico 2: Logaritmo do PIB per capita



A série de estoque de capital público se refere ao estoque líquido – construções – da administração pública. As séries estão avaliadas em R\$ 1999 e são determinadas para o período de 1950-2003.

Gráfico 3: Logaritmo do Estoque de Capital Público



O critério de escolhas das séries foi baseado na constatação que a maioria dos trabalhos empíricos realizados para economia brasileira, Ferreira (1996), Ferreira e Malliagos (1998), Candido Jr. (1999) e Candido Jr (2006), utilizam apenas essas duas séries. Para que os resultados obtidos neste trabalho possam ser comparados com os trabalhos anteriores optou-se, então, por utilizar somente estas.

Ferreira e Vasconcelos (2006) ao estimar a elasticidade produto-gasto público utilizando apenas as séries PIB per capita e estoque de capital público encontraram um valor acima de um. Por conseguinte, os autores incluíram as séries estoque de capital privado e capital humano para evitar um possível viés de variável omitida. A série capital privado foi extraída do IPEADATA, enquanto a série para capital humano foi obtida em Barro-Lee (2000), a amostra compreende os anos de 1960-2000.

O argumento sobre o possível viés de variável omitida pode ser levado em consideração. No entanto, os dados de capital humano, medido como anos de estudo, obtidos em Barro-Lee (2000) apresentam frequência quinquenal. A utilização de qualquer procedimento com o objetivo de mudança na frequência da série faz com o que a mesma perca suas características, tornando contestável a validade dos testes de raiz unitária e cointegração.

A utilização apenas das séries PIB per capita e estoque de capital público será realizada com base na hipótese que as estimativas tenham sido influenciada pelas mudanças estruturais observadas em ambas as séries. Os gráficos das séries indicam que ambas apresentam mudanças estruturais. Tais mudanças refletem a acentuada redução das taxas de investimento público e de crescimento do PIB per capita. Portanto, o presente trabalho tem como objetivo analisar a relação de longo prazo entre as variáveis controlando para as mudanças estruturais ocorridas.

6. Metodologia

As hipóteses do modelo clássico de regressão necessitam que as variáveis consideradas sejam estacionárias e os erros tenham média zero e variância finita. Na presença de variáveis não estacionárias a regressão pode ser espúria. Considere o modelo

$$y_t = \mu_0 + \mu_1 z_t + e_t.$$

Existem quatro casos a considerar:

Caso 1: Ambos y_t e z_t são estacionárias. Quando ambas as variáveis são estacionárias, o modelo de regressão clássico é apropriado.

Caso 2: As seqüências $\{y_t\}$ e $\{z_t\}$ são integradas de ordem diferentes. Equações de regressões usando tais variáveis não possuem significado.

Caso 3: As seqüências não estacionárias, $\{y_t\}$ e $\{z_t\}$, são integradas de mesma ordem e os resíduos contém uma tendência estocástica. Este é o caso no qual a regressão é espúria. Neste caso, é geralmente recomendado que a equação de regressão seja estimada em primeira diferença.

Caso 4: As seqüências $\{y_t\}$ e $\{z_t\}$ são não estacionárias e integradas de mesma ordem e a seqüência de resíduos é estacionária. Nessas circunstâncias, as seqüências $\{y_t\}$ e $\{z_t\}$ são cointegradas.

Uma situação de interesse especial aparece se várias variáveis apresentam uma tendência estocástica comum. Neste caso elas apresentam uma relação particularmente forte que pode ser de interesse do ponto de vista econômico. Seguindo Granger (1981) e Engle & Granger (1987), as variáveis são ditas cointegradas se elas têm uma tendência estocástica comum. Se relações de cointegração estão presentes em um sistema de variáveis, a forma VAR não é o modelo mais conveniente a ser utilizado. Neste caso é

útil considerar parametrizações específicas que suportam a análise da estrutura de cointegração. Os modelos resultantes são conhecidos como modelos vetoriais de correção de erros (VECMs).

6.1. VARs e VECMs

6.1.1. O Modelo

Para um conjunto de K variáveis de séries de tempo $y_t = (y_{1t}, \dots, y_{Kt})'$, um modelo VAR captura suas interações dinâmicas. O modelo básico de ordem p (VAR(p)) tem a forma

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t \quad (1)$$

onde os A_i 's são matrizes de coeficientes ($K \times K$) e $u_t = (u_{1t}, \dots, u_{Kt})'$ é um processo ruído branco com matriz de covariância positiva definida invariante no tempo $E(u_t u_t') = \Sigma_u$. Em outras palavras, os u_t 's são vetores estocásticos independentes $u_t \sim (0, \Sigma_u)$. O processo é estável se

$$\det(I_K - A_1 z - \dots - A_p z^p) \neq 0 \text{ para } |z| \leq 1, \quad (2)$$

isto é, o polinômio definido pelo determinante do operador autorregressivo não possui raízes dentro círculo unitário. Se o polinômio possui uma raiz unitária (i.e., o determinante é zero para $z = 1$), então algumas ou todas as variáveis são integradas. Assumindo que elas sejam no máximo $I(1)$. Se as variáveis possuem uma tendência estocástica comum, é possível que existam combinações lineares delas que sejam $I(0)$.

Embora o modelo (1) seja suficientemente geral para acomodar variáveis com tendências estocásticas, ele não é o mais apropriado para analisar as relações de cointegração dado que estas não aparecem explicitamente. A forma VECM

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t \quad (3)$$

é mais conveniente para análise de cointegração. Aqui $\Pi = -(I_K - A_1 - \dots - A_p)$ e $\Gamma_i = -(A_{i+1} + \dots + A_p)$ para $i = 1, \dots, p-1$. O VECM é obtido da forma VAR em níveis (1) subtraindo y_{t-1} de ambos os lados e rearranjando os termos. Como Δy_t não contém tendências estocásticas pela hipótese que todas as variáveis podem ser no máximo I (1), o termo Πy_{t-1} é o único que inclui variáveis I (1). Por conseguinte, Πy_{t-1} deve ser I (0) também. Então, ele contém as relações de cointegração. Os Γ_j ($j = 1, \dots, p-1$) são geralmente conhecidos como os parâmetros de curto prazo, e Πy_{t-1} é conhecido como o termo de longo prazo. Também é possível determinar as matrizes dos parâmetros em nível A_j a partir dos coeficientes do VECM. Mais precisamente, $A_1 = \Gamma_1 + \Pi + I_K$, $A_i = \Gamma_i - \Gamma_{i-1}$ para $i = 1, \dots, p-1$ e $A_p = -\Gamma_{p-1}$.

Se o processo VAR(p) possui raízes unitárias, isto é, $\det(I_K - A_1 - \dots - A_p) = 0$ para $z = 1$, a matriz $\Pi = -(I_K - A_1 - \dots - A_p)$ é singular. Suponha que $rk(\Pi) = r$. Então Π pode ser escrita como o produto de ($K \times r$) matrizes α e β com $rk(\alpha) = rk(\beta) = r$ como segue: $\Pi = \alpha\beta'$. Pré-multiplicando um vetor $I(0)$ por alguma matriz o resultado é novamente um processo $I(0)$. Então, $\beta' y_{t-1}$ é $I(0)$ porque ele pode ser obtido pré-multiplicando $\Pi y_{t-1} = \alpha\beta' y_{t-1}$ por $(\alpha'\alpha)^{-1}\alpha'$. Por conseguinte, $\beta' y_{t-1}$ contém as relações de cointegração. Segue que existem $r = rk(\Pi)$ relações de cointegração linearmente independentes entre os componentes de y_t . O rank de Π é, portanto, o rank de cointegração do sistema e β é a matriz de cointegração.

6.1.2. Termos Determinísticos

Uma maneira de incluir termos determinísticos é simplesmente adicioná-los a parte estocástica,

$$y_t = \mu_t + x_t, \quad (4)$$

onde μ_t é a parte determinística e x_t é um processo estocástico que pode ter uma representação VAR ou VECM. Sob a hipótese que μ_t é um termo de tendência linear, isto é, $\mu_t = \mu_0 + \mu_1 t$, a representação VAR(p) de y_t é dada por

$$y_t = v_0 + v_1 t + A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t. \quad (5)$$

Esta representação é facilmente derivada pré-multiplicando (4) por $A(L) = I_K - A_1 L - \dots - A_p L^p$, onde L é o operador de defasagens. Note que $A(L)x_t = u_t$.

A representação VECM(p-1) equivalente tem a forma

$$\Delta y_t = v_0 + v_1 t + \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t. \quad (6)$$

Restrições sobre v_0 e v_1 permite a absorção da parte determinística nas relações de cointegração.

6.2. Estimação

6.2.1. Estimação de um VAR sem Restrição

Dada a amostra y_1, \dots, y_T e os valores pré-amostrais y_{-p+1}, \dots, y_0 , as K equações do VAR podem ser estimadas separadamente por Mínimos Quadrados Ordinários (MQO). O estimador resultante tem a mesma eficiência que o estimador de Mínimos Quadrados Generalizados. Utilizando a notação $Y = [y_1, \dots, y_T]$, $A = [A_1 : \dots : A_p]$, $U = [u_1, \dots, u_T]$ e $Z = [Z_0, \dots, Z_{T-1}]$, onde

$$Z_{t-1} = \begin{bmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{bmatrix}$$

O modelo (1) pode ser escrito como

$$Y = AZ + U \quad (7)$$

e o estimador de MQO de A é

$$\hat{A} = [\hat{A}_1 : \dots : \hat{A}_p] = YZ'(ZZ')^{-1} \quad (8)$$

Sob as hipóteses padrões, o estimador de MQO é consistente e assintoticamente normalmente distribuído,

$$\sqrt{T} \text{vec}(\hat{A} - A) \xrightarrow{d} N(0, \Sigma_{\hat{A}}). \quad (9)$$

$$\text{vec}(\hat{A}) \overset{a}{\sim} N(\text{vec}(A), \Sigma_{\hat{A}}/T) \quad (10)$$

A matriz de covariância da distribuição assintótica é $\Sigma_{\hat{A}} = p \lim(ZZ'/T)^{-1} \otimes \Sigma_u$, então

$$\text{vec}(\hat{A}) \overset{a}{\sim} N(\text{vec}(A), (ZZ')^{-1} \otimes \Sigma_u). \quad (11)$$

Para um processo I(0) normalmente distribuído y_t , o estimador de MQO é idêntico ao estimador de Máxima Verossimilhança (MV) condicional aos valores iniciais.

O estimador de MQO também possui a distribuição assintótica mostrada em (9) para sistemas não estacionários com variáveis integradas. Contudo, neste caso a matriz de covariância $\Sigma_{\hat{A}}$ é singular. Em outras palavras, se existem variáveis integradas ou cointegradas, alguns coeficientes ou combinações lineares de coeficientes estimados

convergem a uma taxa maior que $T^{1/2}$. Portanto, os testes t , χ^2 , e F usuais para inferência sobre os parâmetros podem não ser válidos neste caso. Como um exemplo considere o processo autorregressivo univariado $y_t = \rho y_{t-1} + u_t$. Se y_t é I (1) e, por conseguinte, $\rho = 1$, o estimador de MQO de ρ não tem uma distribuição limite padrão. A quantidade $\sqrt{T}(\hat{\rho} - \rho)$ converge em probabilidade para zero.

A matriz de covariância Σ_u pode ser estimada de maneira usual. Denote por \hat{u}_t os resíduos de MQO, isto é, $\hat{u}_t = y_t - \hat{A}Z_{t-1}$, as matrizes

$$\hat{\Sigma}_u = \frac{1}{T - Kp} \sum_{t=1}^T \hat{u}_t \hat{u}_t' \quad \text{e} \quad \tilde{\Sigma}_u = \frac{1}{T} \sum_{t=1}^T \hat{u}_t \hat{u}_t' \quad (12)$$

são estimadores possíveis. Ambos os estimadores são consistentes e assintoticamente normalmente distribuídos independentemente de \hat{A} , isto é, $\sqrt{T}(\hat{\Sigma}_u - \Sigma_u)$ e $\sqrt{T}(\tilde{\Sigma}_u - \Sigma_u)$ têm distribuição assintótica normal se condições de momento suficientes são impostas.

6.2.2. Estimação de VECMs

Estimador de Máxima Verossimilhança

Se o rank de cointegração do sistema sob consideração é conhecido, trabalhar com a forma VECM (3) é conveniente para impor uma restrição correspondente. Ao derivar os estimadores dos parâmetros de (3), a seguinte notação será utilizada: $\Delta Y = [\Delta y_1, \dots, \Delta y_T]$, $Y_{-1} = [y_0, \dots, y_{T-1}]$, $U = [u_1, \dots, u_T]$, $\Gamma = [\Gamma_1 : \dots : \Gamma_{p-1}]$ e $X = [X_0, \dots, X_{T-1}]$ com

$$X_{t-1} = \begin{bmatrix} \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p+1} \end{bmatrix}$$

Para uma amostra com T observações e p valores pré-amostrais, o VECM (3) pode ser escrito compactamente como

$$\Delta Y = \Pi Y_{-1} + \Gamma X + U. \quad (13)$$

Dada uma matriz específica Π , o estimador de MQO de Γ pode ser obtido como

$$\hat{\Gamma} = (\Delta Y - \Pi Y_{-1}) X' (X X')^{-1}. \quad (14)$$

Substituindo (14) em (13) e rearranjando os termos, temos

$$\Delta Y M = \Pi Y_{-1} M + \hat{U}, \quad (15)$$

onde $M = I - X'(X X')^{-1} X$. Para um dado inteiro r , $0 < r < K$, um estimador $\hat{\Pi}$ de Π com $rk(\hat{\Pi}) = r$ pode ser obtido pelo método conhecido como análise de correlação canônica ou, equivalentemente, por uma regressão de rank reduzido baseada no modelo (15). Seguindo Johansen (1995), a estimação pode ser realizada definindo

$$S_{00} = T^{-1} \Delta Y M \Delta Y', \quad S_{01} = T^{-1} \Delta Y M Y_{-1}', \quad S_{11} = T^{-1} Y_{-1} M Y_{-1}'$$

e resolvendo o problema dos autovalores generalizados

$$\det(\lambda S_{11} - S_{01}' S_{00}^{-1} S_{01}) = 0 \quad (16)$$

Seja os autovalores ordenados por $\lambda_1 \geq \dots \geq \lambda_r$ com a matriz correspondente de autovetores $V = [b_1, \dots, b_r]$ satisfazendo $\lambda_i S_{11} b_i = S_{01}' S_{00}^{-1} S_{01} b_i$ e normalizado tal que $V' S_{11} V = I_r$. O estimador de rank reduzido de $\Pi = \alpha \beta'$ é então obtido escolhendo

$$\hat{\beta} = [b_1, \dots, b_r]$$

e

$$\hat{\alpha} = \Delta Y M Y_{-1}' \hat{\beta} (\hat{\beta}' Y_{-1} M Y_{-1}' \hat{\beta})^{-1}, \quad (17)$$

isto é, $\hat{\alpha}$ pode ser visto como o estimador de MQO do modelo

$$\Delta Y M = \alpha \hat{\beta}' Y_{-1} M + \tilde{U}.$$

O correspondente estimador de Π é $\hat{\Pi} = \hat{\alpha} \hat{\beta}'$. O estimador factível de Γ é $\hat{\Gamma} = (\Delta Y - \hat{\Pi} Y_{-1}) X' (X X')^{-1}$. Sob a hipótese Gaussiana esses estimadores são estimadores de MV condicional aos valores pré-amostrais. Eles são consistentes e conjuntamente assintoticamente normais sob as hipóteses gerais,

$$\sqrt{T} \text{vec}([\hat{\Gamma}_1 : \dots : \hat{\Gamma}_{p-1}] - [\Gamma_1 : \dots : \Gamma_{p-1}]) \xrightarrow{d} N(0, \Sigma_{\hat{\Gamma}})$$

e

$$\sqrt{T} \text{vec}(\hat{\Pi} - \Pi) \xrightarrow{d} N(0, \Sigma_{\hat{\Pi}}).$$

A distribuição assintótica de $\hat{\Gamma}$ é não singular, portanto, a inferência padrão pode ser usada para os parâmetros de curto prazo. Por outro lado, a matriz de covariância $\Sigma_{\hat{\Gamma}}$ tem rank Kr e é, portanto, singular se $K < r$.

Nesta abordagem o estimador $\hat{\beta}$ torna-se único pela normalização dos autovetores, e $\hat{\alpha}$ é ajustado de acordo. Contudo, essas não são restrições de identificação. Para estimar as matrizes α e β consistentemente, é necessário impor restrições de identificação. Sem tais restrições somente o produto $\alpha \beta' = \Pi$ pode ser estimado consistentemente. Um exemplo de restrição que tem recebido alguma atenção na literatura assume que a primeira parte de β é uma matriz identidade, isto é, $\beta' = [I_K : \beta'_{(K-r)}]$. Para $r=1$, esta restrição equivale a normalização do coeficiente da primeira variável para 1. Esta normalização requer algum cuidado ao escolher a ordem das variáveis. A razão é que pode existir uma relação de cointegração somente entre um

subconjunto de variáveis do sistema. Portanto, normalizar um coeficiente arbitrário pode resultar na divisão de uma estimativa por um parâmetro correspondente que é atualmente zero porque a variável associada não pertence à relação de cointegração. Se o rank de cointegração $r > 1$, a normalização é tal que

$$\beta = \begin{bmatrix} I_r \\ \beta_{(K-r)} \end{bmatrix}$$

Dado que $rk(\beta) = r$, existe uma submatriz não singular ($r \times r$) de β' que motiva a normalização. Note que $\Pi = \alpha\beta' = \alpha BB^{-1}\beta'$ para qualquer matriz ($r \times r$) não singular B. Por conseguinte, escolhendo B tal que ela corresponda a uma submatriz de β' resulta na decomposição de Π , em que β contém uma submatriz identidade.

A normalização garante a identificação dos parâmetros $\beta_{(K-r)}$ e, dessa forma, a inferência torna-se possível. Se restrições de identificações são impostas $T(\hat{\beta} - \beta)$ e $\sqrt{T}(\hat{\alpha} - \alpha)$ convergem em distribuição. Por conseguinte, o estimador de β converge a uma taxa mais rápida T e, portanto, é dito ser superconsistente. O estimador de α converge a taxa usual \sqrt{T} e tem distribuição assintótica normal.

Os estimadores dos parâmetros $\beta_{(K-r)}$ têm distribuição assintótica normal multivariada sob normalização apropriada. Mais precisamente, particionando y_t como

$$y_t = \begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \end{bmatrix},$$

onde $y_t^{(1)}$ e $y_t^{(2)}$ são ($r \times 1$) e $((K - r) \times 1)$, respectivamente, e definindo $Y_{-1}^{(2)} = [y_0^{(2)}, \dots, y_{T-1}^{(2)}]$, temos

$$\text{vec}\{(\hat{\beta}'_{(K-r)} - \beta'_{(K-r)})(Y_{-1}^{(2)}MY_{-1}^{(2)'})^{1/2}\}$$

$$= \left[(Y_{-1}^{(2)} M Y_{-1}^{(2)'})^{1/2} \otimes I_{K-r} \right] \text{vec}(\hat{\beta}'_{(K-r)} - \beta'_{(K-r)}) \\ \xrightarrow{d} N(0, I_{(K-r)} \otimes (\alpha \Sigma_u^{-1} \alpha)^{-1}).$$

Este resultado pode ser escrito como

$$\text{vec}(\hat{\beta}'_{(K-r)}) \approx N(\text{vec}(\beta'_{(K-r)}), (Y_{-1}^{(2)} M Y_{-1}^{(2)'})^{-1} \otimes (\alpha \Sigma_u^{-1} \alpha)^{-1}). \quad (18)$$

Repassando os parâmetros desconhecidos na matriz de covariância por seus estimadores de MV, as razões t são obtidas simplesmente dividindo os elementos de $\hat{\beta}'_{(K-r)}$ pela raiz quadrada dos respectivos elementos da diagonal de

$$\hat{\Omega} = (Y_{-1}^{(2)} M Y_{-1}^{(2)'})^{-1} \otimes (\hat{\alpha} \tilde{\Sigma}_u^{-1} \hat{\alpha})^{-1}. \quad (19)$$

A expressão em (18) também pode ser usada para testar as hipóteses sobre $\beta'_{(K-r)}$. Para uma dada matriz $R_{(J \times r(K-r))}$ de rank J e um vetor $r_{(J \times 1)}$, a hipótese

$$H_0 : R \text{vec}(\beta'_{(K-r)}) = r \quad (20)$$

pode ser checada por um teste de Wald. O teste estatístico tem distribuição assintótica χ^2 com J graus de liberdade sob H_0 .

$$\lambda_W = [R \text{vec}(\hat{\beta}'_{(K-r)}) - r]' (R \hat{\Omega} R')^{-1} [R \text{vec}(\hat{\beta}'_{(K-r)}) - r] \xrightarrow{d} \chi^2_{(J)}.$$

Note que o estimador dos parâmetros de um VAR em nível \hat{A} pode ser computado via estimativas de estimativas de Π e Γ . Este estimador tem a vantagem de impor restrições de cointegração na versão em nível do VAR estimado. O estimador da matriz de covariância $\tilde{\Sigma}_u$ a partir dos resíduos do VECM estimado possui a mesma distribuição assintótica do estimador obtido a partir dos resíduos de um VAR em nível.

Estimador de Dois Estágios para a Matriz de Cointegração

Outro método para estimar a matriz de cointegração utiliza o fato que o VECM (3) pode ser escrito da seguinte forma

$$\Delta y_t - \Pi_1 y_{t-1}^{(1)} - \Gamma X_{t-1} = \Pi_2 y_{t-1}^{(2)} + u_t, \quad (21)$$

onde $y_{t-1}^{(1)}$ é $(r \times 1)$, $y_{t-1}^{(2)}$ é $((K - r) \times 1)$, $\Gamma = [\Gamma_1 : \dots : \Gamma_{p-1}]$ e $X_{t-1} = [\Delta y'_{t-1}, \dots, \Delta y'_{t-p+1}]$. As matrizes Π_1 e Π_2 são $(K \times r)$ e $(K \times (K - r))$, respectivamente, tal que $[\Pi_1 : \Pi_2] = \Pi = \alpha \beta'$. Normalizando a matriz de cointegração tal que $\beta' = [I_K : \beta'_{(K-r)}]$, temos $\Pi_1 = \alpha$ e $\Pi_2 = \alpha \beta'_{(K-r)}$. Pré-multiplicando (21) por $(\alpha' \Sigma_u^{-1} \alpha)^{-1} \alpha' \Sigma_u^{-1}$ e definindo

$$w_t = (\alpha' \Sigma_u^{-1} \alpha)^{-1} \alpha' \Sigma_u^{-1} (\Delta y_t - \alpha y_{t-1}^{(1)} - \Gamma X_{t-1}),$$

temos

$$w_t = \beta'_{(K-r)} y_{t-1}^{(2)} + v_t, \quad (22)$$

onde $v_t = (\alpha' \Sigma_u^{-1} \alpha)^{-1} \alpha' \Sigma_u^{-1} u_t$ é um processo ruído branco com média zero e matriz de covariância $\Sigma_v = (\alpha' \Sigma_u^{-1} \alpha)^{-1}$.

Se α , Γ e Σ_u são dados, $\beta'_{(K-r)}$ pode ser estimada de (21) por MQO. Como na prática os parâmetros são desconhecidos, o seguinte procedimento em dois estágios pode ser usado. No primeiro passo, elimina-se os parâmetros de curto prazo repassando-os por seus estimadores de MQO dado Π , o modelo considerado toma a forma

$$\Delta Y M = \Pi Y_{-1} M + \hat{U}. \quad (23)$$

Usando este modelo para estimar Π por MQO. Denotando o estimador por $\tilde{\Pi} = [\tilde{\Pi}_1 : \tilde{\Pi}_2]$ e o estimador correspondente da matriz da covariância dos resíduos por $\tilde{\Sigma}_u = T^{-1} (\Delta Y - \tilde{\Pi} Y_{-1}) M (\Delta Y - \tilde{\Pi} Y_{-1})'$ e usando $\tilde{\alpha} = \tilde{\Pi}_1$, temos

$$\tilde{W} = (\tilde{\alpha}' \tilde{\Sigma}_u^{-1} \tilde{\alpha})^{-1} \tilde{\alpha}' \tilde{\Sigma}_u^{-1} (\Delta Y - \tilde{\alpha} - Y_{-1}^{(1)}),$$

onde $Y_{-1}^{(1)} = [y_0^{(1)}, \dots, y_{T-1}^{(1)}]$. O segundo passo consiste em estimar $\beta'_{(K-r)}$ por MQO de

$$\tilde{W}M = \beta'_{(K-r)} Y_{-1}^{(2)}M + \tilde{V}, \quad (24)$$

isto é,

$$\tilde{\beta}'_{(K-r)} = \tilde{W}MY_{-1}^{(2)}(Y_{-1}^{(2)}MY_{-1}^{(2)'})^{-1}, \quad (25)$$

onde $Y_{-1}^{(2)} = [y_0^{(2)}, \dots, y_{T-1}^{(2)}]$.

Este estimador de dois estágios tem a mesma distribuição assintótica que o estimador de MV, $vec(\tilde{\beta}'_{(K-r)})$ tem distribuição assintótica aproximada normal, $N(vec(\beta'_{(K-r)}), \Omega)$, onde de (25) um possível estimador da matriz de covariância é

$$\tilde{\Omega} = (Y_{-1}^{(2)}MY_{-1}^{(2)'})^{-1} \otimes \tilde{\Sigma}_v.$$

Aqui $\tilde{\Sigma}_v = T^{-1}(\tilde{W} - \tilde{\beta}'_{(K-r)} Y_{-1}^{(2)})M(\tilde{W} - \tilde{\beta}'_{(K-r)} Y_{-1}^{(2)})'$ é o estimador da matriz de covariância dos resíduos a partir de (25).

O estimador de dois estágios pode ser usado na maneira usual para obter estimadores dos outros parâmetros no modelo empregando (17) e as fórmulas seguindo aquela equação. O estimador pode ser estendido para o caso no quais termos determinísticos estejam presentes no modelo.

6.3. Estacionariedade

Um processo y_t é estacionário se seus primeiro e segundo momentos não variam ao longo do tempo. Em outras palavras, y_t é estacionário se

1. $E(y_t) = \mu_y$ para todo $t \in T$ e
2. $E[(y_t - \mu_y)(y_{t-h} - \mu_y)] = \gamma_h$ para todo $t \in T$ e todo inteiro h tal que $t - h \in T$.

A primeira condição significa que todos os membros de um processo estacionário possuem a mesma média constante. Por conseguinte, uma série de tempo gerada por um processo estocástico estacionário deve flutuar em torno de uma média constante. A segunda condição garante que as variâncias também não variam ao longo do tempo, para $h=0$, a variância $\sigma_y^2 = E[(y_t - \mu_y)^2] = \gamma_0$ não depende de t . Além disso, as covariâncias $E[(y_t - \mu_y)(y_{t-h} - \mu_y)] = \gamma_h$ não dependem de t mas apenas da distancia no tempo h entre os dois membros do processo.

Alguns processos são chamados de tendência estacionários se eles podem se tornar estacionários subtraindo-se uma função tendência determinística tal como uma função linear da forma $\mu_0 + \mu_1 t$, onde μ_0 e μ_1 são parâmetros fixos.

6.3.1. Autocorrelações Amostrais e Autocorrelações Parciais

As autocorrelações amostrais, $\tilde{\rho}_h = \tilde{\gamma}_h / \tilde{\gamma}_0$ ou $\hat{\rho}_h = \hat{\gamma}_h / \hat{\gamma}_0$, são obtidas a partir de

$$\tilde{\gamma}_h = \frac{1}{T} \sum_{t=h+1}^T (y_t - \bar{y})(y_{t-h} - \bar{y})$$

ou

$$\hat{\gamma}_h = \frac{1}{T-h} \sum_{t=h+1}^T (y_t - \bar{y})(y_{t-h} - \bar{y}),$$

onde $\bar{y} = T^{-1} \sum_{t=1}^T y_t$ é a média amostral. Para séries com processos de geração de dados estacionários, as autocorrelações amostrais tipicamente desaparecem rapidamente quando h aumenta.

As autocorrelações amostrais são estimadas das atuais autocorrelações se o processo é estacionário. Se ele é um passeio aleatório, isto é, se todos os membros são

mutuamente independentes e identicamente distribuídos, dessa forma, y_t e y_{t-h} são estocasticamente independentes para $h \neq 0$, então as autocorrelações estimadas têm distribuição assintótica normal padrão, $\sqrt{T}\tilde{\rho}_h \xrightarrow{d} N(0,1)$, e $\tilde{\rho}_h \approx N(0,1/T)$. Por conseguinte, $[-2/\sqrt{T}, 2/\sqrt{T}]$ é um intervalo de confiança aproximado de 95% em torno de zero.

As autocorrelações parciais também são quantidades que podem conter informações úteis sobre as propriedades do processo de geração de dados de uma dada série de tempo. A autocorrelação parcial entre y_t e y_{t-h} é a autocorrelação condicional dado $y_{t-1}, \dots, y_{t-h+1}$. Formalmente,

$$a_h = \text{Corr}(y_t, y_{t-h} \mid y_{t-1}, \dots, y_{t-h+1}).$$

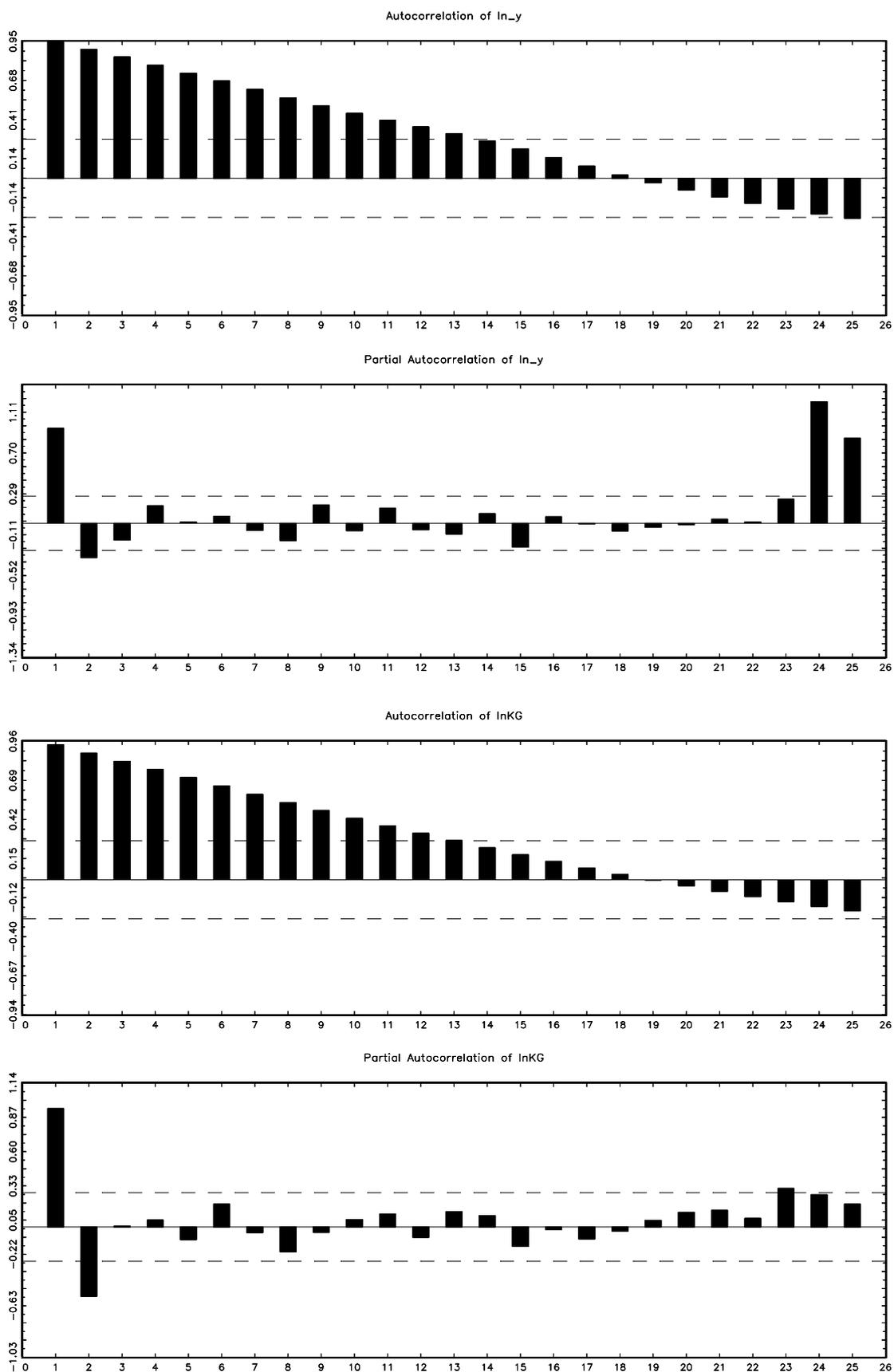
A correspondente quantidade amostral \hat{a}_h é obtida como o estimador de Mínimos Quadrados Ordinários (MQO) do coeficiente α_h em um modelo autorregressivo

$$y_t = \nu + \alpha_1 y_{t-1} + \dots + \alpha_h y_{t-h} + u_t.$$

Para processos estacionários, as autocorrelações parciais também se aproximam de zero quando h tende ao infinito.

A figura 1 mostra os gráficos das funções de autocorrelações e autocorrelações parciais dos logaritmos das séries PIB per capita (y) e estoque de capital público (KG). As linhas pontilhadas na figura são limites $\pm 2/\sqrt{T}$ em torno de zero que podem ser usadas para avaliar se as quantidades estimadas são próximas de zero.

Figura 1- Funções de Autocorrelações Amostrais e Autocorrelações Parciais



6.4. Testes de Raiz Unitária

Os gráficos mostrados na figura 1 sugerem que as séries são não estacionárias. No entanto, existem testes estatísticos formais para estacionariedade. O teste Dickey-Fuller Aumentado (ADF) checka a hipótese nula que existe uma raiz unitária contra a hipótese alternativa de estacionariedade do processo de geração de dados, que pode ter uma média diferente de zero e uma tendência linear determinística. O segundo tipo de teste permite, em adição, considerar mudanças estruturais quando estas são observadas.

6.4.1. O Teste ADF

Se o processo de geração de dados é um $AR(p)$, o processo é integrado quando $\alpha(1) = 1 - \alpha_1 - \dots - \alpha_p = 0$. Em outras palavras, a hipótese de interesse a ser testada é $\alpha(1) = 0$. Para testar esta hipótese nula contra a hipótese alternativa de estacionariedade do processo, o teste ADF é baseado em modelos da forma

$$\Delta y_t = \phi y_{t-1} + \sum_{j=1}^{p-1} \alpha_j^* \Delta y_{t-j} + u_t, \quad (26)$$

onde $\phi = -\alpha(1)$ e $\alpha_j^* = -(\alpha_{j+1} + \dots + \alpha_p)$.

Neste modelo o par de hipóteses, $H_0 : \phi = 0$ versus $H_1 : \phi < 0$, é testada com base na estatística t do coeficiente ϕ a partir da estimação de (26) por MQO [Fuller (1976) e Dickey-Fuller (1979)]. H_0 é rejeitada se a estatística t é menor que o valor crítico relevante. Sob H_0 a série possui uma raiz unitária e é não estacionária, enquanto ela é considerada estacionária se a hipótese nula é rejeitada.

O teste estatístico possui uma distribuição limite não-padrão. Os valores críticos são obtidos por simulação e estão disponíveis em Fuller (1976) e Davidson e MacKinnon (1993). A distribuição limite depende dos termos determinísticos que tenham sido incluídos. Portanto, valores críticos diferentes são usados quando uma constante ou um termo de tendência linear é adicionado ao modelo.

Tabela 1 – Teste de Raiz Unitária Dickey-Fuller Aumentado (ADF)

Variáveis	Nível		1ª Diferença	
	$t_{(\mu)}$	$t_{(t)}$	$t_{(\mu)}$	$t_{(t)}$
ln_y	-2.2225	-0.6975	-4.4651	-5.0599
ln KG	-3.2107	-0.2027	-2.1957	-4.7181
Valores Críticos				
1%	-3.43	-3.96	-3.43	-3.96
5%	-2.86	-3.41	-2.86	-3.41
10%	-2.57	-3.13	-2.57	-3.13

Fonte: Elaboração própria

Observações: As estatísticas $t_{(\mu)}$ e $t_{(t)}$ se referem aos modelos com termo constante e com termo constante e tendência respectivamente.

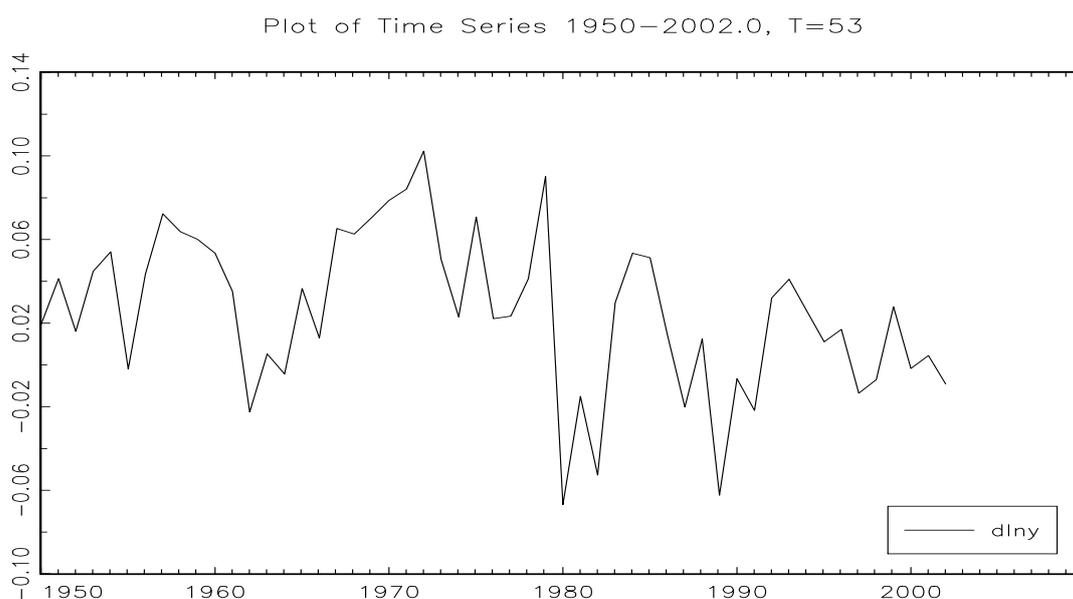
Utilizou-se uma estrutura de defasagens igual a um para as séries em nível e zero para as séries em primeira diferença.

Os valores críticos foram obtidos em Davidson e MacKinnon (1993).

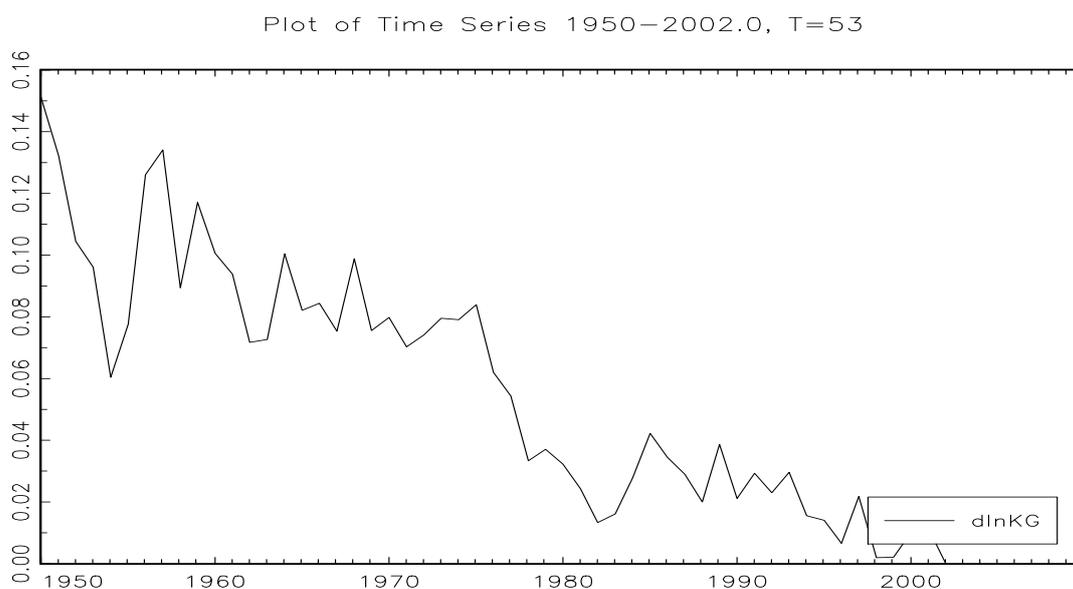
Os resultados do teste de Dickey-Fuller Aumentado (ADF) estão apresentados na Tabela 1. Eles indicam que não se rejeita a presença de uma raiz unitária para nenhuma das variáveis consideradas; isto é, as variáveis do modelo são não estacionárias.

O teste ADF para as variáveis logaritmo do PIB (ln_y) e logaritmo do estoque de capital público (ln KG) indicam que a hipótese nula da presença de raiz unitária nas séries é significativa com o número de defasagens igual a um, de acordo com os critérios de Hannan-Quin e Schwarz, incluindo um intercepto e um termo de tendência linear.

Dado que todas as variáveis do modelo exibiram raiz unitária em nível, passa-se, então, à definição da ordem de integração de cada variável. Para isso, aplica-se o teste ADF nas séries em primeira diferença. Nesse caso, se a hipótese nula for rejeitada então a variável é integrada de ordem um, $I(1)$. Caso contrário, diferencia-se novamente e aplica-se o teste até que a hipótese nula seja rejeitada. Note que, neste procedimento, se um termo de tendência linear é necessário para testar y_t , somente uma constante será usada no teste para Δy_t porque, se $y_t = \mu_0 + \mu_1 t + x_t$, então $\Delta y_t = \mu_1 + \Delta x_t$.

Gráfico 4: Logaritmo do PIB per capita em primeira diferença

Fonte: Elaboração própria

Gráfico 5: Logaritmo do Estoque de Capital Público em primeira diferença

Fonte: Elaboração própria

O resultado do teste, apresentado na Tabela 1, mostra que a variável \ln_y é $I(1)$, enquanto não se pode rejeitar a hipótese de raiz unitária na série $\Delta \ln KG$. Os gráficos 2 e 3 indicam que termos de tendência linear estão presentes nas séries \ln_y e $\ln KG$. Portanto, era de se esperar que as séries em primeira diferença não apresentassem tendência. O termo de tendência deveria desaparecer quando as séries

fossem diferenciadas. No entanto, o gráfico da série $\Delta \ln KG$ parece apresentar uma tendência declinante ao longo do tempo. Note que como as séries PIB per capita e estoque de capital público estão em forma de logaritmo, as séries em primeira diferença representam as taxas de crescimento.

Os dois níveis da taxa de crescimento do estoque de capital público podem sugerir uma mudança estrutural nesta série. Ao realizar testes de raiz unitária, um cuidado especial deve ser tomado se quebra estruturais tenham ocorrido. Quando existem quebras estruturais, os vários testes estatísticos de Dickey-Fuller são viesados em direção da não rejeição de uma raiz unitária. Este viés significa que o teste de Dickey-Fuller é viesado em direção da hipótese nula de uma raiz unitária, mesmo se a série é estacionária dentro de cada subperíodo.

6.4.2. Teste para Processos com Mudança de Nível

Se existe uma mudança de nível no processo de geração de dados, ela deve ser levada em consideração na realização do teste de raiz unitária porque o teste ADF tem baixo poder se a mudança é simplesmente ignorada. Uma abordagem possível é assumir que mudança é determinística. Neste caso, uma “função degrau”, que será denotada por $f_t(\theta)' \gamma$, pode ser adicionada ao modelo.

$$y_t = \mu_0 + \mu_1 t + f_t(\theta)' \gamma + x_t,$$

onde θ e γ são parâmetros desconhecidos e os erros x_t são gerados por um processo $AR(p)$, $\alpha^*(L)(1-\rho L)x_t = u_t$ com $\alpha^*(L) = 1 - \alpha_1^* L - \dots - \alpha_{p-1}^* L^{p-1}$. Funções degrau podem, por exemplo, ser baseadas em

$$f_t^{(1)} = d_{1t} = \begin{cases} 0, & t < T_B \\ 1, & t \geq T_B \end{cases}$$

$$f_t^{(2)}(\theta) = \begin{cases} 0, & t < T_B \\ 1 - \exp\{-\theta(t - T_B + 1)\}, & t \geq T_B \end{cases}$$

ou

$$f_t^{(3)}(\theta) = \left[\frac{d_{1,t}}{1 - \theta L}, \frac{d_{1,t-1}}{1 - \theta L} \right].$$

A primeira dessas funções é simplesmente uma variável *dummy* com mudança na data T_B . Esta função não envolve qualquer parâmetro extra θ . No termo $f_t \gamma$, o parâmetro γ é um escalar. A segunda função é baseada na distribuição exponencial, que permite uma mudança não linear gradual para o novo nível na data T_B . No termo $f_t^{(2)}(\theta) \gamma$, ambos θ e γ são escalares. A primeira é necessariamente positiva, enquanto a segunda pode assumir qualquer valor. Finalmente, a terceira função pode ser vista como uma função racional no operador de defasagens aplicado a uma variável *dummy* d_{1t} . O atual termo de mudança é $[\gamma_1(1 - \theta L)^{-1} + \gamma_2(1 - \theta L)^{-1} L] d_{1t}$. Aqui θ é um escalar entre 0 e 1 e $\gamma = (\gamma_1, \gamma_2)'$. Uma maneira alternativa para escrever esta função é

$$f_t^{(3)}(\theta)' \gamma = \begin{cases} 0, & t < T_B \\ \gamma_1, & t = T_B \\ \gamma_1 + \sum_{j=1}^{t-T_B} \theta^{j-1} (\theta \gamma_1 + \gamma_2), & t > T_B \end{cases}$$

Esta expressão oferece a possibilidade de mudanças não lineares mais gerais.

Se uma tendência linear esta presente no modelo, os parâmetros relevantes $\eta = (\mu_0, \mu_1, \gamma)'$ são estimados pela minimização da soma dos quadrados dos erros em primeira diferença,

$$\Delta y_t = \mu_1 + f_t(\theta)' \gamma + v_t \quad (t = 2, \dots, T),$$

onde $v_t = \alpha^*(L)^{-1} u_t$. Em outras palavras, a estimação é realizada sob a hipótese nula de raiz unitária minimizando

$$Q_p(\eta, \theta, \alpha^*) = (Y - Z(\theta)\eta)' \sum (\alpha^*)^{-1} (Y - Z(\theta)\eta),$$

onde α^* é um vetor de coeficientes $\alpha^*(L)$, $\sum (\alpha^*) = Cov(V) / \sigma_u^2$, $V = (v_1, \dots, v_T)'$ é o vetor de erros do modelo, $Y = [y_1, \Delta y_2, \dots, \Delta y_T]'$ e $Z = [Z_1 : Z_2 : Z_3]$ com $Z_1 = [1, 0, \dots, 0]'$, $Z_2 = [1, 1, \dots, 1]'$ e $Z_3 = [f_1(\theta), \Delta f_2(\theta), \dots, \Delta f_T(\theta)]'$.

Embora, as séries ajustadas $\hat{x}_t = y_t - \hat{\mu}_0 - \hat{\mu}_1 t - f_t(\hat{\theta})' \hat{\gamma}$ possam ser usadas na abordagem do teste ADF, Lanne et al. (2002) propôs um procedimento diferente que ajusta os parâmetros aos erros estimados e que funciona melhor em pequenas amostras. Denotando o estimador do polinômio AR obtido pela minimização de $Q_p(\eta, \theta, \alpha^*)$ por $\hat{\alpha}^*(L)$, Lanne et al. (2002) definiu $\hat{w}_t = \hat{\alpha}^*(L)\hat{x}_t$ e baseou o teste de raiz unitária em um modelo de regressão auxiliar

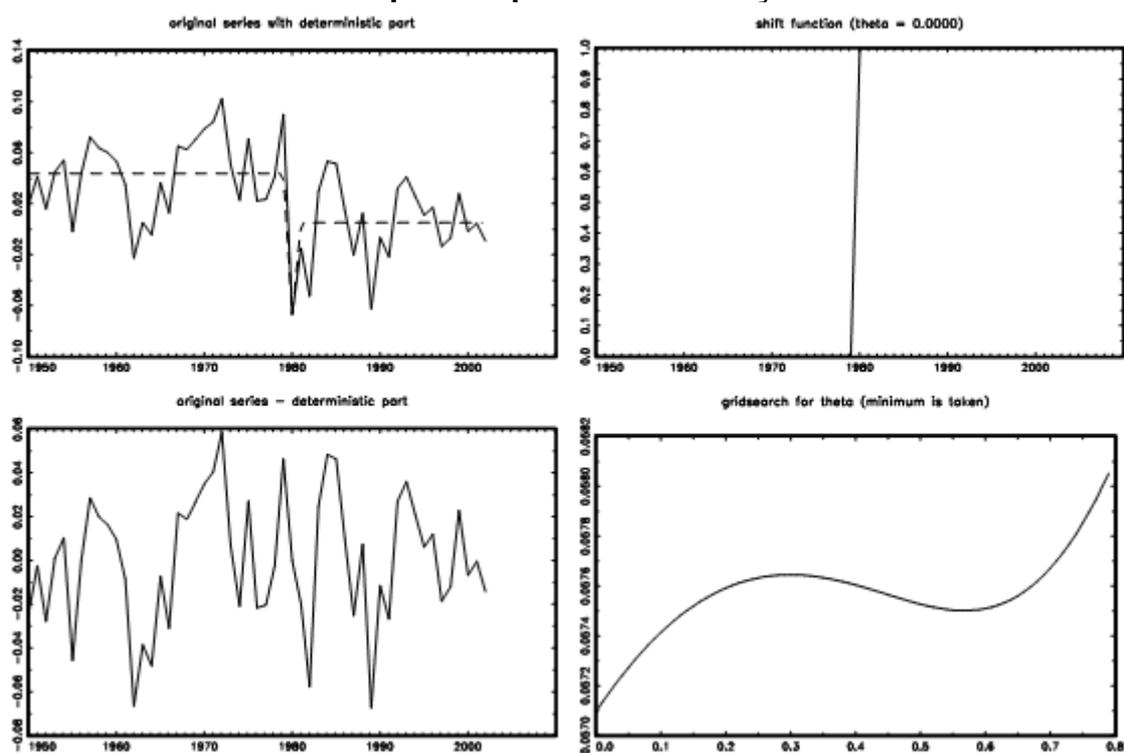
$$\Delta \hat{w}_t = v + \phi \hat{w}_{t-1} + [\hat{\alpha}^*(L) \Delta f_t(\hat{\theta})'] \pi_1 + [\hat{\alpha}^*(L) \Delta F_t(\hat{\theta})'] \pi_2 + \sum_{j=1}^{p-1} \alpha_j^* \Delta \hat{x}_{t-j} + r_t \quad (27)$$

$$(t = p+1, \dots, T)$$

onde $F_t(\theta) = df_t / d\theta$ é a primeira derivada de $f_t(\theta)$ com respeito ao parâmetro θ e r_t denota um termo de erro. O teste de raiz unitária é novamente como um teste t usual do estimador de ϕ baseado na estimação de MQO deste modelo. Como no caso do teste ADF, a distribuição assintótica é não-padrão. Os valores críticos estão tabulados em Lanne et al. (2002). Novamente uma distribuição assintótica diferente é obtida se termos determinísticos são incluídos.

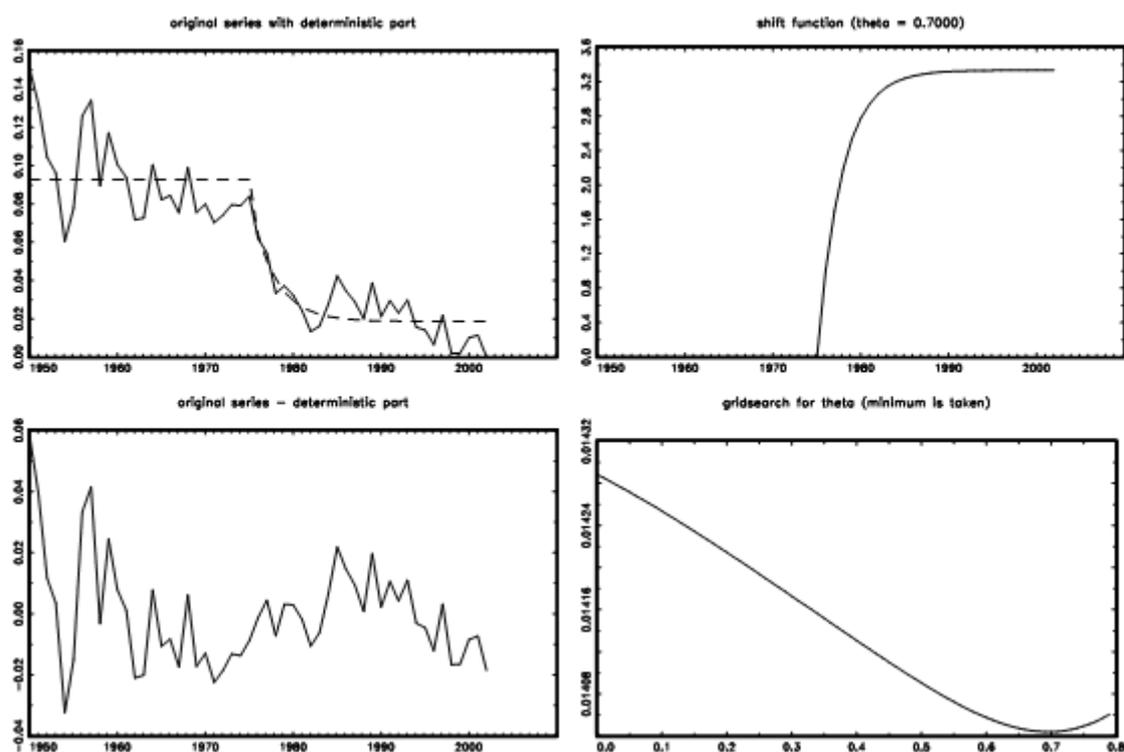
As datas consideradas para realização do teste de raiz unitária com mudança estrutural, determinadas endogenamente pelo teste, foram 1980 e 1976 para o logaritmo do PIB per capita e o logaritmo do estoque de capital público, respectivamente.

Gráfico 6: Teste de Raiz Unitária com Mudança de nível para o PIB per capita em primeira diferença



Fonte: Elaboração Própria

Gráfico 7: Teste de Raiz Unitária com Mudança de nível para o Estoque de Capital Público em primeira diferença



Fonte: Elaboração Própria

Tabela 2 – Teste de Raiz Unitária com Mudança de Nível

Variável	Estatística - t
$\Delta \ln_y$	-3.6451
$\Delta \ln KG$	-3.6893
Valores Críticos	
1%	-3.48
5%	-2.88
10%	-2.58

Fonte: Elaboração própria.

Observações: A estatística t se refere ao modelo com termo constante sem tendência.

Utilizou-se uma estrutura de defasagens igual a zero.

Os valores críticos foram obtidos em Lanne et al. (2002)

O resultado do teste de raiz unitária com mudança de nível mostrado na tabela 2 sugere a rejeição da hipótese nula de raiz unitária. Neste caso, a variável $\ln KG$ possui a mesma ordem de integração da variável \ln_y , ambas as séries são integradas de ordem um, $I(1)$.

6.5. Especificando o Rank de Cointegração

Se as variáveis são $I(1)$, um rank de cointegração r deve ser escolhido em adição ao número de *lags*. Procedimentos de testes seqüenciais baseados em testes da Razão de Verossimilhança (LR) são possíveis ferramentas estatísticas para esta escolha. A seguinte seqüência de hipóteses pode ser considerada:

$$\begin{array}{lll}
 H_0(0) : rk(\Pi) = 0 & \text{versus} & H_1(0) : rk(\Pi) > 0 \\
 H_0(1) : rk(\Pi) = 1 & \text{versus} & H_1(1) : rk(\Pi) > 1 \\
 \vdots & & \vdots \\
 H_0(K-1) : rk(\Pi) = K-1 & \text{versus} & H_1(K-1) : rk(\Pi) = K
 \end{array} \tag{28}$$

A seqüência de testes termina, e o correspondente rank de cointegração é selecionado quando a hipótese nula não pode ser rejeitada pela primeira vez. Se a primeira hipótese nula nesta seqüência, $H_0(0)$, não pode ser rejeitada, um VAR em primeira diferença é considerado. Por outro lado, se todas as hipóteses nulas podem ser rejeitadas, incluindo $H_0(K-1)$, um VAR em nível será considerado.

Sob as hipóteses Gaussianas, a distribuição da estatística LR sob $H_0(r_0)$ é não padrão. Ela depende da diferença $K - r_0$ e dos termos determinísticos incluídos no processo de geração de dados. Em particular, a inclusão de termos de tendência determinística e variáveis *dummy* têm impacto sob a distribuição da hipótese nula dos testes LR. Portanto, testes do tipo LR têm que ser derivados sob diferentes hipóteses sobre os termos determinísticos. Sob a hipótese que o número de *lags* tenha sido especificado apropriadamente, a distribuição limite do teste não depende dos termos dinâmicos de curto prazo.

Considerando o modelo

$$y_t = \mu_0 + \mu_1 t + x_t, \quad (29)$$

onde x_t é um processo VAR(p). Existem três casos de particular interesse do ponto de vista prático. Primeiro, se $\mu_1 = 0$, existe apenas uma média constante sem termo de tendência determinístico. Neste caso, $y_t - \mu_0 = x_t$, e então $\Delta y_t = \Delta x_t$, e da forma VECM de x_t , a média ajustada de y_t tem a forma VECM

$$\Delta y_t = \Pi(y_{t-1} - \mu_0) + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t \quad (30)$$

ou, se um termo de intercepto é utilizado

$$\begin{aligned} \Delta y_t &= v_0^* + \Pi y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t \\ &= \Pi^* \begin{bmatrix} y_{t-1} \\ 1 \end{bmatrix} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t, \end{aligned} \quad (31)$$

onde $\Pi^* = [\Pi : v_0^*]$ é $(K \times (K+1))$ com $v_0^* = -\Pi \mu_0$. Dada a ausência do termo de tendência determinística, o intercepto pode ser absorvido na relação de cointegração. Então, $\Pi^* = \alpha \beta^{* \prime}$ tem rank r . Ambas as versões VECM podem ser usadas para testar o

rank de cointegração. Johansen (1995) considera a versão com intercepto e fornece valores críticos para o teste LR, que é conhecido como o teste do traço. O teste estatístico é da forma

$$LR(r_0) = -T \sum_{j=r_0+1}^K \log(1-\lambda_j),$$

onde os λ_j são os autovalores obtidos pela estimação de ML da equação (31).

Um segundo caso ocorre se uma tendência linear determinística está presente no GDP e, por conseguinte, $\mu_1 \neq 0$. Se a tendência está confinada a alguma variável individual, mas está ausente das relações de cointegração, temos $\beta' \mu_1 = 0$, isto é, a tendência é ortogonal a matriz de cointegração; então $\Pi(y_{t-1} - \mu_0 - \mu_1(t-1)) = \Pi(y_{t-1} - \mu_0)$. Por conseguinte, para este caso, temos

$$\Delta y_t - \mu_1 = \Pi(y_{t-1} - \mu_0) + \sum_{j=1}^{p-1} \Gamma_j (\Delta y_{t-j} - \mu_1) + u_t, \quad (32)$$

onde $\Delta(y_t - \mu_0 - \mu_1 t) = \Delta y_t - \mu_1$ foi usado. Coletando todos os termos constantes no intercepto

$$\Delta y_t = v_0 + \Pi y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t, \quad (33)$$

onde $v_0 = -\Pi \mu_0 + (\sum_{j=1}^p j A_j) \mu_1$. O teste de LR baseados na versão de intercepto do VECM em (28) são tratados por Johansen (1995).

Note que a matriz $\beta_{(K \times r)}$ tem que satisfazer $\beta' \mu_1 = 0$ com $\mu_1 \neq 0$. Este requerimento implica que $r < K$. Então, somente testes da hipótese nula $rk(\Pi) = r_0 < K - 1$ podem ser realizados. Este resultado é uma implicação do fato que uma tendência linear é assumida somente para uma variável, enquanto o modelo estável (onde $r = K$) com um intercepto não possui tendência linear.

O caso final de importância prática surge se um termo de tendência linear irrestrita é incluído em (29). Usando $x_t = y_t - \mu_0 - \mu_1 t$, temos

$$\Delta y_t - \mu_1 = \Pi(y_{t-1} - \mu_0 - \mu_1(t-1)) + \sum_{j=1}^{p-1} \Gamma_j (\Delta y_{t-j} - \mu_1) + u_t, \quad (34)$$

e rearranjando os termos determinísticos, temos

$$\Delta y_t = v + \Pi^+ \begin{bmatrix} y_{t-1} \\ t-1 \end{bmatrix} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t, \quad (35)$$

onde $\Pi^+ = \alpha[\beta': \eta]$ é uma matriz ($K \times (K + 1)$) de rank r com $\eta = -\beta' \mu_1$ e $v = -\Pi \mu_0 + (I_K - \Gamma_1, \dots, \Gamma_{p-1}) \mu_1$. Neste caso, tanto as variáveis como as relações de cointegração possuem uma tendência linear determinística. O modelo (35) leva em conta o fato que o termo de tendência linear determinística pode ser absorvido na relação de cointegração.

Tabela 3 – Teste de Cointegração sem Considerar a Mudança Estrutural

Número de lags	Hipótese Nula	Valor do Teste	Valores Críticos	
			90%	95%
2	$r = 0$	26.82	23.32	25.73
	$r = 1$	7.81	10.68	12.45

Fonte: Elaboração própria

O resultado do teste de cointegração de Johansen mostrado na tabela 3 rejeita a hipótese nula $r = 0$ e não rejeita a hipótese alternativa $r = 1$. O resultado do teste indica que o $rank(\Pi) = 1$, ou seja, existe um vetor cointegrante.

Mudanças Estruturais

Se existe uma mudança estrutural no processo de geração de dados e se a mudança pode ser capturada adicionando uma variável *dummy* a parte determinística do processo, incluir tais termos no modelo provoca uma mudança na distribuição assintótica do teste de Johansen para o rank cointegração. De fato, a distribuição do teste depende de onde as mudanças tenham ocorrido no período amostral. Para o caso de mudanças no período τ , considera-se o seguinte modelo:

$$y_t = \mu_0 + \mu_1 t + \delta d_{t\tau} + x_t. \quad (36)$$

A quantidade $d_{t\tau}$ é uma variável *dummy* definida como

$$d_{t\tau} = \begin{cases} 0, & t < \tau, \\ 1, & t \geq \tau, \end{cases}$$

isto é, $d_{t\tau}$ é uma variável *shift dummy* representando uma mudança no período τ . O processo do erro não observado x_t é assumido ter a forma $VECM(p-1)$. Então o processo observado y_t tem a representação $VECM$

$$\Delta y_t = v + \Pi^{shift} \begin{bmatrix} y_{t-1} \\ t-1 \\ d_{t-1,\tau} \end{bmatrix} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \sum_{j=0}^{p-1} \gamma_j \Delta d_{t-j,\tau} + u_t, \quad (37)$$

onde $\Pi^{shift} = \alpha[\beta': \eta: \theta]$ é $(K \times (K+2))$ de rank r , $\eta = -\beta' \mu_1$, $\theta = -\beta' \delta$ e

$$\gamma_j = \begin{cases} \delta, & j=0 \\ -\Gamma_j \delta, & j=1, \dots, p-1. \end{cases}$$

$\Delta d_{t-j,\tau}$ é uma variável *impulse dummy* com valor 1 no período $t = \tau + j$ e 0 caso contrario.

Tabela 4 – Teste de Cointegração Considerando a Mudança Estrutural

Número de <i>lags</i>	Hipótese Nula	Valor do Teste	Valores Críticos	
			90%	95%
2	$r = 0$	31.14	26.30	28.42
	$r = 1$	8.42	13.21	15.21

Fonte: Elaboração própria

O teste de cointegração considerando as mudanças estruturais mostrado na tabela 3 indica a rejeição da hipótese nula $r = 0$ e não rejeita a hipótese alternativa $r = 1$. Confirmando a existência de uma relação de cointegração entre as variáveis.

7. Resultados

Os modelos serão estimados por dois métodos, pelo método de Máxima Verossimilhança (MV) e pelo Método de Dois Estágios. Além dos modelos estimados esta seção trará o resultado do teste de Chow para estabilidade estrutural e a análise de impulso-resposta.

7.1. Estimativas

A tabela 5 apresenta os resultados dos modelos estimados por Máxima Verossimilhança. No modelo sem levar em consideração as quebras estruturais o valor do coeficiente estimado do logaritmo do estoque de capital público foi estimado com o sinal correto, é estaticamente significativo, e sua magnitude está de acordo com os resultados encontrados na literatura.

Tabela 5: Relações de Cointegração Estimadas por Máxima Verossimilhança

ln_y	lnKG	const	trend	shift trend (76)	shift trend (80)
(1.000)	-0.420	0.590	0.027		
(0.000)	(0.168)	(4.136)	(0.011)		
(0.000)	(-2.499)	(0.143)	(2.388)		
(1.000)	2.235	-73.872	-0.097	0.131	0.031
(0.000)	(1.629)	(39.903)	(0.139)	(0.050)	(0.051)
(0.000)	(1.372)	(-1.851)	(-0.702)	(2.599)	(0.608)

Fonte: Elaboração própria

Ao considerar as mudanças estruturais ocorridas nas séries, o coeficiente do logaritmo do estoque de capital público foi estimado com o sinal incorreto e é estatisticamente não significativo. Esse resultado indica a inadequação da utilização de regressões baseadas em correlações canônicas para controlar mudanças estruturais. A inclusão de termos determinísticos, variáveis *dummy*, na matriz de cointegração provoca a perda de robustez do modelo.

A tabela 6 apresenta os resultados dos modelos estimados pelo estimador de dois estágios. Novamente, no modelo sem considerar as mudanças estruturais o coeficiente do logaritmo do estoque de capital foi estimado com o sinal correto e é estatisticamente significativo. No entanto, seu valor é de magnitude superior ao coeficiente estimado por máxima verossimilhança. Esse resultado (0.56) é igual aos obtidos por Ferreira e Malliagros (1998) e Candido Jr. (2006).

Tabela 6 Relações de Cointegração Estimadas pelo Método de Dois Estágios

ln_y	lnKG	const	trend	shift trend (76)	shift trend (80)
(1.000)	-0.561	5.237	0.011		
(0.000)	(0.091)	(2.234)	(0.006)		
(0.000)	(-6.174)	(2.344)	(1.829)		
(1.000)	-0.403	1.895	-0.015	-0.003	0.011
(0.000)	(0.148)	(3.628)	(0.013)	(0.005)	(0.005)
(0.000)	(-2.721)	(0.522)	(-1.203)	(-0.692)	(2.405)

Fonte: Elaboração própria

Ao considerar as mudanças estruturais no modelo estimado por dois estágios o coeficiente do logaritmo do estoque de capital público foi estimado com o sinal correto, é estatisticamente significativo, e seu valor é inferior aos valores obtidos pelo modelo que não considera as mudanças estruturais estimado pelo método de dois estágios e pelo modelo estimado por máxima verossimilhança. Portanto, obteve-se os resultados esperados, confirmando a importância de considerar as mudanças estruturais ocorridas nas séries.

7.2. Teste de Chow

Seguindo a metodologia de Candelon e Lutkepohl (2001) será utilizado o teste Chow para checar a estabilidade estrutural do modelo. Os autores demonstram por simulações que em pequenas amostras a distribuição do teste estatístico sobre a hipótese

de estabilidade pode ser substancialmente diferente da distribuição assintótica χ^2 . Em alguns casos a distorção é tão grande que o teste se torna ineficaz. Eles consideram então versões de *bootstrap* do teste, e afirmam que estes se tornam muito mais confiáveis em pequenas amostras.

Considera-se duas versões do teste de Chow, o *sample-split test* (SS) e o *break-point test* (BP). Assumindo que a quebra tenha ocorrido no período T_B . O modelo sob consideração é estimado da amostra completa de T observações e da amostra das primeiras T_1 e das últimas T_2 observações, onde $T_1 < T_B$ e $T_2 \leq T - T_B$. Os resíduos resultantes são denotados por \hat{u} , $\hat{u}^{(1)}$ e $\hat{u}^{(2)}$, respectivamente. Além disso, defina $\hat{\Sigma} = T^{-1} \sum_{t=1}^T \hat{u}_t \hat{u}_t'$, $\hat{\Sigma}_{1,2} = T_1^{-1} \sum_{t=1}^{T_1} \hat{u}_t \hat{u}_t' + T_2^{-1} \sum_{t=T-T_2+1}^T \hat{u}_t \hat{u}_t'$, $\hat{\Sigma}_{(1)} = T_1^{-1} \sum \hat{u}_t^{(1)} \hat{u}_t^{(1)'} e$
 $\hat{\Sigma}_{(2)} = T_2^{-1} \sum \hat{u}_t^{(2)} \hat{u}_t^{(2)'}$.

Usando esta notação, o teste estatístico SS pode ser escrito como

$$\lambda_{SS} = (T_1 + T_2) \log \det \hat{\Sigma}_{1,2} - T_1 \log \det \hat{\Sigma}_{(1)} - T_2 \log \det \hat{\Sigma}_{(2)} \rightarrow \chi_{(k)}^2$$

onde k é o número de restrições impostas ao assumir um modelo de coeficiente constante para o período amostral completo, isto é, k é a diferença entre a soma do número de coeficiente estimados no primeiro e no último subperíodos e o número de coeficientes estimados na amostra completa. A hipótese de constância do parâmetro é rejeitada se o valor do teste estatístico λ_{SS} é grande. O teste estatístico BP tem a seguinte forma:

$$\lambda_{BP} = \frac{1 - (1 - R_r^2)^{1/s}}{(1 - R_r^2)^{1/s}} \cdot \frac{Ns - q}{nk} = F(nk, Ns - q)$$

onde

$$s = \left(\frac{n^2 k^2 - 4}{n^2 + k^2 - 5} \right)^{1/2}, \quad q = \frac{nk}{2} + 1, \quad N = t - k_1 - k - (n - k + 1) / 2$$

com k_1 sendo o número de regressores no modelo restrito, e

$$R_r^2 = 1 - \left(\frac{T_1}{T} \right)^n \left| \hat{\Sigma}_{(1)} \right| \left(\left| \hat{\Sigma} \right| \right)^{-1}$$

Novamente a hipótese nula da constância do parâmetro é rejeitada para altos valores de λ_{BP} .

As versões de *bootstrap* dos testes são obtidas estimando o modelo de interesse, denotando os resíduos por \hat{u}_t , computando os resíduos centrados $\hat{u}_1 - \bar{u}, \dots, \hat{u}_T - \bar{u}$, onde $\bar{u} = T^{-1} \sum \hat{u}_t$, e os resíduos gerados por bootstrap u_1^*, \dots, u_T^* são retirados aleatoriamente dos resíduos centrados. Essas quantidades são então usadas para computar a *bootstrap time series* recursivamente partindo de valores pré-amostrais Y_{-p+1}, \dots, Y_0 . O modelo de interesse é então reestimado com e sem as restrições de estabilidade e a versão de *bootstrap* da estatística de interesse, λ_{SS}^* ou λ_{BP}^* , é computada. Repetindo esses passos um grande número de vezes, um valor crítico é então obtido, λ_{crit}^* , da distribuição empírica do teste estatístico com *bootstrap* e a hipótese de estabilidade é rejeitada se $\lambda > \lambda_{crit}^*$. Alternativamente, o *p-value* do teste pode ser estimado como uma fração do número de vezes em que o valor do teste estatístico com *bootstrap* excede o valor de λ .

A tabela 7 apresenta os testes de estabilidade de Chow para o modelo multivariado.

Tabela 7: Resultado do Teste de Chow

	1976	1980
break point Chow test	23.3426	20.4466
bootstrapped p-value	0.0200	0.0300
asymptotic χ^2 p-value	0.0055	0.0153
degrees of freedom	9	9
sample split Chow test	5.4451	4.1159
bootstrapped p-value	0.6000	0.7500
asymptotic χ^2 p-value	0.4881	0.6610
degrees of freedom	6	6

Fonte: Elaboração própria.

Observação: 100 replicações de *bootstrap*.

Para ambas as datas o teste λ_{BP} rejeita a estabilidade dos parâmetros ao nível de significância de 5%. Portanto, conclui-se que o modelo deve ser estimado levando em conta as quebras estruturais em ambas as séries.

7.3. Análise de Impulso-Resposta

Se o processo y_t é $I(0)$, os efeitos de choques nas variáveis de um sistema são mais facilmente visualizados na representação de média móvel (MA) de Wold

$$y_t = \Phi_0 u_t + \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \dots \quad (38)$$

em que $\Phi_0 = I_K$ e

$$\Phi_s = \sum_{j=1}^s \Phi_{s-j} A_j, \quad s=1,2,\dots \quad (39)$$

pode ser computado recursivamente dos coeficientes da forma reduzida do VAR em níveis. Os coeficientes desta representação podem ser interpretados como funções impulso-resposta. Os (i, j) elementos das matrizes Φ_s , considerados como uma função de s , fornecem a resposta esperada de $y_{i,t+s}$ a uma mudança unitária em y_{jt} , mantidos constantes todos os valores passados de y_t . Como a mudança em y_{it} , dados $\{y_{t-1}, y_{t-2}, \dots\}$, é medida pela inovação u_{it} , os elementos de Φ_s representam as funções impulso-resposta dos componentes de y_t com respeito as inovações u_t . No caso $I(0)$, $\Phi_s \rightarrow 0$ quando $s \rightarrow \infty$. Por conseguinte, o efeito do choque é transitório e desaparece ao longo do tempo. Essas funções impulso-resposta são chamadas de funções de impulso-resposta dos erros de previsão porque os u_{it} são os erros de previsão um período a frente.

Embora a representação de Wold não exista para processos não estacionários cointegrados, as matrizes de impulso-resposta podem ser computadas da mesma maneira que em (39) baseado em VARs com variáveis integradas ou em versões de VECM. Neste caso, Φ_s pode não convergir para zero quando $s \rightarrow \infty$; conseqüentemente, alguns choques podem ter efeitos permanentes. Contudo, como Φ_s não tende a zero quando $s \rightarrow \infty$, a função impulso-resposta acumulada não existe. Para sistemas cointegrados a matriz $(I_K - A_1 - \dots - A_p)$ é singular. Pela versão de Johansen do Teorema da Representação de Granger, se y_t é gerado por um VECM na forma reduzida $\Delta y_t = \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t$, ele tem a forma a seguinte representação MA:

$$y_t = \Xi \sum_{i=1}^t u_i + \Xi^*(L)u_t + y_0^*, \quad (40)$$

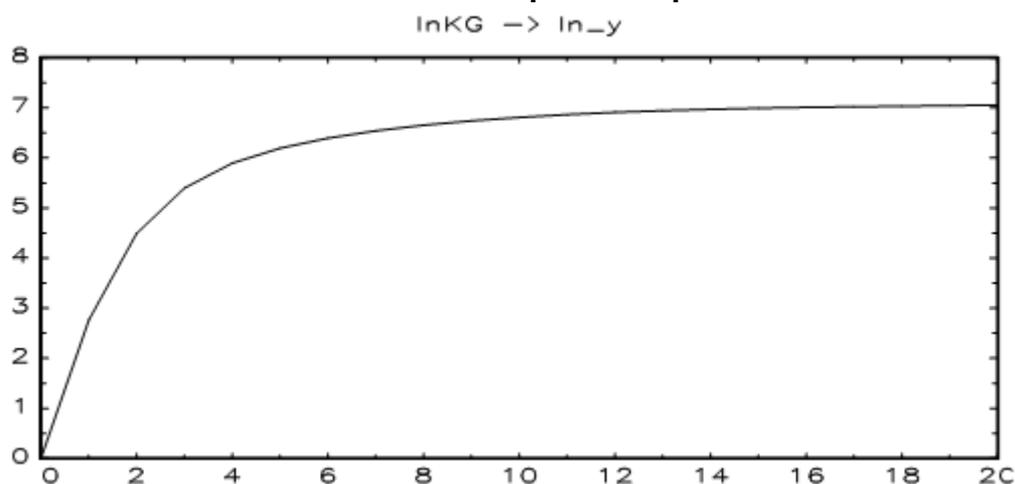
onde $\Xi = \beta_{\perp} \left(\alpha'_{\perp} \left(I_K - \sum_{i=1}^{p-1} \Gamma_i \right) \beta_{\perp} \right)^{-1} \alpha'_{\perp}$ e $\Xi^*(L) = \sum_{j=0}^{\infty} \Xi_j^* L^j$ é um polinômio de ordem infinita no *lag* operador com matrizes de coeficientes Ξ_j^* que tendem a zero quando $j \rightarrow \infty$. O termo y_0^* contém todos os valores iniciais. Note que a matriz Ξ tem rank $K - r$ se o rank de cointegração do sistema é r , e ela representa os efeitos de longo prazo das funções impulso-resposta dos erros de previsão, enquanto as matrizes Ξ_j^* s contém os efeitos transitórios.

Como as funções impulso-resposta dos erros de previsão baseados em Ξ e Ξ_j^* estão sujeitas as mesmas críticas do processo VAR estável, choques apropriados devem ser identificados por uma análise de impulso-resposta significativa. Se u_t é repassado por $A^{-1}B\varepsilon_t$, as funções de impulso-resposta ortogonalizadas de curto prazo podem ser obtidas por $\Xi_j^* A^{-1}B$ em uma maneira análoga ao caso VAR estacionário. Além disso, os efeitos de longo prazo dos choques ε são dados por

$$\Xi A^{-1}B \tag{41}$$

Esta matriz tem rank $K - r$ porque $rk(\Xi) = K - r$ e A e B são não singulares. Então a matriz (44) pode ter no máximo r colunas de zeros. Por conseguinte, pode existir no máximo r choques com efeitos transitórios, e pelo menos $k^* = K - r$ choques tem efeitos permanentes. Dado o rank reduzido da matriz (44), cada coluna de zeros representam apenas k^* restrições independentes. Então, se existem r choques transitórios, as correspondentes coluna de zeros representam $k^* r$ restrições independentes. Para identificar os choques permanentes precisa-se de $k^*(k^* - 1)/2$ restrições adicionais. Similarmente, $r(r - 1)/2$ restrições contemporâneas adicionais são necessárias para identificar os choques transitórios. Juntas existem um total de $k^* r + k^*(k^* - 1)/2 + r(r - 1)/2 = K(K - 1)/2$ restrições. Portanto, assumindo $A = I_K$, existem restrições suficientes para identificar a matriz B.

Gráfico 8: Função Impulso-Resposta de um Choque de Uma Unidade de Desvio-Padrão no Estoque de Capital Público



O gráfico 8 mostra a resposta da variável \ln_y a um choque de uma unidade de desvio-padrão na variável $\ln KG$. Como era de se esperar o PIB per capita responde positivamente a choques no estoque de capital público. Note que a resposta do PIB per capita é crescente durante os quatros primeiros anos, indicando que existe um efeito defasado de um choque no estoque de capital sobre o PIB per capita, que poderia explicar a diferença nas datas das mudanças estruturais nas duas séries.

8. Conclusão

O presente trabalho analisa a relação de longo prazo entre as variáveis PIB per capita e estoque de capital público utilizando um modelo vetorial de correção de erros, considerando as mudanças estruturais ocorridas nas séries. Os resultados dos testes de raiz unitária com mudança estrutural indicam que tais mudanças ocorreram. E os resultados obtidos pelo método de estimação em dois estágios confirmam a importância de se controlar as mudanças estruturais observadas em ambas as séries.

A elevação dos preços do petróleo durante a década de 70 e o aumento das taxas de juros internacionais combinados às políticas econômicas adotadas para a superação da crise externa levaram a economia brasileira a uma profunda recessão no início dos anos 80. As repercussões de tais fatos sobre as séries de PIB per capita e estoques de capital público estão consubstanciadas na mudança das tendências das séries ao longo do tempo.

Enquanto perdurou o modelo de crescimento liderado pelos investimentos do setor público, os investimentos em infra-estrutura representavam aproximadamente 4% do PIB. Com o crescente endividamento externo e a deterioração das finanças públicas, essa participação caiu para 1,85% ao longo dos anos 80. O PIB, após crescer a uma taxa média de quase 10% ao ano, no período compreendido entre o imediato pós-guerra até o final dos anos 70, apresenta uma taxa média de crescimento inferior a 3% ao longo das últimas décadas.

Considerando as mudanças estruturais ocorridas nas séries conseguiu-se estimar a elasticidade produto-gasto público inferior aos valores encontrados por Ferreira e Malliagos (1998). Este resultado está também bastante próximo ao obtido por Ferreira e Araújo (2006), mesmo sem incluir as variáveis estoque de capital privado e capital humano, demonstrando a importância de controlar a ocorrência de mudanças estruturais nas séries.

Uma sugestão para trabalhos futuros seria a inclusão variável estoque de capital privado para evitar um possível viés de variável omitida considerando que, pelos mesmos motivos apresentados, pode ter ocorrido uma mudança estrutural nesta série.

Referencias Bibliográficas

- Abreu, M. et al. (1990). "A Ordem do Progresso". Campus, 12^a ed, cap 11, p. 295-322.
- Aschauer, David (1989). "Is Public Expenditure Productive?" *Journal of Monetary Economics*, v.23, p.177-200.
- Assis, Luciana C. e Dias, Jofilson (2004). "Política Fiscal, Nível Tecnológico e Crescimento Econômico no Brasil: Teoria e Evidência Empírica". XXXII Encontro Brasileiro de Economia. Anpec.
- Barro, R. J (1990) "Government Spending in a Simple Model of Endogenous Growth". *Journal of Political Economic*, 98.
- Candelon, Bertrand e Helmut Lutkepohl (2001). "On the Reliability of Chow-Type Tests for Constancy in Multivariate Dynamic Models." *Economic Letters*, 73, pp 155-160.
- Candido Jr, Oswaldo. (2001). "Os Gastos Públicos no Brasil São Produtivos." *Planejamento e Políticas Públicas*, nº 23. IPEA.
- Candido Jr, Oswaldo. (2006). "Efeitos do Investimento Público sobre o Produto e a Produtividade: Uma Análise Empírica". *Texto para Discussão nº 1204*. IPEA
- Davidson, R. and MacKinnon, J. (1993). *Estimation and Inference in Econometrics*, Oxford University Press, London.
- Enders, Walter (1995). "Applied Econometric Time Series". Jhon Wiley & Sons, Inc.
- Engle, R. F. and Granger, C. W. J. (1987). "Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, 22, p. 313-344.
- Ferreira, P. C. (1996). "Investimento em Infra-estrutura no Brasil: Fatos Estilizados e Relações de Longo Prazo." *Pesquisa e Planejamento Econômico*, v.26, n.2, p.231- 252.
- Ferreira, P. C. e T. Malliagos (1998). "Impactos Produtivos da Infra-estrutura no Brasil: 1950-1995." *Pesquisa e Planejamento Econômico*, v.28, nº 2, pp. 315-338.
- Ferreira, P. C. e C. H. Araújo (2006). "On the Economic and the Fiscal Effects of Infrastructure Investment in Brazil." *Ensaios Econômicos*, nº 613.
- Ghali, K. H. (1999). "Government Size and Economic Growth: Evidence from Multivariate Cointegration Analysis". *Applied Econometrics*, 34, p. 975-987.
- Hansen, Bruce E. (2001). "The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity." *Journal of Economic Perspectives*, v.15, Número 4, pp 117-128.

Hsieh, E. and Lai, K. (1994). "Government Spending and Economic Growth". *Applied Economics Review*, 80, p. 86-91.

Johansen, S. (1995) "Likelihood-based Inference in Cointegration Vector Autoregressive Models". Oxford University Press, Oxford.

Johansen, S., Mosconi, R. and Nielsen, B. (2000). "Cointegration Analysis in the Presence of Structural Breaks in the Determinist Trend". *Econometrics Journal* 3: 216-249.

Kim, I-M. and Mandala, G. S. (1994). "Unit Roots, Cointegration, and Structural Changes. Cambridge University Press, Cambridge.

Lanne, M., Lütkepohl, H. and Saikkonen, P. (2001). Test procedures for unit roots in time series with level shifts at unknown time, Discussion paper, Humboldt-Universität Berlin.

Lutkepohl, H. e Marcus Kratgig (2004). "Applied Time Series Econometric", Cambridge University Press.

Munnel (1990), A. H. "How Does Public Infrastructure Affect Regional Economic Performace. *New England Economic Review*, p. 11-32, Sept. 1990.

Saikkonen, P. and Lütkepohl, H. (2002). Testing for a unit root in a time series with a level shift at unknown time, *Econometric Theory* 18: 313-348.

Yuk, Wing (2005). "Government Size and Economic Growth: Time Series Evidence for United Kingdom, 1830-1993". Working Paper EWP0501. University of Victoria, Canada.