

Um Modelo Econométrico com Parâmetros Variáveis para Carga Tributária Bruta Brasileira

Mário Jorge Mendonça*
Thiago Guerrera Martins*
Cláudio Hamilton dos Santos*

Resumo: Este artigo tem por objetivo estimar um modelo econométrico linear com parâmetros variáveis para análise da carga tributária bruta brasileira (CTBB) trimestral no período 1995-2008. A utilização desse tipo de modelo é justificada pelo fato de o sistema tributário nacional ter sido submetido a seguidas mudanças durante boa parte do período em questão – mais frequentemente em alíquotas e bases de incidência de tributos pré-existentes, mas vez por outra também por conta da eliminação (temporária ou não) desses últimos e/ou a criação de novos tributos. As principais conclusões do artigo são que: (i) o PIB é certamente a principal variável explicativa da dinâmica da CTBB no período em questão; (ii) a parcela da CTBB “autônoma em relação ao PIB” cresceu significativamente durante todo esse último período (possivelmente indicando contínuas melhorias nos procedimentos de fiscalização adotados pelas autoridades tributárias e/ou aumentos na formalização da economia); e, finalmente, que (iii) a elasticidade-PIB da CTBB permaneceu significativamente inferior à unidade ao longo do período em questão, ao contrário do que sugerem estimativas feitas com modelos com parâmetros constantes.

Palavras-chave: Carga Tributária Bruta, Modelo Linear Dinâmico, Sazonalidade, Filtro de Kalman e Amostragem de Gibbs.

Códigos do JEL: H20, H22, C32

Abstract: This article presents a linear econometric model with variable coefficients for the analysis of the quarterly dynamics of the Brazilian gross tax burden in the 1995-2008 years. The choice of this particular model was motivated by the constant changes made in the Brazilian tax system during these years – most frequently in the scope, design, and size of pre-existing taxes, but once in a while also the elimination of some of these taxes and/or the creation of new ones. The main conclusions of the paper are that: (i) GDP is, quite certainly, the main explanatory variable of the dynamics of the quarterly tax burden during the period in question; (ii) the share of the quarterly Brazilian tax burden that does not depend on GDP (or on any other variable, for that matter) increased quite significantly during the latter period – possibly due to continuous improvements on the ability of the Brazilian government to raise tax revenues and/or increases in the size of the economy’s formal sector; and (iii) the GDP-elasticity of the Brazilian tax burden appears to have fluctuated around values well below unity, contrarily to what estimates based on models with constant coefficients would lead us to believe.

Key Words: Brazilian gross tax burden. Linear Dynamic Models, Seasonality, Kalman Filter, Gibbs’ Sampling

JEL Codes: H20, H22, C32

* Instituto de Pesquisa Econômica Aplicada - IPEA

Um Modelo Econométrico com Parâmetros Variáveis para Carga Tributária Bruta Brasileira

1. Introdução

A carga tributária bruta (CTB) de uma determinada economia é dada pela soma dos recursos que o governo recolhe compulsoriamente do setor privado. Tudo o mais permanecendo constante, aumentos na CTB reduzem a renda disponível do setor privado e, portanto, os recursos disponíveis para o financiamento das despesas de consumo e/ou investimento de famílias e firmas. Por outro lado, incrementos na CTB aumentam (via de regra, pelo menos¹) os recursos à disposição do governo e, portanto, permitem diminuições no nível de endividamento público e/ou aumentos na oferta de bens e serviços públicos à população. A dinâmica da CTB é, portanto, um determinante crucial das dinâmicas de diversas outras grandezas macroeconômicas relevantes.

Infelizmente, a modelagem econométrica da CTB brasileira (e de componentes dessa última) é muito dificultada pelas seguidas mudanças em alíquotas e bases de incidência de tributos pré-existentes – e, vez por outra, também pela eliminação (temporária ou não) desses últimos e/ou a criação de novos tributos – que tem caracterizado a administração tributária do país no período pós-Plano Real². Não surpreende, pois, que a literatura brasileira com especificações econométricas explícitas para a CTB brasileira (CTBB) e/ou para os principais componentes dessa última seja relativamente pequena³ e que os resultados reportados pela mesma sejam freqüentemente apenas exploratórios.

O presente texto tem como objetivo contribuir para a referida literatura apresentando um modelo econométrico linear com coeficientes variáveis para uma série trimestral da CTBB no período 1995-2008⁴. A utilização de um modelo com coeficientes variáveis (discutido, dentre outros, por West e Harrison, 1997) é justificada pela conjectura – que nos parece muito plausível *a priori* – de que as seguidas mudanças no sistema tributário nacional estão associadas a seguidas mudanças também nas elasticidades relevantes, de modo a tornar contra-indicada a utilização de especificações com parâmetros fixos.

A utilização de modelos com coeficientes variáveis para a modelagem de séries tributárias não é inédita na literatura brasileira (ver, por exemplo, Hernández, 1998 e

¹ Note-se que nem todo recurso recolhido compulsoriamente pelo governo da sociedade é receita do governo. As contribuições para o FGTS, por exemplo, são recolhidas compulsoriamente da sociedade – e por isso entram no cálculo da CTB – mas não são “receitas públicas” (visto que são de propriedade dos trabalhadores). Por outro lado, nem toda receita do governo é recolhida compulsoriamente da sociedade. Quando o cidadão paga entrada em um museu público, por exemplo, está aumentando a receita do governo, mas não a CTB. Daí que aumentos na CTB não necessariamente implicam aumentos nas receitas públicas e aumentos nessas últimas não necessariamente implicam aumentos na CTB. Por outro lado, a grande maior parte das receitas públicas brasileiras é de natureza tributária e a grande maior parte da CTB brasileira consiste em receitas públicas.

² E mesmo antes disso, como bem apontam Rezende et al. (2007).

³ Santos et al. (2008) apresentam uma resenha da literatura relevante.

⁴ A metodologia de construção da referida série está disponível em Santos e Costa (2008).

Portugal e Portugal, 2001). Mas, até onde sabemos, ninguém aplicou esses modelos à série trimestral da CTBB produzida por Santos e Costa (2008) seguindo a metodologia das contas nacionais do IBGE. Outra inovação do presente estudo, acreditamos, é nossa opção por modelar o componente sazonal da série da CTBB endogenamente. Apesar de pouco usual na literatura, esse procedimento nos permitiu obter estimativas significativamente mais bem ajustadas aos dados, assim como projeções mais confiáveis. Optamos, ainda, por estimar o modelo a partir de procedimentos Bayesianos [ver Gamerman e Lopes (1997); Kim e Nelson (1999); West e Harrison (1997)] tendo em vista que diferentemente da abordagem clássica a prática Bayesiana não é dependente do teorema central do limite. Isso o que elimina o uso de propriedades assintóticas (Gelman et al., 2003, pag. 696) permitindo a obtenção de estimativas mais confiáveis em pequenas amostras do que métodos frequentistas.

Antecipando nossas conclusões, notamos que, ao longo do período em questão, o comportamento do PIB parece ter sido o principal determinante da dinâmica da CTBB – ainda que a parcela da CTBB “autônoma em relação ao PIB” tenha crescido significativamente durante o período (indicando, a nosso ver, contínuas melhorias nos procedimentos de fiscalização adotados pelas autoridades tributárias e/ou o crescimento exógeno verificado no grau de formalização da economia). Esse crescimento é provavelmente o motivo pelo qual nossa estimativa para a elasticidade-PIB da CTBB foi significativamente menor que um – contrariando a sabedoria convencional vigente baseada nos resultados de modelos lineares com parâmetros fixos (ou, na melhor das hipóteses, com poucas “mudanças de regime”). Com efeito, a imensa maior parte dos modelos que assumem que o “termo constante” da equação de regressão é um parâmetro fixo terminam por concluir que o valor da elasticidade-PIB da CTBB é significativamente *maior* do que um [Santos et al. (2008)]. Em nossa interpretação, esse resultado se deve ao fato de que modelos com coeficientes fixos “forçam todo o peso” do crescimento recente da CTBB no PIB, desenfazando (ou mesmo ignorando) a melhoria do estado das artes da fiscalização tributária e o incremento no grau de formalização da economia.

O restante deste artigo está estruturado da seguinte forma. A seção 2 apresenta alguns fatos estilizados acerca da evolução do tamanho e da composição da CTBB no período 1995-2008. A seguir, na seção 3, apresentamos nossa especificação econométrica inicial (e as motivações teóricas da mesma) e os dados utilizados (detalhes sobre a metodologia econométrica utilizada foram acrescentados em um anexo ao texto). Os resultados obtidos com essa última – e com várias outras especificações similares – e a análise dos mesmos são apresentados na seção 4. Por fim, a seção 5 apresenta breves notas à guisa de conclusão.

2. Em que consiste a carga tributária bruta brasileira?

A literatura brasileira é rica em metodologias alternativas de cálculo da CTBB⁵. Nesse estudo, utilizaremos a metodologia das contas nacionais do IBGE, que tem as vantagens de (i) ser compatível com as diretrizes das Nações Unidas; (ii) ser relativamente bem documentada⁶; (iii) ter se mantido relativamente constante nos últimos 15 anos⁷; e, portanto, (iv) ser comparável internacionalmente⁸. De acordo com o IBGE – ou, mais precisamente, de acordo com a tabela sinótica 20 do Sistema de Contas Nacionais Brasil 2002-2006 (IBGE, 2008a, doravante SCN 2006) –, os tributos brasileiros podem ser divididos em cinco categorias, quais sejam: (i) impostos sobre produtos; (ii) outros impostos ligados à produção; (iii) impostos sobre renda e patrimônio; (iv) contribuições sociais efetivas; e (v) impostos sobre o capital (ver tabela 1). A chamada “carga tributária bruta” é simplesmente a soma dos itens (i) a (v) dividida pelo PIB em um dado período contábil (ver tabela 1). Em 2006, por exemplo, CTBB de R\$ 808,6 bilhões representou 34,1% do PIB daquele ano (que atingiu R\$ 2369,8 bilhões).

Tabela 1: Componentes da CTBB em 2006 (em R\$ bilhões)				
	Total	Federal	Estadual	Municipal
Total	808,6	541,6	219,1	47,8
Impostos sobre produtos	336,6	146,6	171,9	18,0
Outros impostos ligados à produção	30,5	16,3	7,9	6,3
Impostos sobre a Renda e a Propriedade	220,9	181,8	22,5	16,6
Contribuições Previdenciárias (INSS+RPPS+FGTS+PIS/PASEP)	216,8	196,9	15,7	4,2
Impostos sobre o Capital	3,8	0,0	1,1	2,7

Fonte: IBGE (2008a, tabela sinótica 20)

O restante dessa seção discute a evolução do tamanho e da composição desses “grandes grupos” de tributos no período 1995-2008. São dois os nossos propósitos com esse esforço de recuperação histórica. Em primeiro lugar, desejamos sublinhar o fato de que a CTBB é um agregado complexo, formado por literalmente dezenas de tributos com lógicas e histórias diferentes. Em segundo lugar, desejamos enfatizar que as segundas mudanças na legislação tributária implicaram em mudanças muito significativas na composição da CTBB – e, portanto, na resposta relativa dessa última variável a (por exemplo) mudanças no PIB – ao longo do período em análise. Ambos esses pontos são, em

⁵ Números diferentes foram divulgados, por exemplo, pela Secretaria da Receita Federal (e.g. 2009), pelo IBGE (e.g. 2008a) e por Afonso et al. (2007). Santos et al. (2008) discute as diferentes metodologias de estimação utilizadas nesses estudos.

⁶ Ver, em particular, IBGE (2004, anexo 8) e IBGE (2008b, anexo 10).

⁷ Ao contrário da metodologia utilizada pela Secretaria da Receita Federal, que sofreu mudanças radicais em 2008.

⁸ Por exemplo, com os dados divulgados pela publicação “General Government Accounts” da OCDE.

nossa opinião, insumos cruciais para a correta interpretação dos resultados apresentados nas seções seguintes⁹.

2.1 – Impostos sobre Produtos

São sete os principais “impostos sobre produto”, quais sejam, (i) ICMS (estadual); (ii) COFINS; (iii) IPI; (iv) ISS (municipal); (v) Imposto sobre Importações; (vi) IOF; e (vii) CIDE¹⁰. Apenas o ICMS e a COFINS, somados, responderam por mais de três quartos de todos os “impostos sobre produtos” e por pouco menos de um terço da CTBB total em 2006. Ademais, os “fatos geradores” e a incidência desses dois tributos (e mesmo do ISS) podem ser tratados – em uma primeira aproximação – como iguais. Com efeito, ambos os tributos incidem (majoritariamente, pelo menos) sobre o valor agregado das empresas em geral. Parece razoável, portanto, supor que a arrecadação desses tributos segue mais ou menos de perto a dinâmica do PIB¹¹. Note-se, entretanto, que – ao contrário do que ocorreu com o ICMS (estadual) – a legislação da COFINS mudou consideravelmente no período 1995-2008. Alíquotas foram aumentadas em 1999 e novamente em 2003-2004, desta feita em conjunto com uma radical mudança na sistemática de recolhimento da COFINS (com a redução da tributação sobre o faturamento e a criação de um regime de incidência não cumulativa).

Voltando nossa atenção para os demais “impostos sobre produtos”, notamos inicialmente que é natural supor que os “impostos sobre importações” respondam ao valor total dessas últimas (que, por sua vez, é positivamente correlacionado com o PIB). Os demais “impostos sobre produtos” são mais idiossincráticos. A arrecadação do IPI, por exemplo, não depende apenas do PIB industrial (que também é positivamente correlacionado com o PIB total, diga-se), mas também de considerações relativas à política industrial/setorial e de emprego – como atestam, aliás, as recentes isenções concedidas ao setor automobilístico. Situações parecidas ocorrem com o IOF e a CIDE-Combustíveis. A arrecadação do IOF não depende apenas do volume de operações financeiras registradas na economia (que, uma vez mais, é positivamente correlacionado com o PIB total), mas também de considerações relacionadas à política monetária. Ademais, as alíquotas do IOF são majoradas em períodos nos quais a cobrança da CPMF é cancelada – tal como ocorreu no início de 1999 e no início de 2008. Também a arrecadação da CIDE combustíveis – criada apenas em 2002 (em mais um esforço de “ajuste fiscal”) – não depende apenas do consumo de combustíveis e lubrificantes (também positivamente correlacionado com o PIB total) mas também da política do governo para os preços desses últimos e da própria

⁹ E, mais geralmente, para a interpretação dos resultados de quaisquer estudos econométricos feitos sobre variáveis tributárias brasileiras nas últimas duas décadas.

¹⁰ As siglas significam “imposto sobre circulação de mercadorias e serviços” (ICMS), “contribuição para o financiamento da seguridade social” (COFINS), “Imposto sobre produtos industrializados” (IPI), “Imposto sobre serviços” (ISS), “imposto sobre operações financeiras” (IOF), e “Contribuição de Intervenção no Domínio Econômico incidente sobre a importação e a comercialização de gasolina e suas correntes, *diesel* e suas correntes, querosene de aviação e outros querosenes, óleos combustíveis (*fuel-oil*), gás liquefeito de petróleo (GLP), inclusive o derivado de gás natural e de nafta, e álcool etílico combustível” (CIDE).

¹¹ Ou, alternativamente, o “valor adicionado total medido a preços básicos” (i.e o PIB menos a arrecadação dos impostos sobre produtos).

administração tributária das empresas produtoras desses produtos (notadamente a Petrobras).

Em suma, a análise “tributo a tributo” dos impostos sobre produtos sugere fortemente que a arrecadação desses últimos é positivamente correlacionada com o PIB. Indica, ainda, que dificilmente a elasticidade-PIB da arrecadação dos impostos sobre produtos permaneceu constante durante o período 1995-2008, tendo em vista a criação da CIDE-combustíveis e as constantes mudanças em alíquotas e bases de incidência (e mesmo, no caso da COFINS, de regime de tributação) desse último tributo, do IOF, do IPI e da COFINS.

2.2 – Contribuições Previdenciárias e “OILPs”

A grande maior parte das contribuições previdenciárias – tanto de empregados do setor privado (e dos “patrões” desses últimos) quanto de servidores públicos (e das próprias administrações públicas, enquanto empregadoras desses últimos) – incide sobre os “rendimentos do fator trabalho”, ou seja, sobre a folha de pagamentos das firmas e dos governos. Isso é verdade, ainda, para as contribuições para o FGTS – que tecnicamente não são receitas públicas, e sim uma “poupança compulsória” dos trabalhadores (e, portanto, propriedade desses últimos)¹². A inclusão do PIB como variável explicativa da arrecadação desses tributos somente pode ser justificada pela hipótese de que essa última variável está positivamente correlacionada com a massa de salários.

Nem todas as “contribuições previdenciárias” incidem sobre a folha de pagamentos, entretanto. Exceções importantes são o SIMPLES e as contribuições para o PIS-PASEP (que somados atingiram R\$ 31,4 bilhões, ou pouco menos de 15% do total das “contribuições previdenciárias em 2006). A rigor, o SIMPLES (instituído em 1999 e modificado seguidas vezes desde então) é um “regime tributário diferenciado, simplificado e favorecido” que permite a micro e pequenas empresas pagarem um único tributo – sobre o faturamento – em substituição a vários outros (IRPJ, COFINS, Contribuições Previdenciárias Patronais, etc). Ou seja, o SIMPLES é classificado como “contribuição previdenciária” apenas por convenção. As contribuições para o PIS-PASEP, por sua vez, são em grande parte cobradas sobre o valor adicionado das firmas (mas apenas desde 2003, quando tiveram seu regime tributário mudado, tal como ocorreu com a COFINS) – sendo classificadas como “previdenciárias” apenas porque servem para capitalizar “fundos” com algumas características (nominais, pelo menos) previdenciárias, a saber o PIS e o PASEP. A inclusão do PIB como variável explicativa desses tributos se deve, portanto, às hipóteses de que o mesmo é positivamente correlacionado tanto com o faturamento das micro e pequenas empresas quanto com o valor adicionado das médias e grandes empresas.

Já os (relativamente pequenos) “outros impostos ligados à produção” (ou OILPs) se dividem em dois grandes grupos. O primeiro deles é formado por tributos que – tal como grande parte das contribuições previdenciárias – incidem sobre a folha de pagamentos,

¹² Assim sendo, o único motivo para a inclusão das contribuições para o FGTS na carga tributária é o fato das mesmas serem compulsórias.

notadamente as contribuições para o “salário educação” e para o “Sistema S”. Assim sendo, o que foi dito acima para as contribuições previdenciárias (com exceção do SIMPLES e do PIS-PASEP) é válido para esses tributos (cuja arrecadação somada atingiu R\$ 13 bilhões em 2006). Já os demais “OILPs” (cuja arrecadação somou R\$ 17,4 bilhões em 2006) são formados por taxas setoriais diversas (e.g. fiscalização de telecomunicações, vigilância sanitária, poder de polícia e etc) com pouca ou nenhuma relação com o PIB.

Em suma, pode-se dizer que (i) a grande maior parte da arrecadação das “contribuições previdenciárias” responde positivamente ao PIB; (ii) que uma pequena parte das mesmas é autônoma com relação ao PIB; e que (iii) as “quebras estruturais” ocorridas na legislação incidente sobre as contribuições previdenciárias foram relativamente menos importantes do que as ocorridas com os “impostos sobre produtos”.

2.3 – Tributos sobre a “renda, o patrimônio e o capital”

São quatro os principais tributos brasileiros sobre a renda, o patrimônio e o capital (IRPCs), quais sejam, o Imposto de Renda (IR), a Contribuição Social sobre o lucro líquido das pessoas jurídicas (CSLL), o IPVA (estadual) e o IPTU (municipal). Com efeito, somente a arrecadação do IR – que é usualmente dividida em seis componentes, quais sejam, o IR pessoa física; o IR pessoa jurídica; o IR retido na fonte sobre os rendimentos do trabalho; o IR retido na fonte sobre ganhos de capital; o IR retido na fonte sobre remessas de divisas ao exterior; e os outros tipos de IR retido na fonte – atingiu R\$ 136 bilhões em 2006, ou pouco menos de 60,5% da arrecadação total dos IRPCs naquele ano. Para efeito de comparação, cumpre notar que a CSLL – um tributo em quase tudo similar ao IR pessoa jurídica (com a importante diferença de ser destinado primordialmente ao financiamento da seguridade social) – arrecadou R\$ 26,7 bilhões (ou 11,9% do total dos IRPCs), o IPVA arrecadou R\$12,3 bilhões (ou 5,5% dos IRPCs) e o IPTU arrecadou R\$ 13,3 bilhões (ou cerca de 5,9% dos IRPCs) naquele mesmo ano.

Além desses tributos, cumpre destacar, ainda, a agora extinta CPMF (contribuição provisória sobre movimentações financeiras). Cobrada entre 1997 e 2007 (com uma significativa mudança de alíquota em meados de 1999) a CPMF arrecadava pouco mais de 1% do PIB em média e perto de 15% de toda a arrecadação dos IRPCs. Ambas a criação e a extinção da CPMF representaram, assim, importantes “quebras estruturais” na série da CTBB. Finalmente, deve-se citar, ainda, os impostos sobre transmissões causa-mortis ou inter-vivos (estaduais), cuja arrecadação atingiu R\$ 3,7 bilhões em 2006 (ou 1,6% dos IRPCs).

Parece-nos, assim, natural assumir que as arrecadações da IR e CSLL dependem positivamente do PIB (dada a correlação positiva entre essa última variável e os lucros das firmas e a massa salarial da economia). O mesmo é verdadeiro no caso da CPMF (enquanto a mesma foi cobrada). Por outro lado, nos parece lícito supor que as arrecadações (menores) do IPTU, IPVA e dos impostos sobre transmissões causa-mortis e inter-vivos sejam relativamente autônomas em relação ao PIB.

3. As Especificações Econométricas e os Dados Utilizados

Claro está que a arrecadação de um determinado tributo depende positivamente tanto da alíquota deste último quanto do montante do “fato gerador” do mesmo. Dito de outro modo, é natural supor, por exemplo, que (tudo o mais permanecendo constante) a arrecadação do imposto sobre importações crescerá com o volume das importações, ou que a arrecadação do IOF crescerá com o volume das operações de crédito. Note-se, entretanto, que o elevado número de tributos existentes no Brasil faz com que a estimação da CTBB “tributo a tributo” seja por demais trabalhosa, de modo que é comum a utilização do PIB como *proxy* para os fatos geradores de todos os tributos brasileiros (e.g. Portugal e Portugal, 2001; Hernández, 1998; Santos et al., 2008). Naturalmente, espera-se que elevações/quedas no PIB (real) estejam associadas às elevações/quedas na arrecadação tributária (real).

Também a inflação é usualmente listada como uma variável explicativa importante da dinâmica da CTB medida em termos reais. O sinal da mesma é incerto, entretanto. Por um lado, elevações da inflação podem diminuir a arrecadação (real) caso os pagamentos (nominais) de tributos sejam recolhidos com atraso pelo governo (Tanzi, 1977). Por outro lado, elevações na inflação reduzem o valor real das “faixas de isenção” (assumindo que as mesmas fiquem fixas em níveis nominais), por exemplo, do imposto de renda – aumentando assim a base de arrecadação do mesmo e, desta forma, a própria arrecadação tributária. Em suma, é de se supor que a inflação afete as receitas públicas de modos diferentes em contextos diferentes.

Note-se ainda que vários estudos (e.g. Hakkio e Rush, 1991; Bohn, 1991 e 1998) buscam explicar a dinâmica dos fluxos das receitas (e dos gastos) correntes do governo a partir de considerações sobre as implicações dinâmicas desses fluxos sobre o endividamento público. A hipótese básica desses estudos é que os governos fixam suas receitas e gastos de modo a manter o endividamento público “sob controle”, por assim dizer. Parece lícito, portanto, acrescentar o comportamento da “dívida líquida do setor público” (DLSP) entre os determinantes (potenciais, pelo menos) da dinâmica da CTBB.

Tendo em vista as considerações acima, optamos por iniciar nossa investigação econométrica com a especificação do modelo linear dinâmico que aparece na equação (1) abaixo. Note-se que a introdução da “constante” nesta especificação não é ingênua, pois (como vimos na seção 2 acima) é lícito assumir que uma parcela da arrecadação tributária (aquela devida à tributação sobre o patrimônio, por exemplo) não está relacionada a nenhum dos fatores descritos acima¹³.

$$\begin{aligned}
 CTBB_t &= b_0(t) + b_1(t)PIB_t + b_2(t)DLSP_t + b_3(t)INFLA_t + S_t + v_t \\
 b_i(t) &= \varphi_i b_i(t-1) + w_i(t), \\
 i &= 0, \dots, 3
 \end{aligned}
 \tag{1}$$

¹³ É interessante notar, entretanto, que o modelo assume que o termo “constante” varia com o passar do tempo (e daí as aspas que adicionamos ao mesmo).

onde $v(t) \sim N(0, V)$, $w_i(t) \sim N(0, W_i)$ e ϕ_i é o componente autoregressivo da equação de estado para cada b_i . S_t é o componente sazonal, que é modelado de acordo com a representação tipo de Fourier (West e Harrison, 1997), de modo que

$$S_t = \sum_{j=1}^2 \left[a_j \text{sen}\left(\frac{2\pi j t}{4}\right) + b_j \cos\left(\frac{2\pi j t}{4}\right) \right] = Z_t \gamma$$

onde,

$$Z_t = \left(\text{sen}\left(\frac{2\pi t}{4}\right) \quad \text{sen}\left(\frac{2\pi 2t}{4}\right) \quad \cos\left(\frac{2\pi t}{4}\right) \quad \cos\left(\frac{2\pi 2t}{4}\right) \right)^T \text{ e } \gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4).$$

De modo a fazer com que as estimativas obtidas expressem os valores das elasticidades, empregamos aqui a transformação logarítmica, excetuando-se a variável inflação. Os dados trimestrais se referem ao período janeiro de 1995 a dezembro de 2008 e são mostrados graficamente no Anexo 1. Conforme pode ser visto a CTBB apresenta um destacado padrão sazonal tanto em valores reais como em relação ao PIB – e daí a necessidade de incluir o componente sazonal acima no modelo. As variáveis usadas neste estudo são as seguintes:

CTBB : carga tributária bruta brasileira nominal deflacionada pelo IPCA¹⁴ ;

DLSP : razão entre o valor mensal da dívida líquida do setor público consolidado e o PIB (acumulado nos últimos 12 meses e valorizado pelo IGP-DI). Fonte: Banco Central do Brasil;

INFLAÇÃO : taxa de inflação medida pela variação da média do IPCA no trimestre. Fonte: IBGE;

PIB : produto interno bruto medido a preços de mercado e deflacionado pelo IPCA. Fonte: IBGE.

4. Resultados

Nesta seção são apresentados os resultados da estimação do modelo “inicial” definido pela equação (1), assim como o de especificações relacionadas – a metodologia

¹⁴ A série utilizada foi uma versão revista e estendida da série original apresentada em Santos e Costa (2008). As principais revisões se deveram aos fatos do IBGE ter divulgado a CTB anual de 2006 em outubro de 2008 (esse valor não estava disponível quando dos cálculos iniciais de Santos e Costa) e desse mesmo instituto ter revisto consideravelmente os dados da arrecadação trimestral dos impostos sobre produtos em 2007 apresentados nas contas nacionais trimestrais. Ademais, os dados originais de Santos e Costa (2008) vão apenas até 2007, de modo que foi necessário estender a série até 2008. A série revista e estendida está à disposição dos leitores, mediante contato com os autores.

econométrica precisa utilizada em nossas estimações é relativamente complexa e está descrita no anexo 2.

De modo a obter a melhor especificação testamos um conjunto de cinco modelos. A seleção do melhor modelo foi feita com base sobretudo nos critérios EQM e DAM¹⁵ que avaliam a projeção dos modelos fora da amostra (que no nosso caso é implementada quatro passos à frente¹⁶). Em todos os modelos a sazonalidade foi tratada a partir da aplicação da representação de Fourier conforme definida na seção 3. Nosso modelo inicial (modelo 1) inclui todas as variáveis explicativas mencionadas na literatura, quais sejam, o PIB, a taxa de inflação e a DLSP (ver equação 1 da seção 3). O modelo 2 incorpora apenas a inflação e o PIB. O modelo 3 inclui apenas o PIB e a DLSP. O modelo 4 apresenta apenas o PIB como variável explicativa. Por fim, o modelo 5 é uma versão do modelo 1 com parâmetros fixos. As tabelas 1 e 2 mostram, respectivamente, os valores obtidos para EQM e DAM com relação ao grau de ajustamento do modelo implementado dentro da amostra, bem como aquele gerado a partir da previsão feita quatro passos à frente fora dela.

Tabela 1: Ajustamento do modelo.

	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
EQM	0.000162	0.000272	0.0001617	0.000576	0.029299
DAM	0.036311	0.049863	0.0360563	0.071696	0.475630

Tabela 2: Previsão quatro passos à frente fora da amostra

	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
EQM	0.000762	0.000157	0.001166	0.000307	0.004561
DAM	0.027200	0.011594	0.027086	0.017437	0.064972

Tomando por base o grau de ajustamento do modelo, verifica-se pela tabela 1 que não existe uma diferença muito marcante entre os quatro modelos com parâmetros variáveis – o ajustamento dos modelos 1 e 3 mostraram-se apenas ligeiramente maior que os dos demais. Note-se, entretanto, que o modelo com parâmetros constantes (modelo 5) apresentou um ajustamento sensivelmente pior do que os dos modelos com parâmetros variáveis.

Por outro lado, considerando-se o critério de previsão fora da amostra os modelos 2 e 4 apresentam resultados claramente superiores aos demais. Tendo em vista que o modelo 2 mostrou uma ligeira vantagem tanto em relação à capacidade de previsão quanto em

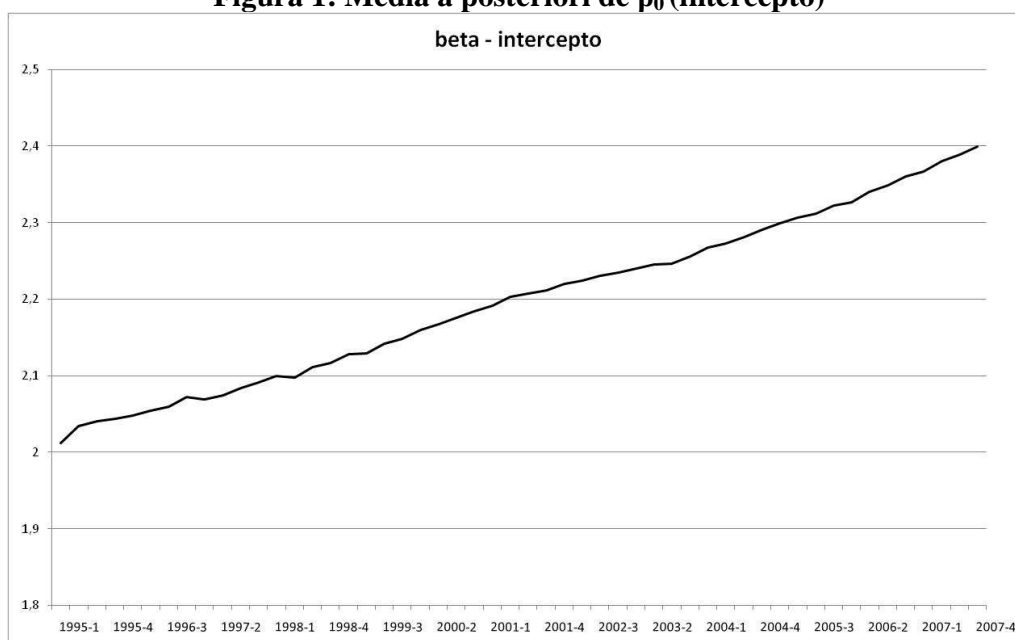
¹⁵. Erro quadrado médio; (EQM) = $N^{-1} \sum_{i=1}^N e_i^2$; Desvio Absoluto Médio (DAM) = $N^{-1} \sum_{i=1}^N |e_i|$.

¹⁶ Deve-se se ter em mente que a previsão é feita tendo em vista o conhecimento prévio das variáveis exógenas no horizonte de previsão.

relação ao ajustamento (em relação ao modelo 4), optamos por tomar este último como nosso modelo referencial.

As figuras 1 a 3 mostram a evolução dos coeficientes de cada uma das variáveis, além do intercepto, para o modelo 2. Inicialmente podemos verificar que o coeficiente relativo à inflação parece ser não significativo em alguns períodos. Em termos práticos, o critério para não significância de um parâmetro na abordagem Bayesiana é que o zero esteja incluído na banda de erro ou banda Bayesiana¹⁷. Embora a banda de erro não seja mostrada na figura 2, podemos adiantar que valor zero muitas vezes aparece dentro do intervalo de confiança¹⁸. Pode-se ainda notar que em alguns momentos a média a posteriori obtida para este coeficiente é zero. Isso significa que não é sempre que a CTBB reage a mudanças na inflação (um comportamento que se verificou em todos os outros modelos).

Figura 1: Média a posteriori de β_0 (intercepto)



¹⁷ Equivalente ao intervalo de confiança no enfoque clássico ou frequentista.

¹⁸ Vale lembrar que tendo em vista que no nosso modelo o parâmetro varia no tempo a banda de erro para o coeficiente é obtida a cada período de acordo com o procedimento de amostragem de Gibbs (Gibbs sampling).

Figura 2: Média a posteriori de β_1 (inflação).

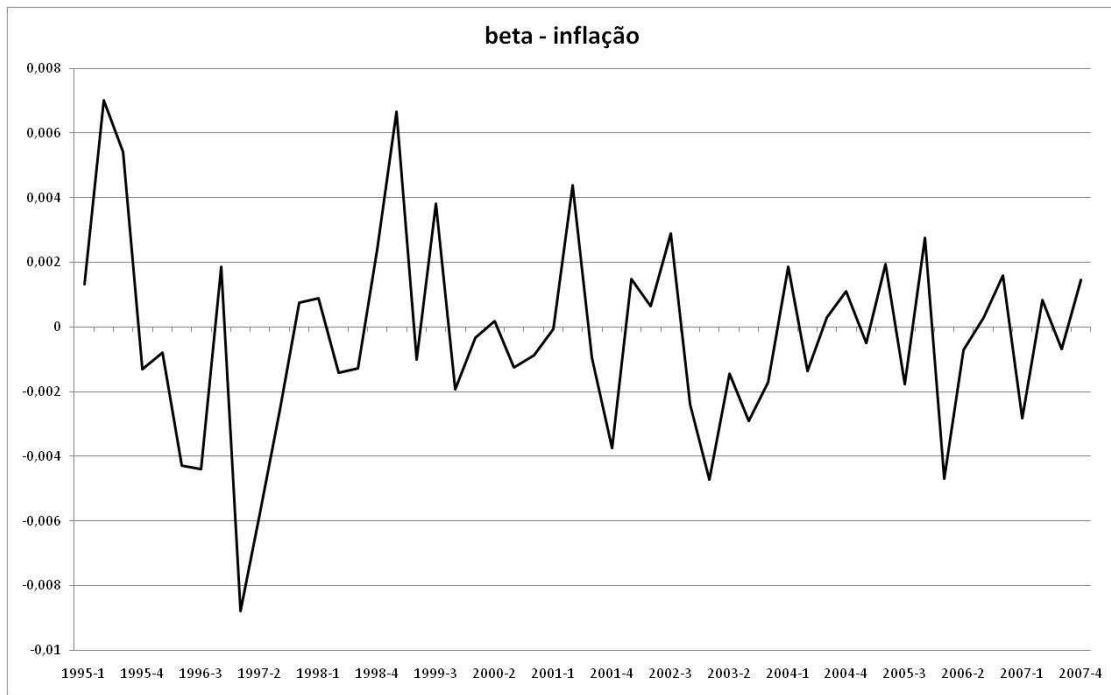
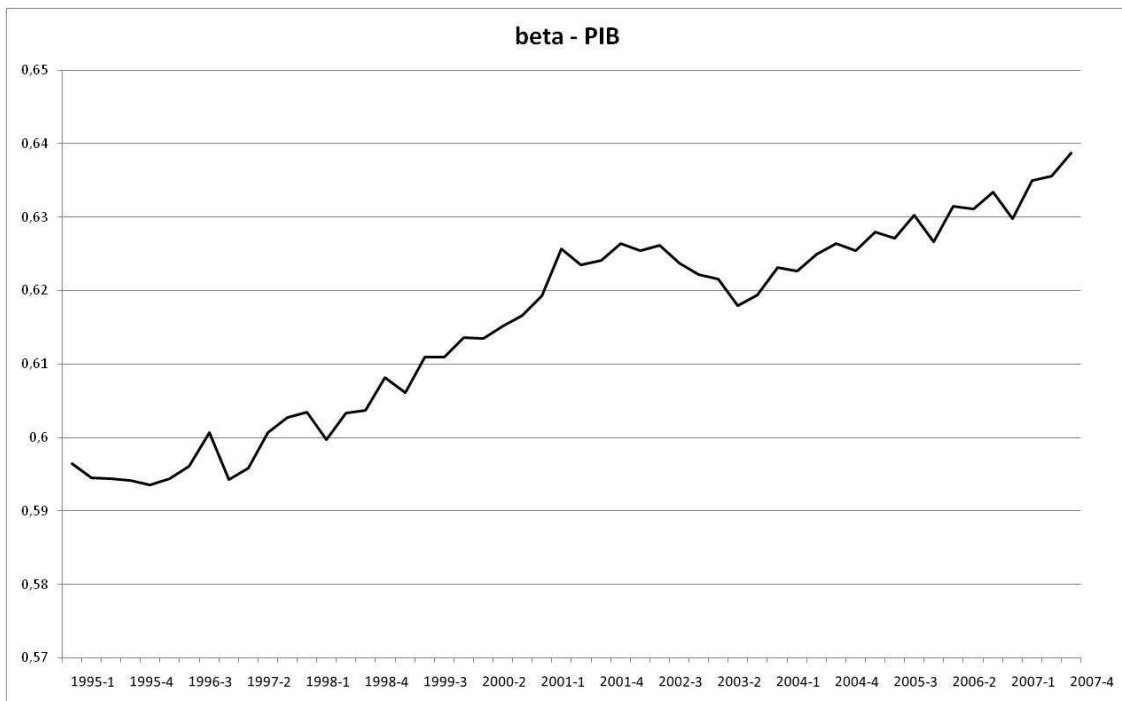


Figura 3: Média a posteriori de β_2 (PIB).



Na figura 3 pode ser visto que muito embora o coeficiente da variável PIB pareça indicar uma trajetória de crescimento contínuo desde 1995, isso se deve mais a escala do gráfico e não ao que está acontecendo de fato. Vale dizer que os valores médios, inicial e final para este coeficiente, são respectivamente, 0,596 e 0,638. Isso mostra que a elasticidade do PIB tem se mantido praticamente constante ao longo de todo o período. Aqui, a elasticidade da carga tributária em relação ao PIB é relativamente inelástica tendo como valor médio global em torno de 0,600. Um ponto positivo é que a trajetória para o parâmetro da variável PIB está contida dentro do quadrante positivo, tal como era esperado para o sinal do coeficiente dessa variável.

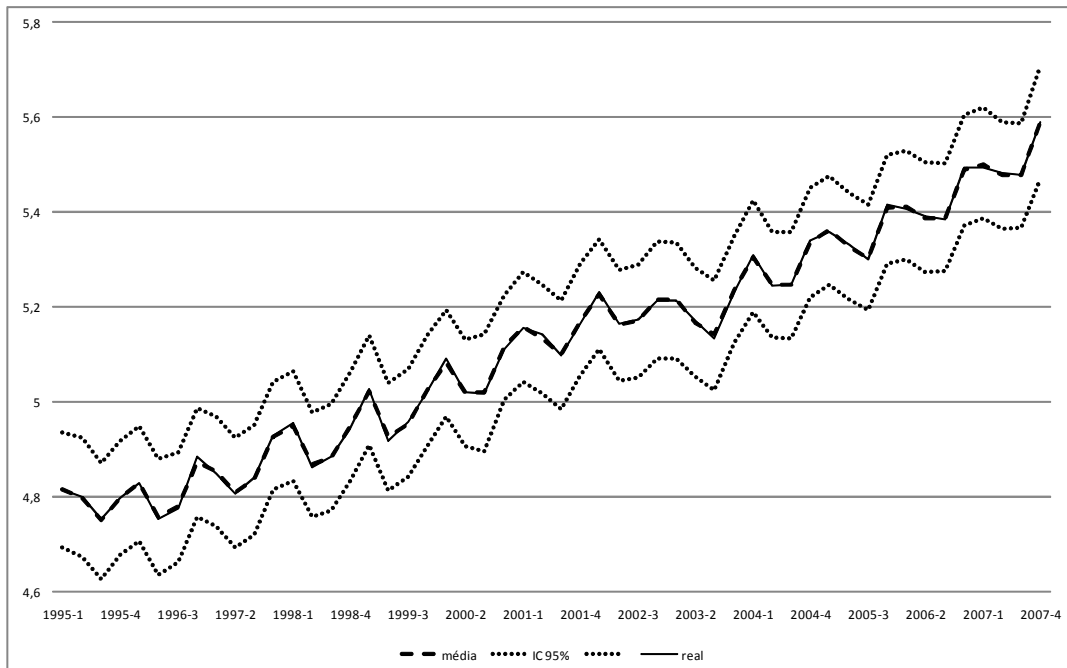
Diferentemente, em relação ao intercepto se observa um crescimento mais firme. Neste caso, o valor médio inicial ficou próximo a 2,00 enquanto o valor médio final se situou em 2,40 tal como se observa na figura 1. Uma interpretação econômica para isso pode se dar a partir do reconhecimento de que fatores não explicitamente incluídos no modelo estariam agindo para fomentar o crescimento da carga tributária. Um deles talvez seja a diminuição do grau de informalidade da economia ou crescimento do emprego formal.

A figura 5 ilustra a previsão quatro passos à frente fora da amostra, enquanto o ajustamento do modelo é apresentado na figura 6. Pode-se observar o alto grau de ajustamento bem como a qualidade da previsão obtida (o erro de previsão anual total foi de R\$ 7,8 bilhões negativos, ou -0,7% do total de R\$ 1066 bilhões da CTBB anual valorizada pelo IPCA verificada no período).

Figura 5: Previsão para o período Q1/08-Q4/08 .



Figura 6: Ajustamento do modelo



A tabela 3 mostra a média assim como o intervalo de confiança de 95% para a variância V da equação do modelo 2, os coeficientes da sazonalidade $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$, assim como as variâncias (W_1, W_2, W_3) e os coeficientes autoregressivos (ϕ_1, ϕ_2, ϕ_3) das equações que descrevem a variação dos parâmetros do modelo ao longo do tempo (ver seção 3). Por fim, na tabela 3 ainda é apresentado o resultado da estatística Gelman-Rubin R para checar a convergência dos parâmetros gerados por meio do processo de estimação do modelo (o chamado algoritmo MCMC – ou Monte Carlo Markov Chains, ver anexo 2). A estatística R é obtida a partir de múltiplas cadeias markovianas obtidas com o referido processo de estimação. A convergência é alcançada quando o valor de R se situa próximo a 1, abaixo do nível crítico. Cadeias independentes foram geradas para 15000 simulações. Em todos os casos as cadeias convergiram.

Tabela 3: Estimativas pontuais e intervalos de 95% de para V, W, Γ e φ .

θ	$E(\theta)$	IC 95%		R
V	0,00609	0,00304	0,01077	1,00118
W_1	0,00482	0,00199	0,00951	1,00154
W_2	0,00243	0,00122	0,00442	1,01031
W_3	0,00100	0,00062	0,00153	1,00084
γ_1	0,04436	-0,01108	0,10107	1,01031
γ_2	0,00760	-1,61458	1,60052	0,99973
γ_3	0,02538	-0,03157	0,08214	1,01253
γ_4	-0,00056	-0,03633	0,03443	1,01639
φ_1	0,73974	0,03306	1,00951	1,06159
φ_2	0,31607	-0,19265	0,82166	1,01035
φ_3	0,81662	0,10041	1,01218	1,08164

6. Considerações finais

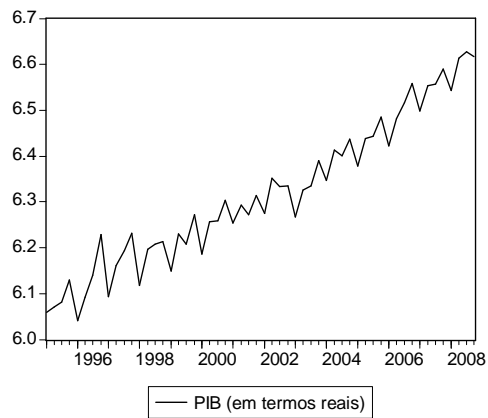
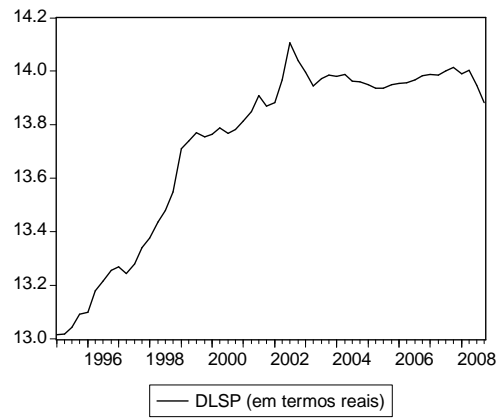
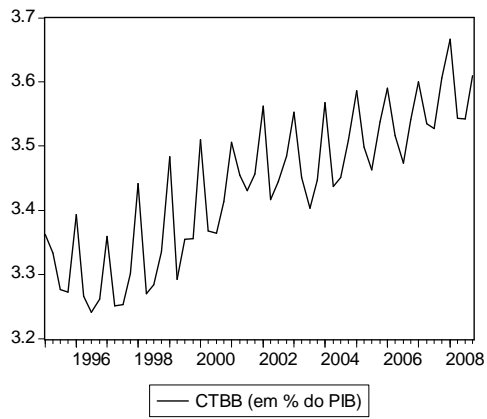
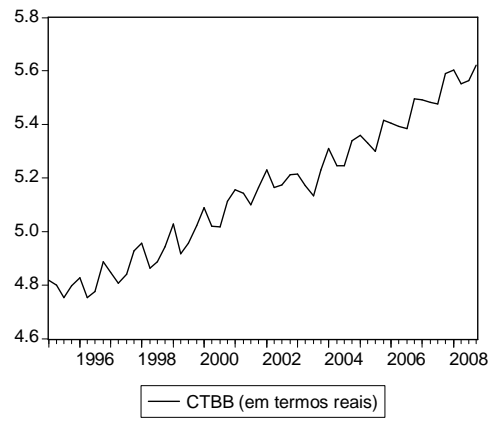
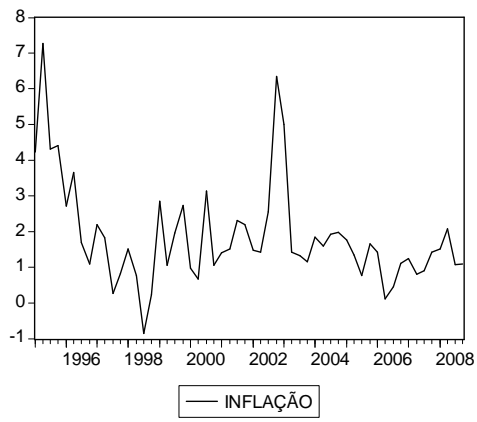
Nas seções anteriores apresentamos (e contextualizamos) um modelo econométrico linear com parâmetros variáveis para a dinâmica da carga tributária bruta brasileira trimestral. Os resultados que obtivemos com o referido modelo – tanto no que toca o ajustamento desse último “dentro da amostra” como no que toca à qualidade das previsões do modelo “fora da amostra” – nos parecem muito satisfatórias. Em particular, o erro de previsão do modelo “fora da amostra” foi inferior a 1% do valor total da CTBB em 2008.

Alguns estudos apresentam valores para a elasticidade da CTBB em relação ao PIB iguais ou superiores a 1,0 (ver, por exemplo, Santos et al. 2008). A questão que se apresenta, então, é o porquê de o nosso modelo diferir em relação a esses estudos. De modo a tratarmos esta questão, estimamos um modelo com intercepto constante (não apresentado aqui), mas permitindo que o coeficiente do PIB variasse no tempo. Neste caso, observamos que o valor da elasticidade do PIB aumenta fortemente, se situando num patamar bastante próximo ao indicado pelos resultados de Santos et al. (2008). Nosso diagnóstico é que métodos econométricos que não permitem variação contínua em todos os parâmetros do modelo não são indicados para estimação de um modelo da CTBB por que, como mencionamos na introdução, a evolução da CTBB é um fenômeno que está sujeito a contínuas mudanças oriundas de vários fatores.

7. Bibliografia

- Afonso, J. R.; Meirelles, B. B.; Castro, K. P. Carga tributária: a escalada continua. *Boletim de Desenvolvimento Fiscal*. IPEA, Brasília, n. 4, p. 25-32, mar. 2007.
- Bohn, H. (1991). "The Sustainability of Budget Deficits with Lump-Sum and with Income-Based Taxation". *Journal of Money, Credit and Banking*, 23(3), 581-604.
- _____ (1998). The Behavior of U.S. Public Debt and Deficits. *The Quarterly Journal of Economics*, 113: 949-963.
- Carter, C. K. and Kohn, R. (1994). *On Gibbs Sampling for State Space Models*, *Biometrika* 81 (3), 541-553.
- Gamerman, D. and Lopes, H. (1997). *Markov Chain Monte Carlo*. Chapman & Hall.
- Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. Bayesian data analysis. Chapman & Hall/CRC, 2003.
- Hakkio, C. and Rush, M. (1991). *Is the Budget Deficit Too Large?* *Economic Inquiry*, 29(7): 429-445.
- Hamilton, J. (1993). *Time Series Analysis*. Princeton University Press
- Hernández, B. (1998). *Um modelo econométrico da conta corrente do governo no Brasil 1951/95*. TD IPEA 543.
- IBGE (2004). *Sistema de Contas Nacionais: Brasil*. Série Relatórios Metodológicos, v.24. Rio de Janeiro.
- _____ (2008a). *Sistema de Contas Nacionais Brasil 2002-2006*. Coordenação de Contas Nacionais. Rio de Janeiro.
- _____ (2008b). *Sistema de Contas Nacionais: Brasil*. Série Relatórios Metodológicos, v.24. Rio de Janeiro. Segunda Edição.
- Kim, C.J. and Nelson, C.R. (1999) *State-Space Models with Regime Switching*. Massachusetts: The MIT Press.
- Portugal, C. e Portugal, M. (2001). *Os efeitos da inflação sobre o orçamento do governo: uma análise empírica*. *Estudos Econômicos*, 31(2): 239-283.
- Rezende, F.R; Oliveira, F; Araújo, E. (2007). *O Dilema Fiscal: Remendar ou Reformar?* Rio de Janeiro: Confederação Nacional da Indústria e Fundação Getúlio Vargas.
- Santos, C.H.M e Costa, F.R. (2008). Uma metodologia de estimação da carga tributária bruta brasileira em níveis trimestrais. *Economia Aplicada*. V.12, n.4, 581-606, 2008
- Santos, C. H. M., Ribeiro, M. B. e Gobetti, S. W. (1998). *A Evolução da Carga Tributária Bruta Brasileira no Período 1995-2007: Tamanho, Composição e Especificações Econométricas Agregadas*. TD IPEA 1350.
- Secretaria da Receita Federal. (2009). *Carga Tributária no Brasil 2008*. Brasília.
- Tanzi, V. (1977). Inflation, lags in collection, and the real value of tax revenue. *International Monetary Fund Staff Papers*. Washington, v.24, n.1,
- West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer

Anexo 1. Gráfico das variáveis



Anexo 2: A Metodologia de Estimação Utilizada nesse trabalho

A.2.1. Modelo de Regressão Linear Dinâmico

A forma geral do modelo DLM pode ser definida da seguinte forma,

$$y_t = X_t \theta_t + S_t + v_t' \quad (2.1)$$

$$\theta_t = \Gamma_1 \theta_{t-1} + \dots + \Gamma_p \theta_{t-p} + \omega_t' \quad (2.2)$$

onde y_t é a variável endógena e X_t é um um vetor $k \times 1$ de variáveis exógenas. O modelo aqui permite ainda a presença de um componente sazonal, S_t , que pode modelado de acordo com a representação tipo de Fourier (West e Harrison, 1997), de modo que

$$S_t = \sum_{j=1}^2 \left[a_j \text{sen} \left(\frac{2\pi j t}{4} \right) + b_j \cos \left(\frac{2\pi j t}{4} \right) \right] = Z_t \varphi$$

onde,

$$Z_t = \left(\text{sen} \left(\frac{2\pi t}{4} \right) \quad \text{sen} \left(\frac{2\pi 2t}{4} \right) \quad \cos \left(\frac{2\pi t}{4} \right) \quad \cos \left(\frac{2\pi 2t}{4} \right) \right)^T \text{ e } \varphi = (a_1 \quad a_2 \quad b_1 \quad b_2) \text{ , } \theta_t \text{ é o}$$

vetor $k \times 1$ de parâmetros que segue um processo autoregressivo de ordem p , onde os coeficientes estão contidos na matriz diagonal Γ_i , $k \times k$, para $i=1, \dots, p$. Iremos assumir que $\omega_t' \sim N(0_{k \times 1}, W_{k \times k})$ e $v_t' \sim N(0, V)$, admitindo ainda que ω_t' e v_t' são não correlacionados. Por simplicidade é assumido também que $W = \text{diag}(w_1, \dots, w_k)$.

A idéia da aplicação da forma de Fourier para tratar sazonalidade está ligada ao fato que qualquer padrão cíclico pode ser representado em termos da combinação linear de funções periódicas¹⁹. Outra motivação para o uso desta abordagem, além da flexibilidade, é a facilidade de se obter interpretação já que as trajetórias em forma de ondas podem ser associadas aos variados padrões observados na prática. Outra forma de tratar a sazonalidade seria introduzi-la diretamente na equação de estado. Neste caso, a equação 2.2 assumiria a seguinte forma

$$\theta_t = \Gamma_1 \theta_{t-1} + \dots + \Gamma_p \theta_{t-p} + S_t + \omega_t' \quad (2.2')$$

Por fim, a sazonalidade pode tomar uma forma dinâmica. Neste caso, poderíamos ter um modelo onde a equação da medida é dada da seguinte forma, $y_t = F_t \theta_t^* + v_t'$, onde $F = [X_t, Z_t]$ e $\theta_t^* = [\theta_t, \varphi_t]$.

¹⁹Uma função $g(t)$ é dita periódica se para algum inteiro $p > 1$, e todo $n, t > 0$ tem-se que $g(t + np) = g(t)$.

Faremos agora uso da seguinte notação para encaixar as equações numa forma mais compacta que reduz o sistema acima numa forma espaço-estado de ordem um,

$$\underline{\theta}_t = \begin{bmatrix} \theta_t \\ \theta_{t-1} \\ \vdots \\ \theta_{t-p+1} \end{bmatrix}_{kp \times 1}, \quad G = \begin{bmatrix} \Gamma_1 & \Gamma_2 & \cdots & \Gamma_{p-1} & \Gamma_p \\ I_k & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & I_k & 0 \end{bmatrix}_{kp \times kp}, \quad \underline{\omega}_t = \begin{bmatrix} \omega_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{kp \times 1} \quad \text{e} \quad F_t' = \begin{bmatrix} X_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{kp \times 1}$$

Visto isso temos que

$$y_t = F_t' \underline{\theta}_t + z_t \varphi + v_t' \quad (3.1)$$

$$\underline{\theta}_t = G \underline{\theta}_{t-1} + \underline{\omega}_t \quad (3.2)$$

onde (3.1) representa a equação da medida ou do movimento enquanto (3.2) define a equação de estado.

A função de verossimilhança de (θ, G, φ, V) é dada por

$$p(y | \theta, G, \varphi, V) = (2\pi)^{-T/2} V^{T/2} \exp\left(-\frac{1}{2V} \sum_{t=1}^T (y_t - F_t' \underline{\theta}_t - z_t \varphi)^2\right)$$

onde $y = (y_1 \quad \dots \quad y_T)$, $\theta = (\underline{\theta}_1 \quad \dots \quad \underline{\theta}_T)$ e $\Gamma = (\Gamma_1 \quad \dots \quad \Gamma_p)$

A.2.2 Procedimento de inferência

Nesta seção um procedimento de inferência baseado no paradigma Bayesiano é realizado. Inicialmente distribuições a priori para todos os parâmetros são apresentadas. Em seguida, o algoritmo MCMC é utilizado para obter amostras a posteriori de todos os parâmetros do modelo.

A.2.2.1. Distribuição a priori

Por simplicidade, distribuições a priori conjugadas são utilizadas para todos os parâmetros. A distribuição a priori para θ_t são especificadas na equação (3.2) e completadas pela informação inicial θ_0 tal que $\theta_0 \sim N(m_0, C_0)$ onde m_0 e C_0 são hiperparâmetros conhecidos. As distribuições a priori para os parâmetros V, W, Γ e φ são as seguintes: i) $V \sim GI(n_v / 2, n_v S_v / 2)$; ii) $w_j \sim GI(n_w / 2, n_w S_w / 2)$, $j = 1, \dots, k$; iii) $\Gamma_j \sim N(m_\Gamma, S_\Gamma)$, $j = 1, \dots, p$ e iv) $\varphi \sim N(m_\varphi, S_\varphi)$ onde $n_v, S_v, n_w, S_w, m_\Gamma, S_\Gamma, m_\varphi$ e S_φ são hiperparâmetros conhecidos.

A.2.2.2. Inferência a posteriori

A distribuição a posteriori conjunta de $(\theta, \Gamma, \varphi, V, W)$ é dada por

$$p(\theta, G, \varphi, V, W | y) \propto \prod_{t=1}^T p(y_t | \underline{\theta}_t, \varphi, V) \prod_{t=1}^T p(\underline{\theta}_t | \underline{\theta}_{t-1}, W, G) p(\underline{\theta}_0 | m_0, C_o) P(\varphi) p(V) p(W) p(\Gamma)$$

a qual é analiticamente intratável e, portanto, inferência a posteriori exata é feita utilizando o esquema MCMC. Neste caso, o vetor da variável de estado são amostrados conjuntamente utilizando o algoritmo FFBS (forward filtering backward sampling) de Carter & Kohn (1994) e Frühwirth-Schnatter (1994). Para o resto de parâmetros Γ , φ , V e W , todas as distribuições condicionais completas são: distribuições normais ou distribuições Gama inversa. Todas as distribuições condicionais completas estão listadas a seguir.

- A distribuição condicional de V é $(V | \theta, W, \varphi, \Gamma) \sim GI(n_V^* / 2, n_V^* S_V^* / 2)$ onde $n_V^* = n_V + T$ e $n_V^* S_V^* = n_V S_V + \sum_{t=1}^T (y_t - F_t \underline{\theta}_t - z_t \varphi)^2$
- A distribuição condicional de $W = \text{diag}(w_1, \dots, w_k)$ é $(w_j | \theta, V, \varphi, \Gamma) \sim GI(n_W^* / 2, n_W^* S_W^* / 2)$ para $j = 1, \dots, k$ onde $n_W^* = n_W + T - 1$ e $n_W^* S_W^* = n_W S_W + \sum_{t=2}^T (\underline{\theta}_t - G \underline{\theta}_{t-1})^T (\underline{\theta}_t - G \underline{\theta}_{t-1})$
- A distribuição condicional de $\Gamma = (\Gamma_1 \dots \Gamma_p)$ é $(\Gamma_j | \theta, V, W, \varphi) \sim N(m_\Gamma^*, S_\Gamma^*)$ para $j = 1, \dots, k$ onde $S_\Gamma^* = (S_\Gamma^{-1} I_{p \times p} + w_j^{-1} H_j^T H_j)^{-1}$ e $m_\Gamma^* = S_\Gamma^{-1} (S_\Gamma^{-1} I_{p \times p} \mathbf{1}_p + w_j^{-1} B_j^T H_j)$ para

$$H_j = \begin{pmatrix} \theta_{p,j} & \theta_{p-1,j} & \dots & \theta_{1,j} \\ \theta_{p+1,j} & \theta_{p,j} & \dots & \theta_{2,j} \\ \vdots & \vdots & \vdots & \vdots \\ \theta_{T-1,j} & \theta_{T-2,j} & \dots & \theta_{T-p,j} \end{pmatrix} \text{ e } B_j = \begin{pmatrix} \theta_{p+1,j} \\ \theta_{p+2,j} \\ \vdots \\ \theta_{T,j} \end{pmatrix}$$

- A distribuição condicional de φ é $(\varphi | \theta, V, W, \Gamma) \sim N(m_\varphi^*, S_\varphi^*)$ onde $S_\varphi^* = (S_\varphi^{-1} I_{4 \times 4} + V^{-1} Z^T Z)^{-1}$ e $m_\varphi^* = S_\varphi^{-1} (S_\varphi^{-1} I_{4 \times 4} \mathbf{1}_4 + V^{-1} Z^T (y - F \theta))$

Os estados $\theta_1, \dots, \theta_T$ são amostrados conjuntamente utilizando o algoritmo FFBS, condicional sobre Γ , φ , V e W a distribuição condicional completa conjunta segue a decomposição retrospectiva $p(\theta | y) = p(\theta_T | D_T) \prod_{t=1}^{T-1} p(\theta_t | \theta_{t+1}, D_t)$ onde

$D_T = \{y_1, y_2, \dots, y_T\}$, $t = 1, \dots, T$ e D_0 representa a informação inicial. Iniciando com $\theta_0 \sim N(m_0, C_0)$, através do Filtro de Kalman pode ser mostrado que:

$\theta_t | D_t \sim N(m_t, C_t)$, onde $m_t = a_{t-1} + A(y_t - f_t - z_t \varphi)$, $C_t = R_t - A_t Q_t A_t'$, $a_t = G m_{t-1}$, $R_t = G C_{t-1} G' + W$, $f_t = F_t' a_t + z_t \varphi$, $Q_t = F_t' R_t F_t + V$ e $A_t = R_t F_t Q_t^{-1}$, para $t = 1, \dots, T$;

θ_T é amostrado de $p(\theta_T | D_T)$ (passo denominado forward filtering). Para $t = T-1, T-2, \dots, 2, 1$, θ_t é amostrado de $p(\theta_t | \theta_{t+1}, D_t) = N(\tilde{a}_t, \tilde{C}_t)$ onde $\tilde{a}_t = m_t + B_t(f_{t+1} - a_{t+1})$, $\tilde{C}_t = C_t - B_t R_{t+1} B_t'$ e $B_t = C_t G' R_{t+1}^{-1}$ (passo denominado backward sampling).

A.2.3. Previsão Modelo de Regressão Linear Dinâmico

Fazer previsão nesta classe de modelos é teoricamente direto. Frequentemente surge o interesse de aprender da densidade preditiva h passos a frente $p(y_{T+h} | y)$, ou seja

$$p(y_{T+h} | y) = \int p(y_{T+h} | \underline{\theta}_{T+h}, \varphi, V) p(\underline{\theta}_{T+h} | \underline{\theta}_T, W, \Gamma) p(\underline{\theta}_T, W, \Gamma, V | y) d\underline{\theta}_{T+h} d\underline{\theta}_T d\varphi dV dW d\Gamma$$

onde $(y_{T+h} | \underline{\theta}_{T+h}, \varphi, V) \sim N(F_{T+h}, V)$, $(\underline{\theta}_{T+h} | \underline{\theta}_T, W, \Gamma) \sim N(\mu_h, V_h)$, $\mu_h = G^h \underline{\theta}_T$ e

$V_h = \sum_{j=1}^h G^{j-1} W (G^{j-1})'$, para $h > 0$. Portanto, se

$\{(\underline{\theta}_T^{(1)}, \varphi^{(1)}, V^{(1)}, W^{(1)}, \Gamma^{(1)}), \dots, (\underline{\theta}_T^{(L)}, \varphi^{(L)}, V^{(L)}, W^{(L)}, \Gamma^{(L)})\}$ é uma amostra de $p(\underline{\theta}_T, W, \Gamma, V | y)$, é fácil amostrar $\underline{\theta}_{T+h}^{(j)}$ de $p(\underline{\theta}_{T+h} | \underline{\theta}_T^{(j)}, W^{(j)}, \Gamma^{(j)})$, para todo $j = 1, \dots, L$, tal que

$$p(y_{T+h} | y) = \frac{1}{L} \sum_{j=1}^L p(y_{T+h} | \underline{\theta}_{T+h}^{(j)}, \varphi^{(j)}, V^{(j)})$$

é a aproximação via Monte Carlo de $p(y_{T+h} | y)$. Analogamente, se $y_{T+h}^{(j)}$ é amostrado de $p(y_{T+h} | \underline{\theta}_{T+h}^{(j)}, \varphi^{(j)}, V^{(j)})$, para $j = 1, \dots, L$, então $\{y_{T+h}^{(1)}, \dots, y_{T+h}^{(L)}\}$ representa uma amostra de $p(y_{T+h} | y)$.