

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE ECONOMIA E RELAÇÕES
INTERNACIONAIS**

Fernanda Cristina Valente

**PREVISÕES COM MODELOS DSGE-VAR:
UMA APLICAÇÃO PARA O BRASIL**

Florianópolis

2018

Fernanda Cristina Valente

**PREVISÕES COM MODELOS DSGE-VAR:
UMA APLICAÇÃO PARA O BRASIL**

Dissertação submetida ao Programa
de Pós Graduação em Economia para
a obtenção do Grau de Mestre.
Orientador: Prof. Dr. Guilherme Valle
Moura

Florianópolis

2018

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Valente, Fernanda Cristina
Previsões com Modelos DSGE-VAR : uma aplicação
para o Brasil / Fernanda Cristina Valente ;
orientador, Guilherme Valle Moura , 2018.
137 p.

Dissertação (mestrado) - Universidade Federal de
Santa Catarina, Centro Sócio-Econômico, Programa de
Pós-Graduação em Economia, Florianópolis, 2018.

Inclui referências.

1. Economia. 2. DSGE-VAR. 3. Previsão. 4.
Análise Bayesiana. I. , Guilherme Valle Moura. II.
Universidade Federal de Santa Catarina. Programa de
Pós-Graduação em Economia. III. Título.

Fernanda Cristina Valente

**PREVISÕES COM MODELOS DSGE-VAR: UMA
APLICAÇÃO PARA O BRASIL**

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestre”, e aprovada em sua forma final pelo Programa de Pós Graduação em Economia.

Florianópolis, 18 de janeiro 2018.

Prof. Dr. Jaylson Jair da Silveira
Coordenador
Universidade Federal de Santa Catarina

Banca Examinadora:

Guilherme Valle Moura, Dr.
Universidade Federal de Santa Catarina

Márcio Poletti Laurini, Dr.
Universidade de São Paulo

André Alves Portela Santos, Dr.
Universidade Federal de Santa Catarina

Maurício Simiano Nunes, Dr.
Universidade Federal de Santa Catarina

AGRADECIMENTOS

Agradeço,

Ao meu orientador, Guilherme Valle Moura, por todos os ensinamentos, incomparável paciência e dedicação.

Aos meus colegas do Mestrado, por toda ajuda, seja nos conhecimentos compartilhados ou nas boas conversas.

Aos meus pais, Eledir e Delson, e minha irmã Amanda, que não mediram esforços para me ajudar nesse caminho. Sou muito grata por toda a compreensão, carinho e apoio.

Ao Kael, pelo seu companheirismo e amor. Agradeço à sua compreensão, por entender as minhas ausências e principalmente por me incentivar.

À Olívia, Chico e Doris pela companhia durante as longas horas de trabalho.

Aos membros da Banca Examinadora, que dedicaram seu tempo para a leitura e avaliação deste trabalho.

À Capes pelo apoio financeiro, fundamental para que eu pudesse me dedicar exclusivamente aos meus estudos.

RESUMO

Este estudo propõe usar uma versão do modelo Dinâmico Estocástico de Equilíbrio Geral (DSGE) de pequena economia aberta desenvolvido por Gali e Monacelli (2005) como uma distribuição a priori na estimação de um modelo de Vetor Autorregressivo (VAR) para a economia brasileira, seguindo a abordagem de Del Negro e Schorfheide (2004) que deu origem aos chamados modelos DSGE-VAR. O principal objetivo é avaliar o desempenho das previsões geradas pelos modelos DSGE-VAR quando comparado com modelos como o Vetor Autorregressivo Bayesiano (BVAR) com a priori de Minnesota, BVAR com a priori difusa e DSGE estimado independentemente. Com intuito de comparar a acurácia das previsões geradas pelos modelos adotou-se o procedimento denominado *Model Confidence Set* (MCS), proposto por Hansen, Lunde e Nason (2011), que gera um subconjunto aleatório que contém o subconjunto dos “melhores” modelos dado um nível de confiança. Os resultados apontam que a inclusão de informações a priori oriundas do modelo DSGE mostram-se benéficas quando o objetivo é gerar previsões, especialmente para horizontes de previsão de médio e longo prazo. Segundo o procedimento MCS, o desempenho do DSGE-VAR mostrou-se superior aos modelos BVAR com a priori difusa e DSGE. Já em relação ao modelo BVAR com a priori de Minnesota, DSGE-VAR mostrou-se competitivo e, em algumas situações, superior.

Palavras-chave: DSGE-VAR. Previsões. Análise Bayesiana.

ABSTRACT

This study proposes to use a version of the small open economy Dynamic Stochastic General Equilibrium (DSGE) model developed by Galí e Monacelli (2005) to provide prior information in the estimation of an autoregression vector (VAR) model for Brazilian economy, following the Del Negro e Schorfheide (2004) approach that gave rise to so-called DSGE-VAR models. The main goal is to evaluate the performance of the predictions generated by the DSGE-VAR models when compared to models such as the Bayesian Vector Autoregression (BVAR) with the Minnesota prior, BVAR with the diffuse prior and independently estimated DSGE model. In order to compare the accuracy of the forecasts generated by the models, it was adopted the Model Confidence Set (MCS) procedure, proposed by Hansen, Lunde e Nason (2011). The MCS procedure yields a model confidence set, that is a collection of models built to contain the best models with a given level of confidence. The results indicates that imposing information coming from a small open economy DSGE model as a prior information in the estimation of a VAR helps improve the model's forecasting performance, especially in the medium and long term. According to the MCS procedure, the performance of the DSGE-VAR was superior to the BVAR models with the diffuse prior and DSGE model. In constrast to the BVAR model with the Minnesota prior, DSGE-VAR proved to be competitive and, in some situations, superior.

Keywords: DSGE-VAR. Forecasts. Bayesian Analysis.

LISTA DE FIGURAS

Figura 1	Funções Impulso-Resposta para um Choque Tecnológico	103
Figura 2	Funções Impulso-Resposta para um Choque Monetário.	105
Figura 3	Funções Impulso-Resposta para um Choque de Preferências	107
Figura 4	Funções Impulso-Resposta para um Choque no Produto do Resto do Mundo	109

LISTA DE TABELAS

Tabela 1	Distribuição a Priori para os Parâmetros do Modelo DSGE	86
Tabela 2	Densidade Marginal dos Dados	89
Tabela 3	Distribuições a Posteriori dos Parâmetros.....	91
Tabela 4	Ganhos (perdas) percentuais no MSE relativo ao BVAR com a priori difusa e \hat{p} -valores MCS para previsões 1 passo a frente	94
Tabela 5	Ganhos (perdas) percentuais no MSE relativo ao BVAR com a priori difusa e \hat{p} -valores MCS para previsões 3 passos a frente	97
Tabela 6	Ganhos (perdas) percentuais no MSE relativo ao BVAR com a priori difusa e \hat{p} -valores MCS para previsões 6 passos a frente	100
Tabela 7	Dados.....	133
Tabela 8	P-valor para a hipótese nula de convergência da cadeia de Markov	134
Tabela 9	Parâmetros Calibrados	135

SUMÁRIO

1	INTRODUÇÃO	17
Parte I	REVISÃO DE LITERATURA	21
2	O ATAQUE À MACROECONOMIA KEYNESI- ANA	23
2.1	MILTON FRIEDMAN E A CURVA DE PHILLIPS	23
2.2	ATAQUE NOVO CLÁSSICO	24
2.3	A SEGUNDA GERAÇÃO NOVO CLÁSSICA: O MO- DELO DE CICLOS REAIS DE NEGÓCIOS	26
3	OS NOVOS KEYNESIANOS E A MODELAGEM DSGE	29
4	MODELOS VAR E BVAR	33
5	MODELOS DSGE-VAR	35
Parte II	MODELO TEÓRICO E METODOLOGIA DE ESTIMAÇÃO	39
6	MODELO DSGE	41
6.1	A PEQUENA ECONOMIA ABERTA	41
6.1.1	Família Representativa	41
6.1.1.1	Inflação, Taxa de Câmbio Real e Termos de Troca	46
6.1.1.2	Diversificação Internacional de Risco	48
6.1.1.3	Paridade Descoberta dos Juros	48
6.1.2	Firmas	50
6.1.2.1	Tecnologia	50
6.1.2.2	Definição de Preços	50
6.2	EQUILÍBRIO	51
6.2.1	Demanda Agregada e Determinação do Produto ..	51
6.2.1.1	Consumo e Produto	51
6.2.1.2	A Balança Comercial	53
6.2.2	O Lado da Oferta	53
6.2.2.1	Produto Agregado e Emprego	53
6.2.2.2	Custo Marginal e Dinâmica Inflacionária	54
6.2.3	Equilíbrio Dinâmico	55
6.3	PEQUENA ECONOMIA ABERTA SOB UMA REGRA DE TAXA DE JUROS	55

7	METODOLOGIA DSGE-VAR	59
7.1	FUNÇÃO DE VEROSSIMILHANÇA	61
7.2	DISTRIBUIÇÃO A PRIORI	64
7.3	DISTRIBUIÇÃO A POSTERIORI	68
7.4	A ESCOLHA DE λ	69
7.5	ALGORITMO METROPOLIS-HASTINGS	70
7.6	IDENTIFICAÇÃO DOS CHOQUES	72
7.7	PREVISÕES FORA DA AMOSTRA	73

Parte III APLICAÇÃO EMPÍRICA E RESULTADOS **77**

8	IMPLEMENTAÇÃO PRÁTICA	79
8.1	DADOS E IMPLEMENTAÇÃO PRÁTICA	79
8.2	A PRIORI DE MINNESOTA	80
8.3	COMPARANDO PREVISÕES	82
8.4	DISTRIBUIÇÃO A PRIORI DE θ	86
8.5	ALGORITMO MCMC E DIAGNÓSTICO DE CONVERGÊNCIA	87
9	RESULTADOS	89
9.1	DETERMINAÇÃO DO MODELO	89
9.2	DISTRIBUIÇÃO A POSTERIORI DE θ	90
9.3	PREVISÕES	92
9.4	FUNÇÕES IMPULSO-RESPOSTA	102
10	CONCLUSÃO	111
	REFERÊNCIAS	113
	APÊNDICE A - DEMONSTRAÇÕES	121
	APÊNDICE B - TABELAS E RESULTADOS ...	133

1 INTRODUÇÃO

Desde o trabalho seminal de Sims (1980), o modelo de Vetor Autorregressivo (VAR) tornou-se uma ferramenta essencial para previsões em macroeconomia. O VAR é um modelo econométrico usado para capturar interdependência linear entre múltiplas séries de tempo. As suas grandes vantagens são a simplicidade e flexibilidade.

Entretanto, dois problemas relacionados aos modelos VAR são frequentemente mencionados na literatura: o primeiro diz respeito à ausência de interpretação estrutural (ou econômica) no caso do VAR em sua forma reduzida. Para superar isso, são necessárias algumas restrições de parâmetros, especialmente relacionadas à matriz de covariância dos erros, a fim de identificar os choques estruturais. O segundo problema está relacionado à parametrização excessiva (o que pode ocorrer devido ao número de variáveis endógenas incluídas ou devido à ordem de defasagem, ou ainda ambos) em comparação ao tamanho da amostra, que geralmente é limitado. Como consequência, os graus de liberdade tornam-se escassos, resultando em estimativas imprecisas e as previsões geralmente exibem comportamento errático e variabilidade excessiva.

Ao propor o modelo VAR Sims (1980) já externou o problema da parametrização excessiva enfrentada por essa classe, sugerindo como alternativa o uso de abordagem Bayesiana. A proposta foi originalmente implementada por Litterman (1979), dando origem ao modelo que ficou conhecido como Vetor Autorregressivo Bayesiano (BVAR). No método de estimação Bayesiano distribuições de probabilidade são usadas para resumir as crenças ou conhecimentos do pesquisador sobre cada parâmetro antes da observação dos dados, chamadas de distribuições a priori. Quando a informação a priori é combinada com a informação contida nos dados (a função de verossimilhança), uma distribuição atualizada é obtida, conhecida como distribuição a posteriori, que é utilizada para realizar inferências sobre os parâmetros de interesse, proporcionando uma maneira de introduzir o “encolhimento” do espaço de parâmetros, reduzindo o problema de parametrização excessiva do VAR.

Ainda que os modelos BVAR tenham ganhado posição de destaque na literatura macroeconômica, principalmente pelo fato de cumprir bem a tarefa de gerar previsões precisas, esses modelos são puramente estatísticos, ou seja, do ponto de vista de análise de política, esses modelos enfrentam problema de identificação. Sims (1980) define que um modelo é dito identificado se distintos pontos no espaço de parâmetros do modelo implicam em padrões observacionalmente distintos de

comportamento para as variáveis do modelo. Portanto, sob o objetivo de realizar análise de política ou inferência estrutural, é necessário obter a forma estrutural do modelo com base na sua forma reduzida. Geralmente isso ocorre somente sob a imposição de restrições, preferencialmente atreladas à teoria econômica. Por esse motivo, tem-se a necessidade da fundamentação teórica.

Paralelamente ao surgimento dos modelos VAR, a macroeconomia Keynesiana e os modelos de larga escala perderam força, dando lugar aos modelos Dinâmico Estocástico de Equilíbrio Geral (DSGE). Os modelos DSGE, decorrentes do trabalho de Kydland e Prescott (1982), surgiram de um esforço para derivar um modelo que fosse imune à crítica de Lucas (1976), isto é, contendo parâmetros que poderiam ser considerados independentes da política econômica. Devido à sua fundamentação teórica, esses modelos tornaram-se bastante populares nos dias de hoje.

Em contrapartida, DSGE impõe muitas restrições (*cross equations restrictions*) aos parâmetros, possivelmente em detrimento de um bom ajuste dos dados. Além disso, de acordo com Del Negro e Schorfheide (2006), dado que evidências empíricas mostram que as previsões de DSGE não são muito boas, muitos condutores de política econômica ainda são céticos de que esses modelos poderiam ser usados para previsões de curto e médio prazo, em razão de que são simplificações de economias complexas e podem sofrer de problemas de má especificação.

Del Negro e Schorfheide (2004) forneceram uma solução à esse dilema ao sugerirem unificar as duas abordagens supramencionadas, VAR e DSGE. A ideia, estendida dos trabalhos de DeJong, Ingram e Whiteman (1993) e Ingram e Whiteman (1994), propõe utilizar o arcabouço teórico dos modelos DSGE como informação a priori Bayesiana na estimação de um modelo VAR, dessa forma, sem permitir que as restrições do DSGE sejam impostas de forma dogmática, permitindo desvios no sentido de manter as propriedades dos dados. A abordagem gera um *continuum* de modelos empíricos denominados DSGE-VAR, variando de uma especificação que impõe categoricamente o modelo DSGE para uma que o ignora completamente.

Este estudo contribui com a literatura ao avaliar as previsões geradas a partir de um modelo DSGE-VAR, seguindo a abordagem de Del Negro e Schorfheide (2004), utilizando dados da economia brasileira. O modelo teórico DSGE segue o modelo de pequena economia aberta desenvolvido por Gali e Monacelli (2005). Nesse sentido, a principal questão é a seguinte: é possível combinar informações dos modelos DSGE

com dados da economia brasileira como forma de melhorar a previsão de agregados macroeconômicos? Portanto, o objetivo fundamental é investigar a magnitude do ganho na acurácia das previsões após a incorporação de informações a priori provenientes de um modelo DSGE na estimação de um VAR. Propõe-se comparar a acurácia das previsões geradas pelo DSGE-VAR com a acurácia das previsões de um BVAR com a priori difusa, um BVAR com a conhecida a priori de Minnesota e o modelo DSGE.

O restante deste trabalho está estruturado em 3 Partes: Parte I é composta por uma revisão de literatura que enfatiza: (i) O Ataque à Macroeconomia Keynesiana, Capítulo 2; (ii) os Novos Keynesianos e modelagem DSGE, Capítulo 3; (iii) Os Modelos VAR e BVAR, Capítulo 4; e (iv) Os Modelos DSGE-VAR, Capítulo 5. Parte II aborda: (i) O Modelo DSGE teórico, Capítulo 6; (ii) A Metodologia DSGE-VAR, Capítulo 7. Parte III discute a aplicação empírica e os resultados. Por fim, o Capítulo 10 expõe as conclusões.

Parte I

**REVISÃO DE
LITERATURA**

2 O ATAQUE À MACROECONOMIA KEYNESIANA

2.1 MILTON FRIEDMAN E A CURVA DE PHILLIPS

Durante o início da década de 1970, a economia mundial foi abalada por uma série de eventos traumáticos. Em especial, a economia dos Estados Unidos viu suas taxas de inflação aumentarem devido aos gastos com a Guerra do Vietnã e o primeiro choque do petróleo em 1973. Como resultado, a economia entrou em processo de estagflação. O cenário, que divergia daquele proposto pelo *trade-off* da curva de Phillips, acabou gerando insatisfação entre economistas e formuladores de política econômica em relação aos modelos macroeconômicos de larga escala, herança da macroeconomia Keynesiana.

A insistência Keynesiana que níveis de desemprego relativamente baixos são alcançáveis através do uso de políticas expansionistas de demanda agregada foi vigorosamente criticada por Milton Friedman, que acabou incentivando uma “contra-revolução” monetarista contra o ativismo político durante as décadas de 50 e 60 (SNOWDON; VANE, 2005).

Em seu trabalho “*The Role of Monetary Policy*”, Friedman (1968) buscou discutir o risco de se atribuir à política monetária um papel maior do que esta pode desempenhar, atacando uma das peças fundamentais da macroeconomia Keynesiana: a curva de Phillips. Para ele a política monetária é incapaz de manter as taxas de desemprego baixas por longos períodos, como sugeria a curva de Phillips ao explorar o *trade-off* entre inflação e desemprego.

Para motivar seu argumento, Friedman (1968) assume uma diferença na percepção entre firmas e trabalhadores. Enquanto as expectativas das firmas são corretas, as dos trabalhadores estão erradas. Em tal contexto, o crescimento na oferta de moeda é não-neutro, gerando deslocamento ao longo da curva de Phillips, porém apenas no curto prazo. Isso porque no próximo período os trabalhadores perceberão seus erros e integrarão o aumento no preço às suas expectativas. Com intuito de manter a taxa de emprego no mesmo nível, a oferta de moeda precisa ser expandida à taxas aceleradas, para que então os trabalhadores mais uma vez cometam erros de expectativa. Porém, se o processo continuar indefinidamente, a inflação se tornará uma hiperinflação, que é uma ameaça ao funcionamento do sistema monetário.

Enquanto a curva de Phillips no curto prazo possui inclinação negativa, no longo prazo ela é vertical ao nível de desemprego que

Friedman (1968) chamou de taxa natural de desemprego.

Segundo Friedman (1968) sempre existirá um *trade-off* temporário entre inflação e desemprego, mas não existe uma relação permanente. O *trade-off* temporário não vem da inflação por si, mas da inflação não antecipada, o que geralmente significa uma taxa de inflação crescente. Para ele, uma taxa crescente de inflação pode reduzir o desemprego, todavia uma taxa alta de inflação, não.

2.2 ATAQUE NOVO CLÁSSICO

Ainda durante a década de 70, outro grupo de economistas forneceu uma crítica muito mais prejudicial à macroeconomia Keynesiana. Seu principal argumento contra os Keynesianos foi que eles haviam falhado ao explorar as implicações de expectativas formadas endogenamente no comportamento dos agentes, sugerindo incorporar as expectativas aos modelos macroeconômicos através de alguma variante da “hipótese de expectativas racionais” de Muth (1961).

De acordo com Muth (1961), a hipótese de expectativas racionais implica que as expectativas subjetivas dos agentes econômicos são, em média, iguais aos valores reais das variáveis. Em outras palavras, trata-se apenas da média das previsões dos agentes econômicos, que será igual à expectativa matemática da variável.

Seguindo contribuições de nomes como Thomas Sargent e Robert E. Lucas Jr., os adeptos das “expectativas racionais” tornaram-se coletivamente conhecidos como escola novo clássica. Uma das principais contribuições de Lucas (1972) foi a inclusão do conceito de curva de oferta agregada “surpresa” de Lucas em um ambiente de informações incompletas. Esse conceito pode ser associado à uma curva de Phillips aumentada de expectativas racionais, representando uma revalidação dos resultados apontados por Friedman (1968).¹

Dessa forma, segundo Snowdon e Vane (2005), pode-se resumir as principais hipóteses da macroeconomia novo clássica envolvendo (i) a hipótese de expectativas racionais, (ii) *market clearing* contínuo e (iii) a hipótese de oferta agregada (“surpresa”).

A partir da contribuição de Lucas (1972), houve uma mudança na agenda de pesquisas da macroeconomia. A macroeconomia Key-

¹Embora as abordagens de Friedman (1968) e Lucas (1972) diferenciem-se em um ponto fundamental da metodologia (ver De Vroey (2001)), uma vez que a primeira segue o conceito Marshalliano de equilíbrio e a última utiliza a abordagem Walrasiana, atribui-se à Lucas (1972) os microfundamentos teóricos para as afirmações de ineficácia de política de Friedman (1968).

nesiana havia redirecionado os esforços da pesquisa para a questão da determinação do produto em um dado ponto do tempo. Com Lucas (1972), a preocupação com ciclos econômicos tomou lugar no topo da agenda de pesquisas.

Em seu trabalho, intitulado “*Understanding Business Cycles*”, Lucas (1977) forneceu uma explicação monetarista “novo clássica” dos ciclos econômicos como um fenômeno de equilíbrio. Em seu modelo, a variação da atividade ao longo do tempo ocorre em função de dois fatores: choques monetários exógenos e a hipótese de agentes com informações incompletas. Além disso, o autor incorpora a hipótese de substituição intertemporal que, segundo De Vroey e Malgrange (2011), Lucas considerou decisiva na explicação da variação do nível de atividade e, posteriormente, foi incorporada aos modelos DSGE. Ademais, nesse modelo, os choques monetários têm efeito real. Entretanto, como argumentado por Friedman (1968), o governo não pode explorá-los, uma vez que ocorrem somente sob surpresas monetárias.

A imagem dos ciclos econômicos como um fenômeno de equilíbrio, que emergiu das contribuições de Lucas, divergia da análise Keynesiana dos ciclos econômicos como fenômenos de desequilíbrio, isto é, a manifestação de uma falha de mercado. Para Lucas (1977), Keynes errou ao descartar o que chama de “disciplina de equilíbrio”, uma premissa básica imposta pela teoria econômica clássica, que consiste de dois postulados: (i) *market clearing* e (ii) os agentes são egoístas, de acordo com Lucas e Sargent (1981). Esses postulados, segundo De Vroey e Malgrange (2011), são considerados como um requisito universal, não estando ligados a propósitos específicos de modelos específicos. De acordo com Lucas (apud DE VROEY; MALGRANGE, 2011), ao faltar com a disciplina de equilíbrio, Keynes fez uma tentativa de explicar importantes aspectos de comportamento humano sem fazer referência ao que as pessoas gostam ou ao que elas são capazes de fazer.

Por outro lado, no plano metodológico, Lucas (1977) desenvolveu importantes avaliações críticas sobre a economia Keynesiana, distinta da economia de Keynes, reconhecendo a contribuição da macroeconomia Keynesiana no sentido de envolver-se em modelos econométricos e testes empíricos, em contraste com o próprio Keynes.

Todavia, Lucas (1976), em sua famosa “Crítica de Lucas”, argumentou que a ausência de uma abordagem baseada em fundamentação teórica para o desenvolvimento das equações estruturais significava que os coeficientes dos modelos estimados provavelmente não seriam invariáveis às mudanças em regimes de políticas ou para outros tipos de alterações estruturais.

Os modelos econométricos de equações simultâneas, herança da macroeconomia Keynesiana e utilizados para avaliar as mudanças de política e gerar previsões, têm como principal problema o fato de que os resíduos são, em geral, correlacionados com as variáveis endógenas, quando o modelo está em sua forma estrutural. Para solucionar este problema, as equações podem ser diminuídas a forma reduzida. Se o objetivo do pesquisador é apenas gerar previsões incondicionais, as equações em forma reduzida são suficientes para resolver esse problema. Entretanto, se o objetivo for entender como a economia reage às mudanças de política, deve-se recuperar a forma estrutural do modelo.

Todavia, geralmente não é possível encontrar estimativas únicas dos parâmetros estruturais com base somente nos parâmetros da forma reduzida, isto é, enfrenta-se o problema de identificação. Portanto, para encontrar os parâmetros estruturais com base nos parâmetros da forma reduzida, é necessário ter algum conhecimento prévio a respeito dos parâmetros de maneira a impor restrições.

A crítica de Lucas (1976) volta-se justamente à forma como os Keynesianos lidam com o problema de identificação. Lucas e Sargent (1981) apontam as principais restrições impostas nos modelos macroeconômicos Keynesianos com o objetivo de identificar os parâmetros estruturais: (a) definição a priori de variáveis iguais a zero; (b) restrição da ordem de correlação serial ρ ; (c) classificação a priori das variáveis em endógenas e exógenas.

O argumento de Lucas (1976) é que a ausência de microfundamentação e teoria de equilíbrio geral na solução Keynesiana do problema de identificação não garante que o modelo levará em consideração o fato de que os agentes poderão mudar suas decisões frente às mudanças de política, o que leva ao descolamento entre a teoria econômica e a econometria discutido pelo autor.

2.3 A SEGUNDA GERAÇÃO NOVO CLÁSSICA: O MODELO DE CICLOS REAIS DE NEGÓCIOS

Durante a década de 80, a explicação novo clássica da instabilidade dos agregados atentou aos choques reais ao invés dos choques monetários, tornando-se conhecida como teoria dos ciclos reais de negócios. Uma nova geração de economistas ligados à abordagem novo clássica, que adotou as mesmas hipóteses de expectativas racionais e *market clearing*, passou a construir modelos onde as flutuações econômicas são resultados de choques reais na economia.

Os trabalhos de Kydland e Prescott (1982) e Long e Plosser (1983) são dois estudos que iniciaram a linha de pesquisa dos ciclos reais de negócios. Ambos tentaram modelar ciclos econômicos como resultado de choques reais na economia.

Uma implicação básica desse programa de pesquisa é que ele começou a partir da posição de que crescimento e flutuações não são vistos como fenômenos distintos para serem estudados com dados e ferramentas analíticas diferentes. De acordo com Snowdon e Vane (2005) o modelo de ciclo reais de negócios, seguindo Frisch (1933), faz distinção entre os mecanismos de impulso e propagação. Um mecanismo de impulso é um choque inicial que causa a uma variável um desvio do seu valor de estado estacionário. Um mecanismo de propagação consiste na força que leva o efeito do choque ao longo do tempo e causa um desvio persistente do estado estacionário.

O modelo de Kydland e Prescott (1982) tinha como características comuns ao modelo de Lucas (1977) a “disciplina de equilíbrio”, a hipótese de expectativas racionais, um ambiente dinâmico-estocástico e o fenômeno de substituição intertemporal. Entretanto, estes dois modelos divergem uma vez que o modelo de Kydland e Prescott (1982) assume choques reais, abandona a hipótese de informações incompletas e trata-se de um modelo quantitativo. É nessa última característica que reside a sua grande importância e o que lhe atribuiu a inauguração de uma nova metodologia, conforme Woodford (1999).

O ponto de partida de Kydland e Prescott (1982) foi um modelo de crescimento neoclássico e o objetivo era mostrar que flutuações econômicas podem ser explicadas como uma consequência do ajustamento de otimização dos agentes econômicos aos choques tecnológicos exógenos. Uma vez que o objetivo proposto por Kydland e Prescott (1982) era avaliar o desenvolvimento da economia dos Estados Unidos entre 1950 e 1970, a validação do modelo ocorre comparando os momentos simulados com os momentos reais da economia. Como resultado, o modelo reproduz satisfatoriamente o comportamento de algumas variáveis, a despeito de outras, cujo resultado mostrou-se pouco satisfatório.

No geral, os resultados de Kydland e Prescott (1982) são impressionantes. Eles conseguem replicar com sucesso várias características importantes das flutuações na economia dos EUA durante um longo período de tempo, com base em um modelo extremamente simples.

Contudo, os modelos de ciclos reais de negócios ainda foram alvos de muitas críticas, entre as quais destaca-se a implicação controversa de neutralidade da moeda. Kydland e Prescott argumentaram que os choques monetários desempenhavam um papel pequeno na explicação dos

ciclos econômicos. Todavia, a neutralidade da moeda é uma hipótese irrealista, uma vez que vai em desacordo à uma crença comum entre economistas que política monetária tem poder de influenciar variáveis reais, pelo menos no curto prazo (ver Friedman e Schwartz (1963), por exemplo). Além disso, essa hipótese torna esses modelos sem utilidade para bancos centrais, visto que sob neutralidade da moeda, essas instituições não teriam utilidade.

3 OS NOVOS KEYNESIANOS E A MODELAGEM DSGE

Ainda que muitos economistas aceitassem as críticas de Lucas, eles gostariam de reabilitar as ideias de Keynes, mas em acordo com os princípios centrais da macroeconomia novo clássica. Essa era a opinião do grupo de economistas que ficou conhecido como “novos Keynesianos”, que surgiu em meados da década de 70 (DE VROEY; MALGRANGE, 2011).

A macroeconomia novo Keynesiana desenvolveu-se em resposta à crise teórica percebida dentro da economia Keynesiana que havia sido exposta por Lucas durante a década de 70. A tarefa primordial que enfrentaram os teóricos novo Keynesianos era remediar as inconsistências no antigo modelo Keynesiano.

Os novo Keynesianos desenvolveram seus modelos em diferentes direções, apesar de compartilharem o mesmo propósito de alterar, se não reverter, as conclusões novo clássicas sobre o sistema de mercado, revivendo a visão de Keynes, afinal a hipótese de *market clearing* contínuo parecia incomodá-los.

Segundo Snowdon e Vane (2005), os economistas novo Keynesianos eram um grupo extremamente heterogêneo, o que faz o uso do termo “escola” novo Keynesiana ser muito mais conveniente do que apropriado. Entre os economistas que fizeram significativas contribuições para a literatura novo Keynesiana, mesmo que alguns deles se oponham a essa classificação, pode-se citar Gregory Mankiw, Lawrence Summers, Olivier Blanchard, Stanley Fischer, Bruce Greenwald, Edmund Phelps, Joseph Stiglitz, Ben Bernanke, Laurence Ball, George Akerlof, David Romer, Janet Yellen, Robert Hall, John Taylor, Dennis Snower e Assar Lindbeck.

Os modelos novo Keynesianos, de acordo com De Vroey e Malgrange (2011), eram tão conceitualmente inovadores e tecnicamente inteligentes quanto os modelos novo clássicos que desejavam refutar. No entanto, eles não conseguiram alterar o novo curso da macroeconomia que Lucas havia iniciado durante a década de 70. Além disso, a maioria dos modelos novo Keynesianos operava dentro de uma estrutura estática enquanto a perspectiva dinâmica estocástica estava se tornando cada vez mais dominante.

Em meados da década de 90 viu-se um declínio dos modelos de ciclos reais de negócios e o surgimento concomitante de um novo tipo de modelo que colocaria fim à batalha metodológica entre novos

Keynesianos e os adeptos dos modelos de ciclo real de negócios: o modelo dinâmico e estocástico de equilíbrio geral, o DSGE.

A contribuição Keynesiana na construção do DSGE foi a introdução de competição imperfeita e rigidez de preços, bem como um foco no papel do banco central. Por outro lado, eles aceitaram os componentes básicos dos modelos de ciclos reais de negócios: choques exógenos, ambiente dinâmico-estocástico, o equilíbrio, substituição intertemporal e expectativas racionais. Esse novo paradigma foi chamado de “nova síntese neoclássica” por muitos autores, como Goodfriend e King (1997), por exemplo.

A relação entre os modelos de ciclos reais de negócios e DSGE pode ser explicada pelo fato do último usar frequentemente o equilíbrio do primeiro como referência para medir a ineficácia decorrente da rigidez nominal e da concorrência monopolística. Na verdade, alguns modelos DSGE definem o hiato do produto como a diferença entre o nível real de produto e o nível natural, sendo que o último é definido como um equilíbrio do modelo de ciclos reais de negócios correspondente. Portanto, os modelos de ciclos reais de negócios podem ser vistos como um ponto de partida para DSGE.

O desenvolvimento de modelos DSGE está intimamente ligado ao esforço para derivar um modelo que fosse imune à crítica de Lucas, ou seja, contendo parâmetros que poderiam ser considerados independentes da política econômica. São modelos microfundamentados e seus parâmetros são funções dos chamados “*deep parameters*”, que são relacionados a preferência, tecnologia e restrições de recursos e podem ser considerados independentes da política monetária.

A competição monopolística foi integrada à modelagem DSGE através do trabalho de Dixit e Stiglitz (1977), enquanto que para modelar rigidez nominal uma abordagem comum na literatura deriva de Calvo (1983).

Calvo (1983) desenvolveu um modelo em que supõe que, a cada período, as firmas estão autorizadas a alterar seus preços, o que ocorre com uma dada probabilidade. Sob essa hipótese, quando a firma pode reajustar seu preço de forma ótima, ela leva em consideração o fato de que não poderá fazê-lo novamente durante um período de tempo. Portanto, as firmas escolhem seus preços para maximizar a soma de seus lucros atuais e futuros, dada a probabilidade de seus preços não serem alterados nos períodos subsequentes. O problema de otimização de preços enfrentados pelas firmas leva, após certas manipulações matemáticas, à derivação de uma das equações chave dos modelos DSGE: a curva de Phillips novo Keynesiana.

Outro desenvolvimento em relação ao DGSE está ligado à política monetária, mais especificamente às regras que os bancos centrais devem seguir. Agora a taxa de juros (não mais a quantidade de moeda) é a variável de controle. Nesse desenvolvimento, dois economistas desempenharam importantes papéis: Taylor e Woodford.

A “regra de Taylor” originou-se do artigo de Taylor (1993), que tentou fornecer uma avaliação empírica da política do Banco Central Americano (FED). A regra consiste em corrigir a taxa de juros levando-se em conta três objetivos: (i) estabilidade de preços, medida pela diferença entre a taxa de inflação observada e a meta de inflação; (ii) o hiato do produto, que é o desvio entre o resultado efetivo potencial; e (iii) um choque de política econômica, que é um choque puramente residual não correlacionado com a inflação ou com o produto. Woodford seguiu a mesma ideia de Taylor, oferecendo diversas contribuições à literatura, ao passo que tentou estabelecer um vínculo completo entre estabilização macroeconômica e bem-estar econômico.

O próximo passo dado em direção ao desenvolvimento do DSGE está ligado ao método estimação. Smets e Wouters (2003) baseando-se no modelo desenvolvido por Christiano, Eichenbaum e Evans (2005), estimaram um modelo DSGE para a zona do euro. A principal contribuição desse trabalho foi o uso de métodos de estimativa Bayesiana para modelos DSGE. A partir disso, bancos centrais em todo o mundo adotaram o modelo de Smets e Wouters (2003) para análise de política e previsões.

Apesar da popularidade crescente dos modelos DSGE, eles ainda são alvos de certa desconfiança. A estrutura teórica dessa classe de modelos impõe muitas restrições - *cross equations restrictions* - aos parâmetros, possivelmente em detrimento de um bom ajuste dos dados que potencialmente geram distorções às previsões. Segundo Del Negro e Schorfheide (2006), a existência de evidências empíricas que mostram que as previsões de DSGE não são muito boas faz com que muitos condutores de política econômica ainda sejam céticos que esses modelos poderiam ser usados para previsões de curto e médio prazo.

4 MODELOS VAR E BVAR

Assim como Lucas (1976), durante a década de 80 Sims (1980) argumentou que as restrições adotadas pelos macroeconomistas Keynesianos, usadas para resolver o problema de identificação dos modelos de larga escala, não são realistas. Especificamente, Sims (1980) criticou a forma como os Keynesianos lidavam com o problema da identificação em relação à classificação a priori das variáveis endógenas e exógenas. Como alternativa, ele propõe que as variáveis não sejam mais divididas entre endógenas e exógenas com base em considerações a priori.

De acordo com Lucas e Sargent (1981), em geral, as variáveis eram classificadas como endógenas por serem amplamente determinadas por ações de agentes privados (como consumo ou gasto com investimento privado) e variáveis exógenas eram aquelas sob controle do governo. Essa divisão tinha a intenção de refletir o significado habitual das palavras *endógeno* - “determinado pelo sistema [econômico]” - e *exógeno* - “afeta o sistema [econômico], mas não é afetado por ele.”

Ao propor que todas as variáveis fossem classificadas como endógenas, Sims (1980) desenvolveu a abordagem de vetores autorregressivos, popularmente conhecido como VAR, que tornou-se uma ferramenta essencial em macroeconomia e é utilizada para descrever a interdependência linear entre múltiplas séries de tempo. As suas grandes vantagens são a simplicidade e flexibilidade.

O modelo de Vetor Autorregressivo Bayesiano (BVAR) veio do argumento de Litterman (1979) que a precisão das estimativas e o desempenho das previsões pode ser melhorado ao incorporar restrições na forma de distribuições a priori nos parâmetros. A formulação a priori de Litterman (1979) é baseada essencialmente nos fatos estilizados dos seus dados que eram variáveis macroeconômicas dos Estados Unidos, e que podem ser bem caracterizadas por processos de raiz unitária. Esta imposição vai ao encontro aos resultados de Nelson e Plosser (1982), que indicam que séries macroeconômicas são consistentes com a hipótese de passeio aleatório. Posto isso, Litterman (1979) propôs “encolher” cada uma das variáveis do VAR em direção a um passeio aleatório univariado.

A partir dos estudos de Doan, Litterman e Sims (1984) e Litterman (1986), essa especificação de informação a priori baseado na ideia de que séries macroeconômicas seguem passeios aleatórios e de que modelos parcimoniosos preveem melhor tornou-se popularmente conhecida como a priori de Minnesota. A origem do nome decorre do fato que

o início dos trabalhos com VAR Bayesiano foi realizado por pesquisadores da Universidade de Minnesota ou no *Federal Reserve Bank of Minneapolis*.

O BVAR possui certa vantagem sobre o VAR no sentido de que reduz o problema da parametrização excessiva, uma vez que “encolhe” os parâmetros em direção a um ponto específico no espaço paramétrico. Embora a priori de Minnesota tenha se mostrado útil em relação a previsibilidade dos modelos VAR, trata-se de um modelo puramente estatístico, isto é, sua estrutura é privada de teoria econômica.

No geral, quando o objetivo é gerar previsões incondicionais, empiricamente BVARs têm sucesso expressivo (por exemplo Litterman (1986) e Kadiyala e Karlsson (1993)). Inferência estrutural e análise de política são, no entanto, inerentemente mais difíceis porque enfrentam o problema da identificação. Isto é, se a finalidade do pesquisador é realizar análise de política ou inferência estrutural, a forma estrutural do modelo precisará ser recuperada com base na sua forma reduzida. Contudo, o problema de identificação reside justamente no fato que geralmente não é possível encontrar estimativas únicas dos parâmetros estruturais com base somente nos parâmetros da forma reduzida. Neste caso, é necessário impor algum tipo de restrição, preferencialmente atrelada à teoria econômica, ou seja, esse problema não pode ser resolvido com base em ferramentas puramente estatísticas. Em vez disso, utilizar o arcabouço da teoria econômica faz-se necessário.

5 MODELOS DSGE-VAR

Em face das vantagens e desvantagens aludidas a respeito dos modelos VAR e DSGE, DeJong, Ingram e Whiteman (1993) e Ingram e Whiteman (1994) propuseram uma metodologia de estimação que unifica essas abordagens, utilizando o melhor de cada uma delas. Intuitivamente, a ideia consiste em utilizar o arcabouço teórico dos modelos DSGE na estimação do VAR, no entanto, sem permitir que as restrições do DSGE sejam impostas de forma categórica. A ideia é implementada usando abordagem Bayesiana, isto é, ela consiste em utilizar o modelo teórico como uma informação a priori na estimação de um modelo VAR.

Em seu estudo, DeJong, Ingram e Whiteman (1993) analisaram funções impulso resposta geradas por um modelo VAR estimado sujeito às restrições impostas por um modelo de equilíbrio geral monetário. Ingram e Whiteman (1994), por sua vez, mostraram que informações a priori oriundas de um modelo de ciclos reais de negócios ajudam a melhorar a precisão das previsões no caso dos movimentos de consumo, produto, horas trabalhadas e investimentos para a economia dos Estados Unidos.

Mais recentemente, Del Negro e Schorfheide (2004) deram origem aos modelos denominados DSGE-VAR ao desenvolverem uma extensão dos trabalhos supracitados: primeiro, eles mostraram formalmente como inferências a posteriori para os parâmetros do VAR podem ser traduzidas em inferências a posteriori para os parâmetros do DSGE; segundo, eles construíram um esquema de identificação para os choques estruturais do VAR, baseado na comparação das respostas contemporâneas do VAR aos choques com as respostas do DSGE; por fim, eles ilustraram como um VAR, com informações a priori de um DSGE, pode ser usado para prever os efeitos de mudanças permanentes nas regras de política.

A abordagem de Del Negro e Schorfheide (2004) dá origem a um *continuum* de modelos empíricos que variam de acordo com o peso relativo das informações a priori do DSGE à amostra. Se por um lado toda a informação a priori é ignorada, o modelo DSGE-VAR colapsa em um modelo BVAR com a priori difusa. Por outro lado, quando o peso atribuído às informações a priori tende ao infinito, o modelo DSGE-VAR tenderá ao modelo DSGE.

Posteriormente, Del Negro e Schorfheide (2009), mostraram como a abordagem DSGE-VAR pode ser utilizada para realizar inferências sobre os parâmetros de modelos sob má especificação e realizar análise

de política monetária com modelos potencialmente mal especificados.

Frente aos bons resultados obtidos com a abordagem de Del Negro e Schorfheide (2004), uma série de estudos utilizando modelos DSGE-VAR foram desenvolvidos em diferentes países. Por exemplo, Watanabe (2007) estimou um modelo DSGE-VAR utilizando variáveis macroeconômicas do Japão. O modelo DSGE utilizado como a priori Bayesiana na estimação do VAR foi baseado no modelo desenvolvido por Christiano, Eichenbaum e Evans (2005). Ao comparar as verossimilhanças marginais dos modelos DSGE-VAR com diferentes razões entre as informações amostradas e informações simuladas, Watanabe (2007) encontrou evidências de má especificação no modelo DSGE ainda que as informações a priori fornecidas pelo DSGE tenham se mostrado úteis.

Lees, Matheson e Smith (2011) aplicaram a metodologia de Del Negro e Schorfheide (2004) ao utilizar um modelo DSGE de pequena economia aberta com meta de inflação explícita como informação a priori na estimação de um VAR com o objetivo de gerar previsões para a Nova Zelândia. Entre os resultados obtidos está a performance superior do modelo DSGE-VAR, em termos de previsão, em relação às previsões oficiais do Banco Central da Nova Zelândia. Além disso, o desempenho do modelo DSGE também mostrou-se competitivo. Por fim, os modelos DSGE-VAR e DSGE aproximam-se, mas não conseguem bater, o desempenho das previsões geradas pelo BVAR com a priori de Minnesota.

Na África do Sul, Gupta e Steinbach (2013) utilizaram a abordagem DSGE-VAR adotando um modelo DSGE de pequena economia aberta com dados da África do Sul e dos Estados Unidos, para gerar previsões e avaliar sua acurácia frente a outros modelos de previsão. O desempenho do modelo DSGE-VAR desenvolvido por Gupta e Steinbach (2013) foi comparado à um modelo DSGE estimado independentemente, um modelo VAR clássico, e modelos BVAR com seis diferentes a priori. Os resultados apontam que o DSGE-VAR apresenta um desempenho competitivo, se não, melhor do que os modelos VAR avaliados, para previsões trimestrais de 1 a 8 passos a frente. No entanto, não há evidências de melhoras no desempenho das previsões ao mover-se para a estrutura DSGE-VAR, a partir de um DSGE estimado de forma independente.

No que tange a literatura brasileira, ainda existem poucos trabalhos que utilizam a abordagem DSGE-VAR. Costa, Porto Junior e Almeida (2017) baseiam-se em um modelo de economia aberta desenvolvido por Gali e Monacelli (2005), utilizando a abordagem DSGE-VAR para avaliar a reação do Banco Central do Brasil (BCB) frente

às oscilações na taxa de câmbio, para um caso de uma economia sob metas de inflação. Dentre os resultados, Costa, Porto Junior e Almeida (2017) encontram evidências que o modelo DSGE-VAR apresenta um ajuste consideravelmente melhor que o modelo DSGE, independente do número de defasagens incluídas no modelo VAR, indicando que as restrições do modelo DSGE não ajustam-se bem aos dados. O presente trabalho diferencia-se dos demais estudos encontrados na literatura brasileira uma vez que propõe gerar previsões a partir dos modelos DSGE-VAR e comparar sua acurácia frente a modelos de previsão alternativos, como o BVAR com a priori de Minnesota e o modelo DSGE estimado de forma independente.

Parte II

MODELO TEÓRICO E METODOLOGIA DE ESTIMAÇÃO

6 MODELO DSGE

O modelo teórico é baseado no modelo desenvolvido por Gali e Monacelli (2005). Isto é, existe um *continuum* de pequenas economias abertas representadas por um intervalo unitário. As decisões de política em cada economia não têm impacto no resto do mundo. Além disso, enquanto diferentes economias estão sujeitas a choques tecnológicos imperfeitamente correlacionados, assume-se que essas economias compartilham preferências, tecnologias e estrutura de mercado idênticas. Há discriminação entre bens domésticos e estrangeiros, ainda que todos os bens possam ser trocados internacionalmente e não há, na economia mundial, política de coordenação internacional.

6.1 A PEQUENA ECONOMIA ABERTA

A pequena economia aberta do modelo consiste em uma família representativa que busca maximizar sua função de utilidade, um *continuum* de firmas produzindo bens diferenciados, onde cada firma opera em um ambiente monopolisticamente competitivo com rigidez de preços à la Calvo e, por fim, a política monetária é especificada por uma regra de Taylor, com produto e inflação como argumentos na regra de política.

6.1.1 Família Representativa

A pequena economia aberta é habitada por uma família representativa que busca maximizar

$$E_0 \sum_{t=0}^{\infty} \beta^t U(C_t, N_t; Z_t) \quad (6.1)$$

em que β é um fator de desconto, N_t denota as horas de trabalho, C_t é um índice de consumo composto e $z_t = \log Z_t$ segue um processo AR(1) estacionário,

$$z_t = \rho_z z_{t-1} + \epsilon_t^z, \quad (6.2)$$

no qual ϵ_t^z é um ruído branco e pode ser interpretado como um choque de preferências. O índice de consumo composto é definido por

$$C_t \equiv [(1 - \alpha)^{\frac{1}{\eta}} (C_{H,t})^{\frac{\eta-1}{\eta}} + \alpha^{\frac{1}{\eta}} (C_{F,t})^{\frac{\eta-1}{\eta}}]^{\frac{\eta}{\eta-1}} \quad (6.3)$$

onde $C_{H,t}$ é um índice de consumo de bens domésticos dado por função elasticidade de substituição constante (CES)

$$C_{H,t} \equiv \left(\int_0^1 C_{H,j}(j)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}} \quad (6.4)$$

em que $j \in [0, 1]$ é a variedade de bens. $C_{F,t}$ é um índice de bens importados,

$$C_{F,t} \equiv \left(\int_0^1 C_{i,t}(i)^{\frac{\gamma-1}{\gamma}} di \right)^{\frac{\gamma}{\gamma-1}}, \quad (6.5)$$

em que $C_{i,t}$ é, por sua vez, um índice da quantidade de bens importados do país i e consumido pela família representativa na economia doméstica,

$$C_{i,t} \equiv \left(\int_0^1 C_{i,j}(j)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}}. \quad (6.6)$$

Note que o parâmetro $\epsilon > 1$ denota a elasticidade de substituição entre as variedades produzidas dentro de um dado país. O parâmetro $\alpha \in [0, 1]$ pode ser interpretado como o grau de abertura da economia. O parâmetro $\eta > 0$ mede a substituição entre os bens domésticos e importados, a partir do ponto de vista do consumidor doméstico, enquanto γ mede a substituição entre os diferentes bens produzidos em diferentes países estrangeiros.

A maximização de (6.1) é sujeita a uma sequência de restrições orçamentárias com a seguinte forma

$$\begin{aligned} \int_0^1 P_{H,t}(j) C_{H,t}(j) dj + \int_0^1 \int_0^1 P_{i,t}(j) C_{i,t}(j) dj di \\ + E_t \{ Q_{t,t+1} D_{t+1} \} \leq D_t + W_t N_t + T_t \end{aligned} \quad (6.7)$$

para $t = 0, 1, 2, \dots$ onde $P_{H,t}(j)$ é o preço da variedade doméstica j . $P_{i,t}(j)$ é o preço da variedade j importada do país i . D_{t+1} é o retorno nominal no período $t+1$ de um portfolio mantido até o final do período t , W_t é o salário nominal e T_t denota uma taxa *lump-sum*. Por fim, $Q_{t,t+1}$ é um fator de desconto estocástico do retorno nominal para um período a frente. Todas as variáveis são expressas em unidades de

moeda doméstica corrente.

A alocação ótima de uma determinada despesa dentro de cada categoria de bens produz funções demanda

$$C_{H,t}(j) = \left(\frac{P_{H,t}(j)}{P_{H,t}} \right)^{-\epsilon} C_{H,t} \quad ; \quad C_{i,t}(j) = \left(\frac{P_{i,t}(j)}{P_{i,t}} \right)^{-\epsilon} C_{i,t} \quad (6.8)$$

para todo $i, j \in [0, 1]$, onde $P_{H,t} \equiv \left(\int_0^1 P_{H,t}(j)^{1-\epsilon} dj \right)^{\frac{1}{1-\epsilon}}$ é o índice de preços doméstico e $P_{i,t} \equiv \left(\int_0^1 P_{i,t}(j)^{1-\epsilon} dj \right)^{\frac{1}{1-\epsilon}}$ é o índice de preços para bens importados do país i (expresso em moeda corrente doméstica) para todo $i \in [0, 1]$. Combinando as alocações ótimas em (6.8) com as definições dos índices de preços e quantidades $P_{H,t}$, $C_{H,t}$, $P_{i,t}$ e $C_{i,t}$, obtêm-se

$$\int_0^1 P_{H,t}(j) C_{H,t}(j) dj = P_{H,t} C_{H,t} \quad (6.9)$$

$$\int_0^1 P_{i,t}(j) C_{i,t}(j) dj = P_{i,t} C_{i,t}. \quad (6.10)$$

Além disso, a alocação ótima das despesas em bens importados por país de origem implica

$$C_{i,t} = \left(\frac{P_{i,t}}{P_{F,t}} \right)^{-\gamma} C_{F,t} \quad (6.11)$$

para todo $i \in [0, 1]$ onde $P_{F,t} \equiv \left(\int_0^1 P_{i,t}^{1-\gamma} di \right)^{\frac{1}{1-\gamma}}$ é o índice de preços para bens importados, também expressos em moeda corrente. Usando as definições de $P_{F,t}$ e $C_{F,t}$ combinadas com (6.11), obtêm-se

$$\int_0^1 P_{i,t} C_{i,t} di = P_{F,t} C_{F,t}. \quad (6.12)$$

Por fim, a alocação ótima das despesas entre bens domésticos e importados é dada por

$$C_{H,t} = (1 - \alpha) \left(\frac{P_{H,t}}{P_t} \right)^{-\eta} C_t \quad (6.13)$$

$$C_{F,t} = \alpha \left(\frac{P_{F,t}}{P_t} \right)^{-\eta} C_t \quad (6.14)$$

onde $P_t \equiv [(1 - \alpha)(P_{H,t})^{1-\eta} + \alpha(P_{F,t})^{1-\eta}]^{\frac{1}{1-\eta}}$ é o índice de preços ao

consumidor (IPC).

Portanto, o total de despesas com consumo da família doméstica é dado por $P_{H,t}C_{H,t} + P_{F,t}C_{F,t} = P_t C_t$. Sendo assim, a restrição orçamentária pode ser reescrita como

$$P_t C_t + E_t\{Q_{t,t+1}D_{t+1}\} \leq D_t + W_t N_t + T_t. \quad (6.15)$$

Adicionalmente, tem-se a condição de transversalidade:

$$\lim_{T \rightarrow \infty} E_t\{D_T\} \geq 0, \quad (6.16)$$

que impedirá a família de envolver-se em esquemas do tipo Ponzi.

A função de utilidade é especificada de acordo com

$$U(C_t, N_t; Z_t) = \left(\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi} \right) Z_t \quad (6.17)$$

onde σ é a elasticidade substituição intertemporal e φ representa a elasticidade inversa da oferta de trabalho.

A partir das condições de otimalidade resultantes da maximização de (6.17) sujeito a (6.15), obtém-se

$$C_t^\sigma N_t^\varphi = \frac{W_t}{P_t}. \quad (6.18)$$

Com a finalidade de derivar a equação de Euler, note que a restrição orçamentária em (6.15) pode ser reescrita, assumindo que a igualmente é mantida, como

$$P_t C_t = D_t + W_t N_t + T_t - E_t\{Q_{t,t+1}D_{t+1}\} \quad (6.19)$$

onde o lado esquerdo representa as despesas com consumo no período t , o termo $D_t + W_t N_t + T_t$ no lado direito representa a renda bruta disponível no período t , enquanto $E_t\{Q_{t,t+1}D_{t+1}\}$ é o investimento no tempo t em um portfólio com retorno nominal D_{t+1} no período $t + 1$. Portanto, a equação (6.19) diz que qualquer renda deixada após o investimento é usada para consumo. O problema intertemporal para a família representativa em relação à escolha ótima da carteira de um período pode ser escrito como

$$\max_{D_{t+1}} \{U(C_t, N_t; Z_t) + E_t\{\beta U(C_{t+1}, N_{t+1}; Z_{t+1})\}\} \quad (6.20)$$

sujeito a

$$P_t C_t = D_t + W_t N_t + T_t - E_t \left\{ \int V_{t,t+1} D_{t+1} d\tau \right\} \quad (6.21)$$

$$E_t \{ P_{t+1} C_{t+1} \} = E_t \left\{ \int \xi_{t,t+1} D_{t+1} d\tau + W_{t+1} N_{t+1} + T_{t+1} - V_{t,t+2} D_{t+2} \right\} \quad (6.22)$$

onde $E_t \{ \int V_{t,t+1} D_{t+1} d\tau \}$ é o preço de mercado do portfolio de um período que produz um retorno aleatório D_{t+1} , onde integrou-se todos os possíveis estados da natureza indexado por τ . $V_{t,t+1}$ é um seguro de um período que retorna uma unidade de moeda corrente se um específico estado da natureza é realizado no período $t + 1$, e zero caso contrário. $\xi_{t,t+1}$ é a probabilidade de que um dado estado da natureza se realize no período $t + 1$. De forma equivalente, o preço pode ser escrito como $E_t \{ \frac{V_{t,t+1}}{\xi_{t,t+1}} D_{t+1} \}$. Portanto, o fator de desconto estocástico pode ser definido como

$$Q_{t,t+1} \equiv \frac{V_{t,t+1}}{\xi_{t,t+1}}. \quad (6.23)$$

A partir das condições de primeira ordem do problema de maximização em (6.20) sujeito a (6.21) e (6.22), e da relação em (6.23), obtém-se

$$\begin{aligned} \frac{V_{t,t+1}}{\xi_{t,t+1}} &= \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left(\frac{Z_{t+1}}{Z_t} \right) \left(\frac{P_t}{P_{t+1}} \right) \\ Q_{t,t+1} &= \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left(\frac{Z_{t+1}}{Z_t} \right) \left(\frac{P_t}{P_{t+1}} \right). \end{aligned} \quad (6.24)$$

Tomando a esperança condicional em ambos os lados de (6.24), tem-se a equação de Euler

$$Q_t = \beta E_t \left\{ \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left(\frac{Z_{t+1}}{Z_t} \right) \left(\frac{P_t}{P_{t+1}} \right) \right\} \quad (6.25)$$

onde $Q_t \equiv E_t \{ Q_{t,t+1} \}$.

Reescrevendo as equações (6.18) e (6.25) de forma log-linearizada

$$w_t - p_t = \sigma c_t + \varphi n_t \quad (6.26)$$

$$c_t = E_t \{ c_{t+1} \} - \frac{1}{\sigma} (i_t - E_t \{ \pi_{t+1} \} - \rho) + \frac{1}{\sigma} (1 - \rho_z) z_t \quad (6.27)$$

onde as letras minúsculas denotam o logaritmo das respectivas variáveis, isto é, $x_t \equiv \log X_t$. Define-se assim $i_t \equiv -\log Q_t$ como a taxa nominal de curto prazo, $\rho \equiv -\log \beta$ a taxa de desconto intertemporal e $\pi_t \equiv p_t - p_{t-1}$ a inflação IPC, com $p_t \equiv \log P_t$.

6.1.1.1 Inflação, Taxa de Câmbio Real e Termos de Troca

Os termos de troca bilaterais entre a economia doméstica e um país i é definido como $S_{i,t} = \frac{P_{i,t}}{P_{H,t}}$, ou seja, o preço de um bem do país i em termos de bens domésticos. Portanto, os termos de troca efetivos são dados por

$$\begin{aligned} S_t &\equiv \frac{P_{F,t}}{P_{H,t}} \\ &= \left(\int_0^1 S_{i,t}^{1-\gamma} di \right)^{\frac{1}{1-\gamma}} \end{aligned} \quad (6.28)$$

o qual pode ser aproximado em torno de um estado estacionário simétrico satisfazendo $S_{i,t} = 1$ para todo $i \in [0, 1]$ por

$$s_t = \int_0^1 s_{i,t} di \quad (6.29)$$

onde $s_t \equiv \log S_t = p_{F,t} - p_{H,t}$.

De forma similar, a log-linearização do IPC em torno do mesmo estado estacionário simétrico resulta em

$$\begin{aligned} p_t &\equiv (1 - \alpha)p_{H,t} + \alpha p_{F,t} \\ &= p_{H,t} + \alpha s_t. \end{aligned} \quad (6.30)$$

Note que a inflação doméstica, definida como a taxa de mudanças no índice de preços domésticos, $\pi_{H,t} \equiv p_{H,t+1} - p_{H,t}$ e a inflação IPC estão ligadas de acordo com a relação

$$\pi_t = \pi_{H,t} + \alpha \Delta s_t \quad (6.31)$$

o que faz o hiato entre as duas medidas de inflação proporcional à mudança percentual nos termos de troca, com o coeficiente de proporcionalidade dado pelo grau de abertura α .

Assumindo que a lei do preço único se mantém para bens individuais por todos os períodos (tanto para preços de importação quanto de

exportação), implicando que $P_{i,t}(j) = \varepsilon_{i,t} P_{i,t}^i(j)$ para todo $i, j \in [0, 1]$, onde $\varepsilon_{i,t}$ é a taxa de câmbio bilateral nominal e $P_{i,t}^i(j)$ é o preço do bem j do país i expresso em termos da sua própria moeda. Introduzindo a definição anterior na definição de $P_{i,t}$ tem-se $P_{i,t} = \varepsilon_{i,t} P_{i,t}^i$, onde $P_{i,t}^i \equiv (\int_0^1 P_{i,t}^i(j)^{1-\epsilon} dj)^{\frac{1}{1-\epsilon}}$ é o índice de preços doméstico do país i . Por sua vez, substituindo na definição de $P_{F,t}$ e log-linearizando em torno do estado estacionário simétrico,

$$\begin{aligned} p_{F,t} &= \int_1^0 (e_{i,t} + p_{i,t}^i) di \\ &= e_t + p_t^* \end{aligned} \quad (6.32)$$

onde $p_{i,t}^i \equiv \int_0^1 p_{i,t}^i(j) dj$ é o (logaritmo) do índice de preços doméstico para o país i , expresso em termos da sua própria moeda, $e_t \equiv \int_0^1 e_{i,t} di$ é o (logaritmo) da taxa de câmbio nominal efetiva, e $p_t^* \equiv \int_0^1 p_{i,t}^i di$ é o (logaritmo) do índice de preços mundial. Note que para o mundo como um todo não há distinção entre IPC e nível de preços doméstico, nem entre suas taxas de inflação correspondentes.

Combinando o resultado em (6.32) com a definição dos termos de troca, obtém-se

$$s_t = e_t + p_t^* - p_{H,t}. \quad (6.33)$$

Define-se a taxa de câmbio real bilateral com o país i como $Q_{i,t} \equiv \frac{\varepsilon_{i,t} P_t^i}{P_t}$, ou seja, a taxa do IPC de dois países, ambos expressos em termos da moeda doméstica corrente. Seja $q_t \equiv \int_0^1 q_{i,t} di$ o (logaritmo) o da taxa real de câmbio efetiva, onde $q_{i,t} \equiv \log Q_{i,t}$. Segue que

$$\begin{aligned} q_t &= \int_0^1 (e_{i,t} + p_t^i - p_t) di \\ &= e_t + p_t^* - p_t \\ &= s_t + p_{H,t} - p_t \\ &= (1 - \alpha) s_t \end{aligned} \quad (6.34)$$

onde a última igualdade só se mantém para uma aproximação de primeira ordem quando $\eta \neq 1$.

6.1.1.2 Diversificação Internacional de Risco

Sob a hipótese de mercados completos para seguros tratados internacionalmente, uma condição análoga a (6.25) deve também ser mantida para a família representativa em outro país, um país i

$$1 = \beta E_t \left\{ Q_t^{-1} \left(\frac{C_{t+1}^i}{C_t^i} \right)^{-\sigma} \left(\frac{Z_{t+1}^i}{Z_{t+1}^i} \right) \left(\frac{\varepsilon_{i,t} P_{i,t}^i}{\varepsilon_{i,t+1} P_{i,t+1}^i} \right) \right\} \quad (6.35)$$

onde a presença do termo de taxa de câmbio reflete o fato de que o seguro adquirido pela família do país i tem um preço $V_{t,t+1}$ e uma unidade de retorno expressos na moeda corrente da pequena economia aberta de referência, e portanto, precisam ser convertidos para a moeda corrente do país i .

Dividindo (6.25) por (6.35) e resolvendo para C_t , obtém-se

$$C_t = \vartheta_i C_t^i Q_{i,t}^{\frac{1}{\sigma}} \quad (6.36)$$

para todo t , e onde ϑ_i é uma constante que geralmente irá depender das condições iniciais. Sem perda de generalidade, assume-se que todos os países começam com condições idênticas, então $\vartheta_i = \vartheta = 1$ para todo i .

Tomando o logaritmo em ambos os lados de (6.36) e integrando sobre todos os possíveis valores de i

$$\begin{aligned} c_t &= c_t^* + \frac{1}{\sigma} q_t \\ &= c_t^* + \left(\frac{1 - \alpha}{\sigma} \right) s_t \end{aligned} \quad (6.37)$$

onde $c_t^* \equiv \int_0^1 c_t^i di$ é o (logaritmo) índice de consumo do mundo, e a segunda igualdade só se mantém para uma aproximação de primeira ordem quando $\eta \neq 1$. Portanto, a hipótese de mercados completos a nível internacional leva a uma simples relação, ligando consumo doméstico com o consumo do resto do mundo e o termos de troca.

6.1.1.3 Paridade Descoberta dos Juros

Permitir que a família representativa invista tanto em títulos domésticos quanto estrangeiros, B_t e B_t^* , permite que a restrição orça-

mentária seja escrita como

$$P_t C_t + Q_{t,t+1} B_{t+1} + Q_{t,t+1} \varepsilon_{t+1} B_{t+1}^* \leq B_t + \varepsilon_t B_t^* + W_t N_t + T_t. \quad (6.38)$$

As condições de otimalidade com relação a esses ativos são

$$1 = \beta E_t \left\{ Q_t^{-1} \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left(\frac{Z_{t+1}}{Z_t} \right) \left(\frac{P_t}{P_{t+1}} \right) \right\} \quad (6.39)$$

$$1 = \beta E_t \left\{ Q_t^{*-1} \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left(\frac{Z_{t+1}}{Z_t} \right) \left(\frac{P_t}{P_{t+1}} \frac{\varepsilon_{t+1}}{\varepsilon_t} \right) \right\} \quad (6.40)$$

Dividindo (6.39) por (6.40)

$$\begin{aligned} 1 &= E_t \left\{ \frac{Q_t^* \varepsilon_t}{Q_t \varepsilon_{t+1}} \right\} \\ \frac{Q_t^*}{Q_t} &= E_t \left\{ \frac{\varepsilon_{t+1}}{\varepsilon_t} \right\}. \end{aligned} \quad (6.41)$$

Log-linearizando (6.41)

$$\begin{aligned} -i_t^* + i_t &= E_t \{ e_{t+1} - e_t \} \\ i_t &= i_t^* + E_t \{ \Delta e_{t+1} \}. \end{aligned} \quad (6.42)$$

A equação (6.42) é a equação da paridade da taxa de juros descoberta, que afirma que a taxa de juros nominal doméstica é igual a taxa de juros nominal do resto do mundo mais a taxa de depreciação esperada da moeda corrente. De (6.34), tem-se

$$\begin{aligned} E_t \{ s_{t+1} \} - s_t &= E_t \{ e_{t+1} \} - e_t + E_t \{ p_{t+1}^* \} - p_t^* - E_t \{ p_{H,t+1} \} + p_{H,t} \\ s_t &= -E_t \{ \Delta e_{t+1} \} - E_t \{ \pi_{t+1}^* \} + E_t \{ \pi_{H,t+1} \} + E_t \{ s_{t+1} \}. \end{aligned} \quad (6.43)$$

Usando (6.42), tem-se a seguinte equação em diferenças

$$s_t = (i_t^* - E_t \{ \pi_{t+1}^* \}) - (i_t - E_t \{ \pi_{H,t+1} \}) + E_t \{ s_{t+1} \}. \quad (6.44)$$

Resolvendo através de substituição recursiva, tem-se

$$s_t = E_t \left\{ \sum_{k=0}^{\infty} [(i_{t+k}^* - \pi_{t+k}^*) - (i_{t+k} - \pi_{H,t+k})] \right\} \quad (6.45)$$

ou seja, (6.45) expressa os termos de troca como a esperança da soma dos diferenciais das taxas de juros reais entre o mercado do resto do mundo e o mercado doméstico.

6.1.2 Firms

6.1.2.1 Tecnologia

A firma na economia doméstica produz um bem diferenciado com tecnologia linear representada por uma função de produção para um dado nível de produção $Y_t(j)$,

$$Y_t(j) = A_t N_t(j) \quad (6.46)$$

onde a produtividade total dos fatores A_t é um processo de raiz unitária exógeno, dado por

$$\ln A_t = \gamma^* + \ln A_t + a_t \quad (6.47)$$

onde a_t segue um processo AR(1) estacionário,

$$a_t = \rho_a a_{t-1} + \epsilon_t^a, \quad (6.48)$$

em que ϵ_t^a é um ruído branco e pode ser interpretado como um choque tecnológico e $j \in [0, 1]$ é um índice para uma específica firma.

Portanto, o custo marginal (expresso em termos dos preços domésticos) será comum entre todas as firmas domésticas e é dado por

$$mc_t = -\nu + w_t - p_{H,t} - a_t \quad (6.49)$$

onde $\nu \equiv -\log(1 - \tau)$, com τ sendo um subsídio ao emprego.

6.1.2.2 Definição de Preços

A rigidez de preços é introduzida usando a rigidez de Calvo (1983). Assim, em qualquer período determinado, apenas uma fração das empresas $(1 - \theta)$, escolhida aleatoriamente, pode reotimizar seus preços, sendo que a probabilidade de otimização de cada empresa é dada de forma independente do tempo decorrido desde a última atualização de preços.

A firma que ajusta seu preço t o faz de acordo com

$$p_{\bar{H},t} = \mu + (1 - \beta\theta) \sum_{k=0}^{\infty} (\beta\theta)^k E_t \{ mc_{t+k} + p_{H,t+k} \} \quad (6.50)$$

onde $p_{H,t+k}$ denota o logaritmo do preço do composto de bens domésticos e $\mu = \log \frac{\epsilon}{\epsilon-1}$ é o *mark up* bruto de equilíbrio.

6.2 EQUILÍBRIO

6.2.1 Demanda Agregada e Determinação do Produto

6.2.1.1 Consumo e Produto

O equilíbrio do mercado de bens da economia doméstica requer

$$\begin{aligned} Y_t(j) &= C_{H,t}(j) + \int_0^1 C_{H,t}^i(j) di = \left\{ \frac{P_{H,t(j)}}{P_{H,t}} \right\}^{-\epsilon} \\ &\times \left[(1 - \alpha) \left(\frac{P_{H,t}}{P_t} \right)^{-\eta} C_t + \alpha \int_0^1 \left(\frac{P_{H,t}}{\varepsilon_{i,t} P_{F,t}^i} \right)^{-\gamma} \left(\frac{P_{F,t}^i}{P_t^i} \right)^{-\eta} C_t^i di \right] \end{aligned} \quad (6.51)$$

para todo $j \in [0, 1]$ e todo t , onde $C_{H,t}^i(j)$ denota a demanda no país i pelo bem j produzido na economia doméstica. Note que a segunda igualdade fez uso de (6.11), (6.13) e (6.14) junto com a hipótese de preferências simétricas entre os países, que implica que $C_{H,t}^i(j) = \alpha \left(\frac{P_{H,t(j)}}{P_{H,t}} \right)^{-\epsilon} \left(\frac{P_{H,t}}{\varepsilon_{i,t} P_{F,t}^i} \right)^{-\gamma} \left(\frac{P_{F,t}^i}{P_t^i} \right)^{-\eta} C_t^i$.

Plugando (6.51) na definição de produto agregado doméstico

$Y_t \equiv [\int_0^1 Y_t(j)^{1-\frac{1}{\epsilon}} dj]^{\frac{\epsilon}{\epsilon-1}}$, tem-se

$$\begin{aligned}
 Y_t &= (1 - \alpha) \left(\frac{P_{H,t}}{P_t} \right)^{-\eta} C_t + \alpha \int_0^1 \left(\frac{P_{H,t}}{\varepsilon_{i,t} P_{F,t}^i} \right)^{-\gamma} \left(\frac{P_{F,t}^i}{P_t} \right)^{-\eta} C_t^i di \\
 &= \left(\frac{P_{H,t}}{P_t} \right)^{-\eta} \left\{ (1 - \alpha) C_t + \alpha \int_0^1 \left(\frac{\varepsilon_{i,t} P_{F,t}^i}{P_{H,t}} \right)^{\gamma-\eta} Q_{i,t}^\eta C_t^i di \right\} \\
 &= \left(\frac{P_{H,t}}{P_t} \right)^{-\eta} C_t \left[(1 - \alpha) + \alpha \int_0^1 (S_t^i S_{i,t})^{\gamma-\eta} Q_{i,t}^{\eta-\frac{1}{\sigma}} di \right] \quad (6.52)
 \end{aligned}$$

onde a última igualdade segue de (6.36), e onde S_t^i denota os termos de troca efetivos para o país i , enquanto $S_{i,t}$ denota os termos de troca bilaterais entre a economia doméstica e o país i .

De forma geral, assumindo que $\int_0^1 s_t^i di = 0$, tem-se a seguinte aproximação log-linear de primeira ordem para (6.52) em torno do estado estacionário simétrico

$$\begin{aligned}
 y_t &= c_t + \alpha \gamma s_t + \alpha \left(\eta - \frac{1}{\sigma} \right) q_t \\
 &= c_t + \frac{\alpha \omega}{\sigma} s_t \quad (6.53)
 \end{aligned}$$

onde $\omega \equiv \sigma \gamma + (1 - \alpha)(\sigma \eta - 1)$. Note que uma condição análoga deve manter-se para todos os países, portanto, para um país genérico i , (6.53) pode ser reescrito como $y_t^i = c_t^i + \frac{\alpha \omega}{\sigma} s_t^i$. Agregando todos os países, uma condição de equilíbrio do mercado do resto do mundo pode ser derivada como

$$\begin{aligned}
 y_t^* &\equiv \int_0^1 y_t^i di \\
 &= \int_0^1 c_t^i di \equiv c_t^* \quad (6.54)
 \end{aligned}$$

onde y_t^* e c_t^* são índices para produto e consumo no resto do mundo, e onde a principal igualdade segue, mais uma vez, do fato de que $\int_0^1 s_t^i di = 0$. Além disso, assume-se que y_t^* segue um processo AR(1) estacionário

$$y_t^* = \rho_y y_{t-1}^* + \epsilon_t^y \quad (6.55)$$

onde ϵ_t^y é um ruído branco e pode ser interpretado como um choque no produto do resto do mundo.

Combinando (6.53) com (6.37) e (6.54), tem-se

$$y_t = y_t^* + \frac{1}{\sigma_\alpha} s_t \quad (6.56)$$

onde $\sigma_\alpha \equiv \frac{\sigma}{1+\alpha(\omega-1)} > 0$.

Por fim, combinando (6.53) com a equação de Euler em (6.27)

$$y_t = E_t\{y_{t+1}\} - \frac{1}{\sigma_\alpha} (i_t - E_t\{\pi_{H,t+1}\} - \rho) + \frac{1}{\sigma_\alpha} (1 - \rho_z) z_t + \alpha \Theta E_t\{\Delta y_{t+1}^*\} \quad (6.57)$$

onde $\Theta = \omega - 1$.

6.2.1.2 A Balança Comercial

Seja $nx_t \equiv \left(\frac{1}{Y}\right) \left(Y_t - \frac{P_t}{P_{H,t}} C_t\right)$ as exportações líquidas, expressas em termos do produto doméstico, como uma função do estado estacionário Y . Uma aproximação de primeira ordem produz $nx_t = y_t - c_t - \alpha s_t$, que combinado com (6.53), implica em uma relação simples entre exportações líquidas e termos de troca

$$nx_t = \alpha \left(\frac{\omega}{\sigma} - 1\right) s_t. \quad (6.58)$$

O sinal da relação entre os termos de troca e as exportações líquidas será ambíguo e dependerá do tamanho relativo de σ , γ e η .

6.2.2 O Lado da Oferta

6.2.2.1 Produto Agregado e Emprego

Equilíbrio no mercado de trabalho exige que

$$N_t = \int_0^1 N_t(j) dj. \quad (6.59)$$

Usando a função de produção (6.46), tem-se

$$N_t = \frac{Y_t}{A_t} \int_0^1 \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} dj. \quad (6.60)$$

Uma aproximação de primeira ordem produz a seguinte relação

$$n_t = y_t - a_t + d_t \quad (6.61)$$

onde $d_t \equiv \int_0^1 \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} dj$ é uma medida de dispersão de preços, e será igual a zero em uma aproximação log-linear de primeira ordem. Logo, tem-se

$$y_t = a_t + n_t. \quad (6.62)$$

6.2.2.2 Custo Marginal e Dinâmica Inflacionária

Na pequena economia aberta, as dinâmicas da inflação doméstica em termos do custo marginal real são descritas pela seguinte equação

$$\pi_{H,t} = \beta E_t \{ \pi_{H,t+1} \} + \lambda \hat{m}c_t \quad (6.63)$$

onde $\lambda \equiv \frac{(1-\beta\theta)(1-\theta)}{\theta}$. Porém, a relação em (6.63) não depende de nenhum parâmetro que caracterize a economia aberta. Por outro lado, a determinação do custo marginal real como uma função do produto doméstico da economia aberta difere um pouco de uma economia fechada. Assim, tem-se

$$\begin{aligned} mc_t &= -\nu + (w_t - p_{H,t}) - a_t \\ &= -\nu + (w_t - p_t) + (p_t - p_{H,t}) - a_t \\ &= -\nu + \sigma c_t + \varphi n_t + \alpha s_t - a_t \\ &= -\nu + \sigma y_t^* + \varphi y_t + s_t - (1 + \varphi)a_t \end{aligned} \quad (6.64)$$

onde na última igualdade utilizou-se (6.62) e (6.37).

Finalmente, utilizando (6.56) para substituir os termos de troca, (6.64) pode ser reescrita como

$$mc_t = -\nu + (\sigma_\alpha + \varphi)y_t + (\sigma - \sigma_\alpha)y_t^* - (1 + \varphi)a_t. \quad (6.65)$$

Sob a hipótese de preços flexíveis é fácil perceber que $mc_t = -\mu$ para todo t . Portanto, o nível natural de produto na economia aberta é dado por

$$y_t^n = \Gamma_0 + \Gamma_a a_t + \Gamma^* y_t^* \quad (6.66)$$

onde $\Gamma_0 \equiv \frac{\nu - \mu}{\sigma_\alpha + \varphi}$, $\Gamma_a \equiv \frac{1 + \varphi}{\sigma_\alpha + \varphi} > 0$, e $\Gamma \equiv -\frac{\alpha \Theta \sigma_\alpha}{\sigma_\alpha + \varphi}$.

6.2.3 Equilíbrio Dinâmico

Seja $\tilde{y}_t \equiv y_t - y_t^n$ o hiato do produto doméstico. Dado (6.65) e o fato que y_t^* é invariante aos desenvolvimentos domésticos, segue que o custo marginal real e o hiato do produto relacionam-se de acordo com

$$\hat{m}c_t = (\sigma_\alpha + \varphi)\tilde{y}_t. \quad (6.67)$$

Combinando a expressão anterior com (6.63), obtém-se a seguinte versão da Curva de Phillips Novo Keynesiana para a economia aberta

$$\pi_{H,t} = \beta E_t\{\pi_{H,t+1}\} + \kappa_\alpha \tilde{y}_t \quad (6.68)$$

onde $\kappa_\alpha \equiv \lambda(\sigma_\alpha + \varphi)$. Usando (6.57) é possível derivar uma versão da equação IS dinâmica para a pequena economia aberta em termos do hiato do produto

$$\tilde{y}_t = E_t\{\tilde{y}_{t+1}\} - \frac{1}{\sigma_\alpha}(i_t - E_t\{\pi_{H,t+1}\} - r_t^n) + \frac{1}{\sigma_\alpha}(1 - \rho_z)z_t \quad (6.69)$$

onde r_t^n é a taxa natural de juros da pequena economia aberta.

6.3 PEQUENA ECONOMIA ABERTA SOB UMA REGRA DE TAXA DE JUROS

Para fechar o modelo assume-se que a autoridade monetária segue uma regra do tipo Taylor, que penaliza os desvios da inflação doméstica da meta, assim como do hiato do produto.

$$i_t = \phi_\pi \pi_{H,t} + \phi_y \tilde{y}_t + \epsilon_t^i \quad (6.70)$$

onde ϵ_t^i é um ruído branco e pode ser interpretado como um choque monetário. Os parâmetros ϕ_π e ϕ_y são coeficientes não negativos escolhidos pela autoridade monetária.

O sistema log-linearizado pode ser reduzido para quatro equações em produto, inflação, taxa nominal de juros e taxa de câmbio real

efetiva

$$y_t = E_t\{y_{t+1}\} - \frac{1}{\sigma_\alpha}(i_t - E_t\{\pi_{H,t}\} - \rho) + \frac{1}{\sigma_\alpha}(1 - \rho_z)z_t + \alpha(\omega - 1)(\rho_y - 1)y_t^* \quad (6.71)$$

$$\pi_{H,t} = \beta E_t\{\pi_{H,t+1}\} + \kappa_\alpha y_t - \kappa_\alpha \alpha \Psi y_t^* - \kappa_\alpha \Gamma_\alpha \tilde{a}_t \quad (6.72)$$

$$i_t = \phi_i i_{t-1} + (1 - \phi_i)(\phi_\pi \pi_{H,t} + \phi_y y_t - \phi_y \alpha \Psi y_t^* - \phi_y \Gamma_\alpha \tilde{a}_t) + \epsilon_t^i \quad (6.73)$$

$$q_t = (1 - \alpha)\sigma_\alpha y_t - (1 - \alpha)\sigma_\alpha y_t^* \quad (6.74)$$

O sistema linear de expectativas racionais dado pelas equações (6.2), (6.48), (6.55) e (6.71)-(6.74), pode ser resolvido usando métodos adequados, como por exemplo, os descritos em Blanchard e Kahn (1980), Klein (2000) e Sims (2002).

A relação entre os desvios do estado estacionário e as observações de produto, inflação, taxa de juros nominal e taxa de câmbio real efetiva é dada pelas seguintes equações de medida

$$\Delta \ln y_t^{obs} = \gamma^* + \Delta y_t + \tilde{a}_t \quad (6.75)$$

$$\pi_{H,t}^{obs} = \pi^* + \pi_{H,t}$$

$$i_t^{obs} = 12[(\rho + \pi^*) + i_t]$$

$$\Delta \ln q_t^{obs} = q^* + \Delta q_t$$

onde π^* e q^* correspondem ao logaritmo do estado estacionário da inflação e da taxa de câmbio real efetiva, respectivamente.

Na análise do exercício empírico subsequente um período t corresponde a um mês. As variáveis observadas produto, inflação e taxa de câmbio têm periodicidade mensal, enquanto a taxa de juros é anualizada. O modelo DSGE possui quatro choques estruturais, contidos no vetor $\epsilon_t = [\epsilon_t^a, \epsilon_t^y, \epsilon_t^i, \epsilon_t^z]'$ e, assume-se que os choques são normalmente distribuídos e independentes entre si. Seus desvios padrão são denotados por σ_a , σ_y , σ_i e σ_z , respectivamente. Os parâmetros do modelo DSGE a serem estimados estão contidos no vetor $\theta \in \Theta$, assim

$$\theta_{DSGE} = [\sigma, \varphi, \theta, \phi_\pi, \phi_y, \rho_a, \rho_y, \rho_z, \gamma^*, \pi^*, q^*, \rho, \sigma_a, \sigma_y, \sigma_i, \sigma_z]'. \quad (6.76)$$

Modelos DSGE log-linearizados com expectativas racionais podem ser representados na forma de modelos de espaço de estado, onde as variáveis observadas são ligadas às variáveis do modelos através de

uma equação de medida. Ao mesmo tempo, a equação de transição fornece a forma reduzida do modelo DSGE, mapeando as variáveis correntes com suas defasagens e choques independentes e identicamente distribuídos. Se a solução do modelo DSGE, linear e Gaussiano, é única, o filtro de Kalman pode ser aplicado para computar o logaritmo da função de verossimilhança.

A solução dos termos esperados na forma estrutural do modelo DSGE leva a representação espaço de estado, com a equação de transição que pode ser descrita como

$$s_t = T(\theta)s_{t-1} + R(\theta)\epsilon_t. \quad (6.77)$$

A equação de medida na forma empilhada pode ser escrita como

$$y_t = Z(\theta)s_t + D(\theta) + \nu_t. \quad (6.78)$$

As matrizes de variância e covariância dos choques são definidas como

$$E[\nu_t\nu_t'] = \Sigma\nu\nu(\theta) \quad E[\epsilon\epsilon_t'] = \Sigma\epsilon\epsilon(\theta) \quad E[\epsilon\nu_t'] = \Sigma\epsilon\nu(\theta). \quad (6.79)$$

O fato de ter-se um número de choques igual ou superior ao número de variáveis observáveis possibilita definir $\nu_t = 0$.

7 METODOLOGIA DSGE-VAR

Existe uma relação umbilical entre os modelos DSGE e VAR: a representação espaço de estados de um modelo DSGE log-linearizado pode ser escrita como um vetor autorregressivo de média móvel (VARMA) em sua forma reduzida, onde os parâmetros são restringidos pelas funções de restrição $T(\theta)$ e $R(\theta)$ na equação de transição em (6.77). Para o caso em que o número de variáveis observadas é igual ao número de choques no modelo, Fernandez-Villaverde, Rubio-Ramirez e Sargent (2005) mostram que o modelo VARMA pode ser invertido e escrito como um VAR(∞). Por sua vez, Giacomini (2013) mostra que o VAR(∞) pode ser aproximado por um VAR(p). Assim, a estimação de um modelo DSGE linearizado pode ser aproximada da estimação de um VAR(p) restrito, sendo que o erro de aproximação é reduzido a medida que aumenta-se a ordem de defasagem p . É importante notar que o VAR(p) restrito herda as propriedades do DSGE, isto é, ambos são versões muito estilizadas do mundo real, por construção. Portanto, VAR irrestritos são mais flexíveis, pois não impõe restrições aos parâmetros. Por outro lado, VAR irrestritos em sua forma reduzida não são úteis para análises de políticas e, além disso, sofrem da maldição da dimensionalidade.

Del Negro e Schorfheide (2004) oferecem uma solução a este problema ao utilizar um modelo DSGE teórico como distribuição a priori na estimação Bayesiana de um VAR, dando origem aos modelos denominados DSGE-VAR.

A abordagem Bayesiana tem como uma das suas principais vantagens o fato de sua estrutura ser baseada em algumas regras simples de probabilidade. Para motivar a simplicidade dessa abordagem considere duas variáveis aleatórias, denotadas por A e B . As regras de probabilidade implicam que:

$$p(A, B) = p(A|B)p(B)$$

onde $p(A, B)$ é a probabilidade conjunta de A e B ocorrerem, $p(A|B)$ é a probabilidade de A ocorrer dado que B ocorreu, e $p(B)$ é a probabilidade marginal de B . Alternativamente, pode-se inverter os papéis de A e B e encontrar uma expressão para a probabilidade conjunta de A e B sendo:

$$p(A, B) = p(B|A)p(A).$$

Igualando as equações para $p(A, B)$ e rearranjando os termos, obtém-se a conhecida Regra de Bayes, que é considerada parte funda-

mental da econometria Bayesiana:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}. \quad (7.1)$$

A econometria preocupa-se em usar dados para aprender sobre algo em que o pesquisador está interessado. Tipicamente, em economia, trabalha-se com modelos que dependem de parâmetros, onde o pesquisador está interessado em estimar tais parâmetros. Assim, considerando y como um vetor ou matriz de dados e θ um vetor ou matriz de parâmetros que contenha os parâmetros para um modelo que procura explicar y , é possível utilizar a Regra de Bayes para aprender sobre θ com base nos dados, y . Portanto, substituindo B por θ e A por y em (7.1), obtém-se:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}. \quad (7.2)$$

A densidade a posteriori, denotada por $p(\theta|y)$, é a função densidade de probabilidade dos dados dado os parâmetros do modelo e é o objetivo central da abordagem Bayesiana, pois é a partir dela que serão realizadas inferências.

Uma diferença importante entre as abordagens da econometria clássica e Bayesiana é o fato que a última trata θ como uma variável aleatória, isso porque a econometria Bayesiana é baseada em uma visão subjetiva de probabilidade, argumentando que a incerteza sobre algo desconhecido pode ser expressa usando regras de probabilidade.

A partir da equação 7.2, ao passo que θ é o que se está interessado em aprender, pode-se suprimir o termo $p(y)$, uma vez que não envolve θ . Assim, 7.2 pode ser escrita como:

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad (7.3)$$

onde \propto é o símbolo de proporcionalidade. Por fim, $p(y|\theta)$ representa a função de verossimilhança dos dados e $p(\theta)$ é a densidade a priori.

A função de verossimilhança, $p(y|\theta)$, é a densidade dos dados condicional aos parâmetros vista como função dos parâmetros para uma base de dados específica. Já a priori, $p(\theta)$, não depende dos dados, isto é, ela resume o conhecimento do pesquisador a respeito dos parâmetros antes de observar a amostra. Finalmente, a posteriori, $p(\theta|y)$, é a densidade de interesse fundamental, que pode ser pensada como uma atualização da informação a priori após a inclusão das informações contidas na verossimilhança.

Utilizar o modelo DSGE como a priori implica utilizar os momentos populacionais calculados a partir do modelo. Define-se então o hiperparâmetro λ como a medida do peso relativo da a priori à amostra. No caso extremo em que $\lambda = 0$ o modelo DSGE-VAR colapsa em um BVAR com a priori difusa, sem atribuir qualquer importância às informações fornecidas pelos DSGE. Por outro lado, quando λ tende ao infinito, as restrições provenientes do DSGE são impostas rigidamente. A abordagem de Del Negro e Schorfheide (2004) fornece um meio de escolher λ .

Outra vantagem da abordagem DSGE-VAR é que os parâmetros estruturais do DSGE são estimados juntamente com os parâmetros do VAR. Geralmente o pesquisador não sabe ao certo os valores dos parâmetros estruturais do DSGE, tem apenas algum conhecimento incerto a respeito desses valores. Portanto, pode-se adotar uma distribuição a priori para os parâmetros do DSGE. Como resultado, uma a priori com estrutura hierárquica pode ser construída da seguinte forma: uma distribuição condicional para os parâmetros do VAR dado os parâmetros estruturais do DSGE e uma distribuição marginal para os parâmetros do DSGE. Dessa forma, uma vez que os parâmetros aparecem na distribuição a priori, eles são atualizados com a informação amostral, via teorema de Bayes.

7.1 FUNÇÃO DE VEROSSIMILHANÇA

A representação VARMA gerada pelo modelo DSGE linearizado do capítulo 6, apresentada nas equações (6.77)-(6.78), pode ser aproximada por um modelo VAR(p). Para que essa aproximação exista é necessário que o número de choques de modelo seja igual ao número de variáveis observadas e, além disso, segundo Giacomini (2013), para que essa relação seja exata, faz necessário que todas as variáveis sejam observadas. Entretanto, o VAR(p) oriundo do modelo DSGE, seja ele aproximado ou exato, impõe várias restrições em um modelo VAR(p) irrestrito, que pode ser escrito como:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + u_t \quad (7.4)$$

onde u_t é um vetor de erros de previsão um passo a frente.

Para construir a função de verossimilhança desse VAR(p) irrestrito, define-se Y uma matriz $T \times n$ com linhas y'_t . Seja $k = 1 + np$, X é uma matriz de dimensão $T \times k$ com linhas $x'_t = [1, y'_{t-1}, \dots, y'_{t-p}]$, U é uma matriz com linhas u'_t , e $\phi = [\phi_0, \phi_1, \dots, \phi_p]'$. Além disso,

assume-se, de acordo com Koop (2003), as seguintes hipóteses:

1. As inovações u_t em (7.4) seguem uma distribuição Normal multivariada $N(0, \Sigma_u)$ condicional às observações do passado de y_t ;
2. Os elementos de X são fixos, ou, caso sejam variáveis aleatórias, são independentes de U com função densidade de probabilidade $p(X|\psi)$, onde ψ é um vetor de parâmetros independente de ϕ e Σ_u .

A função de verossimilhança será determinada pela função densidade de probabilidade conjunta de Y e X condicional aos parâmetros desconhecidos, isto é, $p(Y, X|\phi, \Sigma_u, \psi)$. A segunda hipótese supramencionada implica que pode-se reescrever a função de verossimilhança como:

$$p(Y, X|\phi, \Sigma_u, \psi) = p(Y|X, \phi, \Sigma_u)p(X|\psi). \quad (7.5)$$

Uma vez que $p(X|\psi)$ não depende ϕ e de Σ_u , é considerado apenas uma constante, podendo ser omitido. Para fins de notação, X também será omitido de $p(Y|X, \phi, \Sigma_u)$. Dessa forma, o modelo VAR(p) irrestrito expresso como $Y = X\phi + U$ tem a seguinte função de verossimilhança

$$L(\phi, \Sigma_u|Y) = \prod_{t=1}^T \frac{1}{\sqrt{(2\pi)^n |\Sigma_u|}} \exp\left\{-\frac{1}{2}(Y - X\phi)\Sigma_u^{-1}(Y - X\phi)'\right\} \quad (7.6)$$

A equação (7.6) pode ser escrita como:

$$p(Y|\phi, \Sigma_u) = \frac{1}{(2\pi)^{\frac{nT}{2}} |\Sigma_u|^{\frac{T}{2}}} \exp\left\{-\frac{1}{2}(Y - X\phi)\Sigma_u^{-1}(Y - X\phi)'\right\}. \quad (7.7)$$

O termo $(Y - X\phi)\Sigma_u^{-1}(Y - X\phi)'$ em (7.7) pode ser escrito como

$$\begin{aligned} (Y - X\phi)\Sigma_u^{-1}(Y - X\phi)' &= Tr [(Y - X\phi)\Sigma_u^{-1}(Y - X\phi)'] \\ &= Tr [\Sigma_u^{-1}(Y - X\phi)'(Y - X\phi)]. \end{aligned} \quad (7.8)$$

Ignorando por um momento o restante da equação (7.8) e trabalhando apenas com $(Y - X\phi)'(Y - X\phi)$, aplica-se a distributiva, obtendo-se

$$(Y - X\phi)'(Y - X\phi) = Y'Y - Y'X\phi - (X\phi)'Y + (X\phi)'(X\phi). \quad (7.9)$$

Uma vez que o termo $Y'X\phi$ é um produto interno, denotado por

$\langle Y, X\phi \rangle$, vale a propriedade de comutatividade, então:

$$\begin{aligned} -Y'X\phi - (X\phi)'Y &= -\langle Y, X\phi \rangle - \langle X\phi, Y \rangle \\ &= -\langle X\phi, Y \rangle - \langle X\phi, Y \rangle \\ &= -X'\phi'Y - X'\phi'Y \\ &= -2X'\phi'Y. \end{aligned}$$

Dessa forma, tem-se:

$$(Y - X\phi)'(Y - X\phi) = Y'Y - 2\phi'X'Y + \phi'X'X\phi. \quad (7.10)$$

Somando e subtraindo $\hat{\phi}'X'X\hat{\phi}$, em que $\hat{\phi} = (X'X)^{-1}X'Y$ é o estimador de máxima verossimilhança (MV), tem-se:

$$\begin{aligned} (Y - X\phi)'(Y - X\phi) &= Y'Y - 2\phi'X'Y + \phi'X'X\phi \\ &\quad + \hat{\phi}'X'X\hat{\phi} - \hat{\phi}'X'X\hat{\phi}, \end{aligned} \quad (7.11)$$

onde o segundo termo no lado direito da igualdade em (7.11) pode ser reescrito como:

$$\begin{aligned} 2\phi'X'Y &= 2\phi' \underbrace{X'X(X'X)^{-1}}_I X'Y \\ &= 2\phi'X'X \underbrace{(X'X)^{-1}X'Y}_{\hat{\phi}} \\ &= 2\phi'X'X\hat{\phi}. \end{aligned}$$

Reescrevendo (7.11), tem-se:

$$\begin{aligned} (Y - X\phi)'(Y - X\phi) &= Y'Y - \hat{\phi}'X'X\hat{\phi} + \phi'X'X\phi - 2\phi'X'X\hat{\phi} + \hat{\phi}'X'X\hat{\phi} \\ &= Y'Y - 2\hat{\phi}'X'X\hat{\phi} + \hat{\phi}'X'X\hat{\phi} + (\phi - \hat{\phi})'X'X(\phi - \hat{\phi}) \\ &= (Y - X\hat{\phi})'(Y - X\hat{\phi}) + (\phi - \hat{\phi})'X'X(\phi - \hat{\phi}). \end{aligned}$$

Definindo $S = (Y - X\hat{\phi})'(Y - X\hat{\phi})$, então a verossimilhança em

(7.7) pode ser escrita como

$$\begin{aligned}
 p(Y|\phi, \Sigma_u) &\propto |\Sigma_u|^{-T/2} \exp\left\{-\frac{1}{2}Tr[S\Sigma_u^{-1}]\right\} \\
 &\times \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(\phi - \hat{\phi})'X'X(\phi - \hat{\phi})]\right\}. \quad (7.12)
 \end{aligned}$$

onde (7.12) é o produto do núcleo de uma densidade Normal para $\phi|\Sigma_u$ vezes o núcleo de uma densidade Wishart-invertida para Σ_u .

7.2 DISTRIBUIÇÃO A PRIORI

Distribuições a priori podem, de maneira conveniente, ser representadas por observações simuladas. Essa ideia nasceu de Theil e Goldberger (1961), que propuseram a conhecida estimativa mista. Essas observações simuladas podem ser observações reais obtidas de outros países ou observações simuladas obtidas de modelos macroeconômicos.

Suponha que as observações amostradas são aumentadas com a inclusão de $T^* = \lambda T$ observações artificiais Y^*, X^* , geradas a partir do modelo DSGE descrito no capítulo 6, baseado no vetor de parâmetros θ . A função de verossimilhança para amostra, que combina as observações amostradas e as observações artificiais, é obtida pré-multiplicando (7.12) com

$$p(Y^*(\theta)|\phi, \Sigma_u) \propto \frac{1}{(2\pi)^{\frac{n\lambda T}{2}} |\Sigma_u|^{\frac{\lambda T}{2}}} \exp\left\{-\frac{1}{2}(Y^* - X^*\phi)\Sigma_u^{-1}(Y^* - X^*\phi)'\right\}. \quad (7.13)$$

A fatoração

$$p(Y^*(\theta), Y|\phi, \Sigma_u) = p(Y^*(\theta)|\phi, \Sigma_u)p(Y|\phi, \Sigma_u) \quad (7.14)$$

sugere que o termo $p(Y^*(\theta)|\phi, \Sigma_u)$ pode ser interpretado como uma densidade a priori para ϕ e Σ_u .

Adotar informações simuladas como distribuição a priori possibilita que a a priori no DSGE-VAR possa ser interpretada como a função de verossimilhança dos dados simulados do modelo DSGE. Portanto, da mesma forma que a função de verossimilhança foi escrita como o produto do núcleo de uma densidade Normal para $\phi|\Sigma_u$ vezes o núcleo de uma densidade Wishart-Invertida para Σ_u , o mesmo ocorre com a priori. Isto é, a priori é dita conjugada natural, o que implica que além de produzir uma posteriori de mesma distribuição, a priori tem a

mesma distribuição que a verossimilhança.

Portanto, a partir deste ponto serão realizadas manipulações matemáticas de forma a obter a distribuição a priori escrita como a combinação do núcleo de uma distribuição Normal vezes o núcleo de uma densidade Wishart-invertida. Esquecendo-se por um momento do restante da expressão (7.13) e lidando apenas com o termo $(Y^* - X^*\phi)\Sigma_u^{-1}(Y^* - X^*\phi)'$, este pode ser reescrito como

$$\begin{aligned} (Y^* - X^*\phi)\Sigma_u^{-1}(Y^* - X^*\phi)' &= Tr [(Y^* - X^*\phi)\Sigma_u^{-1}(Y^* - X^*\phi)'] \\ &= Tr [\Sigma_u^{-1}(Y^* - X^*\phi)'(Y^* - X^*\phi)] \end{aligned} \quad (7.15)$$

Aplicando-se a distributiva em $(Y^* - X^*\phi)'(Y^* - X^*\phi)$, obtem-se:

$$(Y^* - X^*\phi)'(Y^* - X^*\phi) = Y^{*'}Y^* - \phi'X^{*'}Y^* - Y^{*'}X^*\phi + \phi'X^{*'}X^*\phi \quad (7.16)$$

Uma vez que o termo $Y^{*'}X^*\phi$ é um produto interno, vale a propriedade de comutatividade, então:

$$(Y^* - X^*\phi)'(Y^* - X^*\phi) = Y^{*'}Y^* - 2\phi'X^{*'}Y^* + \phi'X^{*'}X^*\phi. \quad (7.17)$$

Somando e subtraindo o termo $\phi^{*'}X^{*'}X^*\phi^*$, em que define-se $\phi^* = (X^{*'}X^*)^{-1}X^{*'}Y^*$ como o estimador de MV, então

$$\begin{aligned} (Y^* - X^*\phi)'(Y^* - X^*\phi) &= \\ &= Y^{*'}Y^* - 2\phi'X^{*'}Y^* + \phi'X^{*'}X^*\phi \\ &+ \phi^{*'}X^{*'}X^*\phi^* - \phi^{*'}X^{*'}X^*\phi^*. \end{aligned} \quad (7.18)$$

O segundo termo no lado direito de (7.18) pode ser reescrito da seguinte maneira:

$$\begin{aligned} 2\phi'X^{*'}Y^* &= 2\phi' \underbrace{X^{*'}X^*(X^{*'}X^*)^{-1}}_I X^{*'}Y^* \\ &= 2\phi'X^{*'}X^* \underbrace{(X^{*'}X^*)^{-1}X^{*'}Y^*}_{\phi^*} \\ &= 2\phi X^{*'}X^*\phi^* \end{aligned}$$

Então (7.18) pode ser escrita como:

$$\begin{aligned}
(Y^* - X^* \phi)'(Y^* - X^* \phi) &= Y^{*'} Y^* - \phi^{*'} X^{*'} X^* \phi^* \\
&\quad + \phi' X^{*'} X^* \phi - 2\phi' X^{*'} X^* \phi^* + \phi^{*'} X^{*'} X^* \phi^* \\
&= Y^{*'} Y^* - 2\phi^{*'} X^{*'} X^* \phi^* + \phi^{*'} X^{*'} X^* \phi^* \\
&\quad + (\phi - \phi^*)' X^{*'} X^* (\phi - \phi^*) \\
&= (Y^* - X^* \phi^*)'(Y^* - X^* \phi^*) \\
&\quad + (\phi - \phi^*)' X^{*'} X^* (\phi - \phi^*).
\end{aligned}$$

Finalmente, (7.13) pode ser reescrita como o núcleo de uma densidade Normal vezes o núcleo de uma densidade Wishart-invertida

$$\begin{aligned}
p(Y^*(\theta)|\phi, \Sigma_u) &\propto \frac{1}{(2\pi)^{\frac{n\lambda T}{2}} |\Sigma_u|^{\frac{\lambda T}{2}}} \exp\left\{-\frac{1}{2} Tr[\Sigma_u^{-1} \Sigma_u^*]\right\} \\
&\quad \times \exp\left\{-\frac{1}{2} Tr[\Sigma_u^{-1} (\phi - \phi^*)' X^{*'} X^* (\phi - \phi^*)]\right\} \quad (7.19)
\end{aligned}$$

onde $\Sigma_u^* = (Y^* - X^* \phi^*)'(Y^* - X^* \phi^*)$.

As manipulações matemáticas de (7.13) até (7.19) possibilitaram escrever a priori como o núcleo de uma densidade Normal vezes o núcleo de uma densidade Wishart-invertida, de maneira que esta agora pode ser combinada analiticamente com a função de verossimilhança em (7.12). A combinação da distribuição a priori conjugada natural e da função de verossimilhança produz uma posteriori que também seguirá uma distribuição Normal-Wishart-invertida, condicional ao parâmetro θ .

É importante notar que repetidas aplicações do procedimento de construir uma a priori obtendo realizações aleatórias a partir de um modelo DSGE levam a problemas de variação estocástica, uma vez que os modelos DSGE são modelos estocásticos. A fim de resolver esse problema em $p(Y^*(\theta)|\phi, \Sigma_u)$, Del Negro e Schorfheide (2004) sugerem a substituição dos momentos amostrais $Y^{*'} Y^*$, $Y^{*'} X^*$ e $X^{*'} X^*$ pelos seus valores esperados. Segundo o modelo DSGE, o vetor y_t é estacionário em covariância e seus valores esperados dos momentos amostrais são dados pelos momentos populacionais $\lambda T \Gamma_{yy}^*(\theta)$, $\lambda T \Gamma_{yx}^*(\theta)$ e $\lambda T \Gamma_{xx}^*(\theta)$, onde, por exemplo, $\Gamma_{yy}^*(\theta) = E_\theta[y_t y_t']$. Os momentos populacionais são computados analiticamente a partir da representação espaço de estado (6.77, 6.78) do modelo DSGE e evitam então a necessidade de simular

amostras do modelo teórico

$$E[y_t y_t'] = Z\Omega_{ss}Z' + ZR\Sigma_{\epsilon\nu} + (ZR\Sigma_{\epsilon\nu})' + \Sigma_{\nu\nu} + DD' \quad (7.20)$$

$$E[y_t y_{t-h}'] = ZT^h(\Omega_{ss}Z' + ZR\Sigma_{\epsilon\nu}) + DD'$$

em que $\Omega_{ss} = E[s_t s_t']$ e pode ser obtida resolvendo a seguinte equação de Lyapunov: $\Omega_{ss} = T\Omega_{ss}T' + R\Sigma_{\epsilon\epsilon}R'$.

Formalmente, usar os momentos populacionais implica substituir (7.19) por

$$\begin{aligned} p(Y^*(\theta)|\phi, \Sigma_u) &\propto |\Sigma_u|^{-\frac{\lambda T}{2}} \exp\left\{-\frac{1}{2}Tr [\lambda T\Sigma_u^* \Sigma_u^{-1}]\right\} \\ &\times \exp\left\{-\frac{1}{2}Tr [\Sigma_u^{-1}(\phi - \phi^*)' \lambda T\Gamma_{xx}^* (\phi - \phi^*)]\right\} \end{aligned} \quad (7.21)$$

onde

$$\phi^* = \Gamma_{xx}^{*-1} \Gamma_{xy}^* \quad (7.22)$$

$$\Sigma_u^* = \Gamma_{yy}^* - \Gamma_{yx}^* \Gamma_{xx}^{*-1} \Gamma_{xy}^*. \quad (7.23)$$

Adicionando a priori imprópria $p(\phi, \Sigma_u) \propto |\Sigma_u|^{-(n+1)/2}$, tem-se

$$\begin{aligned} p(Y^*(\theta)|\phi, \Sigma_u) &= c(\theta) |\Sigma_u|^{-\frac{\lambda T + n + 1}{2}} \exp\left\{-\frac{1}{2}Tr [\lambda T\Sigma_u^* \Sigma_u^{-1}]\right\} \\ &\times \exp\left\{-\frac{1}{2}Tr [\Sigma_u^{-1}(\phi - \phi^*)' \lambda T\Gamma_{xx}^* (\phi - \phi^*)]\right\} \end{aligned} \quad (7.24)$$

onde $c(\theta)$ é uma constante de integração escolhida de forma a garantir que a densidade integre em um (a equação de $c(\theta)$ pode ser encontrada no Apêndice A).

Condicional ao vetor de parâmetros do DSGE, a distribuição a priori dos parâmetros do VAR é proporcional a uma distribuição Normal-Wishart-Invertida

$$\Sigma_u | \theta \sim Wi(\lambda T\Sigma_u^*, \lambda T - k, n) \quad (7.25)$$

$$\phi | \Sigma_u, \theta \sim N(\phi^*, \Sigma_u \otimes (\lambda T\Gamma_{xx}^*)^{-1}) \quad (7.26)$$

Note que a estrutura da a priori torna-se completa com a distribuição dos parâmetros do modelo DSGE, ou seja, a priori tem uma estrutura hierárquica

$$p(\phi, \Sigma_u, \theta) = p(\phi, \Sigma_u | \theta) p(\theta). \quad (7.27)$$

A habilidade de computar os momentos populacionais do modelo DSGE e o fato de usar a priori conjugadas torna a abordagem de Del Negro e Schorfheide (2004) bastante eficiente do ponto de vista computacional.

7.3 DISTRIBUIÇÃO A POSTERIORI

A distribuição a posteriori pode ser fatorada em uma densidade a posteriori dos parâmetros do VAR, dado os parâmetros do modelo DSGE, e uma densidade marginal a posteriori dos parâmetros do modelo DSGE, de acordo com

$$p(\phi, \Sigma_u, \theta|Y) = p(\phi, \Sigma_u|Y, \theta)p(\theta|Y). \quad (7.28)$$

Dessa forma, dado que a priori é conjugada natural, a distribuição a posteriori também seguirá uma distribuição Normal-Wishart-invertida condicional ao vetor de parâmetros do modelo DSGE.

A densidade a posteriori é, por definição, proporcional ao produto da verossimilhança e da priori. Portanto, utilizando as expressões da verossimilhança em (7.12) e da a priori em (7.24), a posteriori será dada por (a demonstração completa está no Apêndice A):

$$\begin{aligned} p(\phi, \Sigma_u|Y, \theta) &\propto |\Sigma_u|^{-\frac{(1+\lambda)T+n+1}{2}} \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(1+\lambda)T\tilde{\Sigma}_u]\right\} \\ &\times \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(\phi - \tilde{\phi})'(\lambda T\Gamma_{xx}^* + X'X)(\phi - \tilde{\phi})]\right\}. \end{aligned} \quad (7.29)$$

Portanto, a posteriori é proporcional ao produto de uma densidade Normal para $\phi|\Sigma_u, \theta$ vezes uma densidade Wishart-Invertida para $\Sigma_u|\theta$

$$\Sigma_u|Y, \theta \sim Wi((\lambda + 1)T\tilde{\Sigma}_u, (1 + \lambda)T - k, n) \quad (7.30)$$

$$\phi|\Sigma_u, Y, \theta \sim N(\tilde{\phi}, \Sigma_u \otimes (\lambda T\Gamma_{xx}^* + X'X)^{-1}). \quad (7.31)$$

Usualmente a priori para θ não tem uma forma padrão simples, assim a distribuição marginal a posteriori para θ não terá fórmula fechada. Para calcular a distribuição utiliza-se métodos numéricos, como descrito a seguir na seção 7.5.

7.4 A ESCOLHA DE λ

O hiperparâmetro λ determina a importância das informações a priori em relação às informações amostrais. Se λ é pequeno, a priori é difusa e as observações amostrais dominam as observações artificiais na posteriori. No caso extremo, em que $\lambda = 0$, a média a posteriori ϕ colapsa nas estimativas de máxima verossimilhança. Por outro lado, valores altos para λ fazem com que a priori concentre mais massa de probabilidade ao longo das funções de restrições $\phi^*(\theta)$ e $\Sigma_u^*(\theta)$.

Seja Λ um conjunto possível de valores de λ , isto é $\Lambda \equiv \{\lambda_1, \lambda_2, \dots, \lambda_q\}$. A abordagem de Del Negro e Schorfheide (2004) é comparar o modelo avaliado em cada $\lambda \in \Lambda$, usando como medida a densidade marginal dos dados, $p_\lambda(Y)$. Isto é, a probabilidade dos dados observados dado um certo λ , que funciona de maneira semelhante à um critério de informação e, pode ser obtida através da integração, em relação aos parâmetros, da densidade conjunta dos dados e dos parâmetros

$$p_\lambda(Y) = \int p(Y|\theta, \Sigma_u, \phi)p(\theta, \Sigma_u, \phi)d(\theta, \Sigma_u, \phi). \quad (7.32)$$

Integrações envolvendo o cálculo de densidades marginais dos dados são computacionalmente intensas. Entretanto, note que a densidade a priori conjunta dos parâmetros do VAR e do DSGE, $p(\phi, \Sigma_u, \theta)$, pode ser decomposta como $p(\phi, \Sigma_u|\theta)p(\theta)$ e, a priori dos parâmetros do VAR dado θ tem forma Normal-Wishart-invertida, permitindo que a integral com relação aos parâmetros do VAR seja calculada analiticamente (ver apêndice A), deixando apenas a integral com relação a θ a ser calculada para aproximar $p_\lambda(Y)$.

$$p_\lambda(Y) = \int p_\lambda(Y|\theta)p(\theta)d(\theta). \quad (7.33)$$

O λ ótimo, denotado por $\hat{\lambda}$, é escolhido de forma a maximizar $p_\lambda(Y)$, isto é,

$$\hat{\lambda} = \arg \max_{\lambda \in \Lambda} p_\lambda(Y). \quad (7.34)$$

Del Negro e Schorfheide (2004) ainda sugerem que a densidade marginal dos dados pode ser utilizada para escolher a ordem de defasagem apropriada para o VAR.

7.5 ALGORITMO METROPOLIS-HASTINGS

Quando o objetivo é realizar inferência Bayesiana, é necessário calcular uma distribuição conjunta a posteriori para as variáveis aleatórias sob análise. Porém, geralmente isso requer o cálculo de integrais intratáveis. Nesses casos, uma solução é prosseguir com técnicas de amostragem baseado em métodos de Monte Carlo via Cadeias de Markov (MCMC). Métodos MCMC são uma classe de algoritmos construídos de maneira que seja possível amostrar da cadeia de Markov, cuja distribuição invariante é igual à densidade objetivo desconhecida. Ao usar MCMC, estima-se a distribuição a posteriori e as integrais usando, após a convergência, amostras simuladas da cadeia de Markov.

O algoritmo Metropolis-Hastings (MH) é um método MCMC que gera amostras a partir da distribuição a posteriori desconhecida através de amostras de uma distribuição candidata.

O algoritmo MH inicia com a densidade objetivo (*target density*), normalmente a densidade a posteriori, $p(\theta|Y)$. Uma distribuição candidata é então escolhida. A chave do algoritmo MH é justamente a distribuição candidata $q(\nu|\theta^{i-1})$, que potencialmente depende da realização θ^{i-1} na iteração $i-1$ do algoritmo. Se a realização da densidade candidata não é aceita, então a cadeia não irá se mover e $\theta^i = \theta^{i-1}$. Note que realizações da densidade candidata não são aceitas automaticamente como amostras da distribuição a posteriori. Essas amostras candidatas são aceitas com base em uma probabilidade de aceitação, $\alpha(\nu|\theta^{i-1})$. E somente serão aceitas como amostras da distribuição a posteriori após a convergência da cadeia de Markov para a sua distribuição invariante.

O Algoritmo 1 fornece detalhes de um algoritmo MH genérico.

O primeiro passo é atribuir um valor inicial para cada variável aleatória contida no vetor θ . O algoritmo consiste em três componentes: (i) gerar uma amostra candidata ν a partir da distribuição candidata $q(\nu|\theta^{i-1})$; (ii) computar a probabilidade de aceitação via $\alpha(\nu|\theta^{i-1})$, com base na distribuição proposta e na densidade a posteriori; (iii) aceitar a amostra candidata com probabilidade α , ou rejeitá-la com probabilidade $1 - \alpha$.

Uma vez que $p(\theta|Y) \propto p(Y|\theta)p(\theta)$, pode-se substituir as densidades a posteriori no cálculo da probabilidade de aceitação $\alpha(\nu|\theta^{i-1})$ pelo produto da verossimilhança e da priori, tornando desnecessário conhecer a densidade marginal dos dados $p(Y)$.

Intuitivamente, a função de aceitação do MH é projetada para encontrar o equilíbrio entre a tendência do amostrador de visitar áreas

Algoritmo 1 Algoritmo Metropolis-Hastings Genérico

Inicializar $x^{(0)}$

for $i = 1 : N$

1. Propor ν a partir da densidade $q(\nu|\theta^{i-1})$.
2. Probabilidade de aceitação

$$\alpha(\nu|\theta^{i-1}) = \min \left\{ 1, \frac{p(Y|\nu)p(\nu)/q(\nu|\theta^{i-1})}{p(Y|\theta^{i-1})p(\theta^{i-1})/q(\theta^{i-1}|\nu)} \right\}$$

3. Amostrar $u \sim \text{Uniforme}(0, 1)$

if $u < \alpha$

$\theta^i = \nu$

else

$\theta^i = \theta^{i-1}$.

end if

end for

de maior probabilidade sob a densidade a posteriori e o fato que o amostrador deve explorar todo o suporte de $p(\theta|Y)$, evitando ficar preso. A forma particular da função de aceitação do MH é o que garante que o algoritmo MH satisfaz a *condição de reversibilidade*¹, que garante que a distribuição estacionária do algoritmo MH é de fato a densidade objetivo que se está interessado.

Como em qualquer método MCMC, as realizações geradas pelo algoritmo MH são consideradas como amostras da densidade objetivo apenas depois que a cadeia ultrapassou o estágio de transição e o efeito dos valores iniciais se tornaram tão pequenos que podem ser ignorados. De fato, a convergência para a distribuição invariante ocorre sob as condições necessárias de irredutibilidade e aperiodicidade.

Para implementar o algoritmo MH, é necessário especificar a densidade candidata. Uma prática comum na literatura, principalmente na estimação de modelos DSGE, é assumir que a distribuição candidata $q(\nu|\theta^{i-1})$ pode ser expressa como um passeio aleatório $\nu = \theta^{i-1} + \eta$, em que η é uma realização de uma distribuição centrada em zero. O fato da candidata ser igual ao valor atual mais um ruído fez com que esse algoritmo MH ficasse conhecido como algoritmo *Random Walk Metropolis-Hastings* (RWMH).

A variável aleatória incremental η é definida com média zero e

¹ver Chib e Greenberg (1995) para um detalhamento maior.

variância $c^2\hat{\Sigma}$ e, dado a natureza simétrica da distribuição candidata, a probabilidade de aceitação torna-se

$$\alpha(\nu|\theta^{i-1}) = \min \left\{ \frac{p(\nu|Y)}{p(\theta^{i-1}|Y)}, 1 \right\}. \quad (7.35)$$

Uma realização ν é aceita com probabilidade um se a posteriori em ν é maior do que o valor da posteriori em θ^{i-1} .

Para implementar o algoritmo RWMH o pesquisador ainda precisa especificar c e $\hat{\Sigma}$. Uma abordagem popular, usada em Schorfheide (2000), é definir $\hat{\Sigma}$ como sendo o negativo do inverso da matriz Hessiana, avaliado na moda do log a posteriori, obtida através de uma rotina de otimização numérica, realizada antes do MCMC. Já o parâmetro c é tipicamente ajustado para garantir uma taxa de aceitação razoável do algoritmo. Segundo Herbst e Schorfheide (2015), a maioria dos pesquisadores define como aceitável uma taxa de aceitação entre 0.20 e 0.40.

7.6 IDENTIFICAÇÃO DOS CHOQUES

Para computar as respostas dinâmicas de produto, inflação, taxa de juros e taxa de câmbio às mudanças não antecipadas na política monetária e outros choques estruturais é necessário determinar um mapeamento entre os choques estruturais ϵ_t e os erros de previsão um passo a frente u_t . Seja, $\Sigma^{\frac{1}{2}}$ a decomposição de Cholesky de Σ_u . Sabe-se que em qualquer VAR estrutural exatamente identificado, a relação entre u_t e ϵ_t pode ser caracterizada como

$$u_t = \Sigma^{\frac{1}{2}}\Omega\epsilon_t \quad (7.36)$$

onde Ω é uma matriz ortogonal. O impacto inicial de ϵ_t nas variáveis endógenas, y_t no VAR é dado por

$$\left(\frac{\partial y_t}{\partial \epsilon_t'} \right)_{VAR} = \Sigma^{\frac{1}{2}}\Omega. \quad (7.37)$$

O problema da identificação surge do fato de que os dados não dão informações sobre a escolha da matriz de rotação Ω , isto é, a função de verossimilhança é invariante para Ω .

Del Negro e Schorfheide (2004) propõe uma abordagem em que utilizam o fato que os choques no modelo DSGE são exatamente identi-

ficados, no sentido de que para cada valor de θ existe uma única matriz $A_0(\theta)$, obtida da representação espaço de estado (6.78) que determina o efeito contemporâneo de ϵ_t em y_t . Usando uma fatoração QR de $A_0(\theta)$, a resposta inicial de y_t ao choque estrutural pode ser unicamente decomposta em

$$\left(\frac{\partial y_t}{\partial \epsilon'_t} \right)_{DSGE} = A_0(\theta) = \Sigma^{\frac{1}{2}*}(\theta)\Omega^*(\theta) \quad (7.38)$$

onde $\Sigma^{\frac{1}{2}*}(\theta)$ é uma matriz triangular inferior e $\Omega^*(\theta)$ é ortogonal. Para identificar o VAR, mantém-se a triangulação da sua matriz de covariância Σ_u e substitui a rotação Ω na Equação (7.37) pela função $\Omega^*(\theta)$ que aparece em (7.38).

A implementação desse procedimento de identificação ocorre da seguinte forma: uma vez que é possível gerar realizações a partir da distribuição a posteriori conjunta de ϕ , Σ_u , e θ , para cada realização (i) usar ϕ para construir uma representação MA de y_t em termos dos choques na forma reduzida u_t ; (ii) computar a decomposição de Cholesky de Σ_u , e (iii) calcular $\Omega = \Omega^*(\theta)$ para obter a representação em termos dos choques estruturais ϵ_t .

7.7 PREVISÕES FORA DA AMOSTRA

O objetivo fundamental das previsões utilizando abordagem Bayesiana é a distribuição preditiva, a distribuição $p(y_{T+1:T+h}|Y_T)$ das observações futuras, $Y_{T+1:T+h} = (y'_{T+1}, \dots, y'_{T+h})'$ condicional as observações atualmente observadas, $Y_T = \{y_t\}_{t=1}^T$.

Uma vez que deseja-se realizar previsões h -passos a frente da variável dependente, geradas a partir do seguinte modelo

$$Y_{T+h} = \bar{X}\phi + u^* \quad u^* \sim N(0, \Sigma_u) \quad (7.39)$$

em que \bar{X} é uma matriz com dimensão $h \times k$, então densidade preditiva pode ser representada por

$$p(Y_{T+h}|Y) = \int \int \int \overbrace{p(Y_{T+h}|Y, \phi, \Sigma_u, \theta, \bar{X})}^A \times \underbrace{p(\phi|\Sigma_u, \theta, Y)p(\Sigma_u|y, \theta)p(\theta|Y)}_B d\phi d\Sigma_u d\theta \quad (7.40)$$

Como u é independente de Y , logo Y_{T+h} também é independente de Y , então:

$$p(Y_{T+h}|Y, \phi, \Sigma_u, \theta, \bar{X}) = p(Y_{T+h}|\phi, \Sigma_u, \theta, \bar{X}). \quad (7.41)$$

Além disso, no restante das demonstrações, \bar{X} será suprimido com o objetivo de simplificar a notação.

Em (7.40) A é a verossimilhança para h novos dados e pode ser escrito como núcleo de uma distribuição Normal-Wishart-invertida. Já B é a distribuição posteriori e tem a forma de uma distribuição Normal para ϕ dado Σ_u e θ e uma distribuição Wishart-invertida para Σ_u dado θ . Então a densidade preditiva pode ser escrita como

$$p(Y_{T+h}|Y) = \int \int \int \underbrace{K_N(\phi|\Sigma_u, \theta, Y_{T+h})K_{Wi}(\Sigma_u|\theta, Y, Y_{T+h})}_{\text{verossimilhança}} \underbrace{p(\phi|\Sigma_u, \theta, Y)p(\Sigma_u|y, \theta)p(\theta|Y)}_{\text{posteriori}} d\phi d\Sigma_u d\theta. \quad (7.42)$$

Considerando a distribuição preditiva para um passo frente, Y_{T+1} , como resultado das integrais em ϕ e Σ_u obtém-se uma distribuição t multivariada com parâmetros $(\lambda + 1)T - k$, σ , $(\lambda + 1)T\tilde{\Sigma}_u$ e ν , cuja demonstração pode ser encontrada no apêndice A. Entretanto, a integral em relação à θ não possui fórmula fechada, exigindo que utilize-se métodos de amostragem como o RWMH.

Já para $h > 1$ são necessários esquemas de simulação para gerar uma amostra a partir da densidade preditiva. É bastante simples aplicar o procedimento conhecido como *sampling the future* de Thompson e Miller (1986) para modelos DSGE-VAR. O algoritmo pode ser descrito como:

Algoritmo 2 Simulando a distribuição preditiva com uma posteriori Normal-Wishart-invertida

1. Gerar $\theta^{(j)}$ de $p(\theta|Y)$ usando o algoritmo Metropolis-Hastings Passeio Aleatorio;
 2. Gerar $\Sigma_u^{(j)}$ de $p(\Sigma_u|\theta, Y)$;
 3. Gerar $\phi^{(j)}$ de $p(\phi|\Sigma_u, \theta, Y)$;
 4. Amostrar U_{T+i} a partir de $N(0, \Sigma_u)$ para $i = 1, \dots, h$ e alimentar (7.39) recursivamente;
 5. Repetir o quarto passo M_1 vezes para os mesmos (ϕ, Σ_u, θ) ;
 6. Repetir os passos de um a quatro M_2 vezes.
-

Parte III

APLICAÇÃO EMPÍRICA E RESULTADOS

8 IMPLEMENTAÇÃO PRÁTICA

8.1 DADOS E IMPLEMENTAÇÃO PRÁTICA

A fim de empregar o modelo DSGE de pequena economia aberta descrito no capítulo 6 como informação a priori para um VAR para as variáveis produto, inflação, taxa de juros nominal e taxa de câmbio real efetiva da economia brasileira, adotou-se dados mensais de janeiro de 2003 até dezembro de 2016. Para o produto utilizou-se o Índice de Atividade Econômica do Banco Central (IBC-Br), enquanto que para inflação utilizou-se o Índice de Preços ao Consumidor Amplo (IPCA). Para a taxa de juros utilizou-se a taxa Selic efetiva. Finalmente, para a taxa de câmbio utilizou-se o Índice de Taxa de Câmbio Efetiva Real. A Tabela 7 que contém informações a respeito dos dados pode ser encontrada no apêndice B.

A implementação prática ocorreu de acordo com os passos descritos a seguir¹:

1. O primeiro passo foi especificar a priori para os parâmetros do modelo DSGE. A escolha da a priori para os parâmetros estimados foi determinada por implicações teóricas do modelo e evidências de estudos prévios. Essa escolha também refletiu a imposições de restrições de não-negatividade ou intervalos para determinados parâmetros. Por exemplo, as distribuições Beta foram escolhidas para parâmetros cujo intervalo está definido entre 0 e 1. Já a distribuição Normal foi escolhida para representar os parâmetros cujo domínio são os números reais);
2. Uma vez que as distribuições a priori dos parâmetros do DSGE foram definidas, o modelo foi solucionado utilizando o algoritmo de Sims (2002), a fim de encontrar a representação espaço de estados, conectando o modelo teórico às variáveis observadas através da equação de medida. Usando o otimizador numérico *csmmwel*² de Christopher Sims, encontrou-se a moda do logaritmo posteriori. O negativo do inverso da matriz Hessiana, avaliado nesse ponto de máximo da distribuição a posteriori, foi utilizado como matriz de covariância da densidade candidata no algoritmo *Random Walk* Metropolis-Hastings (RWMH). Além disso, determinou-se

¹Para obter os resultados elaborou-se rotinas próprias no MATLAB.

²Disponível em <http://sims.princeton.edu/yftp/optimize/>.

o parâmetro $c = 0.3$ de forma a obter-se uma taxa de aceitação de 27%;

3. Com a densidade candidata em mãos, o algoritmo RWMH foi usado para explorar a distribuição a posteriori de θ (função de verossimilhança marginal de θ vezes a priori de θ). Uma vez que os parâmetros do VAR, condicionais a θ , têm distribuição a posteriori com fórmula fechada, obter realizações dessa distribuição a posteriori é simples. Diagnósticos de convergência foram realizados com base em Geweke (1992);
4. Baseado nessas realizações, construiu-se um *grid* para diferentes valores de λ , $\Lambda = [\lambda_1, \dots, \lambda_q]$, e utilizou-se o estimador de média harmônica modificada de Geweke (1999) para obter-se a aproximação numérica da densidade marginal dos dados $p_\lambda(Y)$ e, então, encontrou-se o valor ótimo $\hat{\lambda}$ que maximiza a densidade marginal dos dados;
5. Uma vez que o valor ótimo $\hat{\lambda}$ foi determinado, novamente utilizou-se a densidade marginal dos dados a fim de selecionar a ordem de defasagem ótima do modelo VAR, de forma a maximizar $p_{\hat{\lambda}}(Y)$;
6. Foram selecionadas as realizações de θ , ϕ e Σ_u que correspondem a $\hat{\lambda}$ e a ordem de defasagem ótima do VAR e realizadas inferências sobre o modelo DSGE-VAR;
7. Finalmente, a partir das realizações de ϕ , Σ_u , e θ , foi possível gerar previsões fora da amostra a partir de modelos DSGE-VAR com diferentes valores de λ e compará-las com as previsões geradas pelos modelos DSGE e BVAR com a priori Minnesota, através do procedimento de comparação proposto por Hansen, Lunde e Nason (2011).

8.2 A PRIORI DE MINNESOTA

Esta seção descreve brevemente a metodologia da a priori de Minnesota, utilizada como modelo de previsão alternativo empregado para comparar a acurácia das previsões obtidas com os modelos DSGE-VAR. Destaca-se que a seção baseia-se fortemente nas discussões disponíveis em Koop e Korobilis (2010).

Considerando a Equação (7.4), é possível rescrevê-la como $y = (I_n \otimes Z)\alpha + \epsilon$, onde $\epsilon \sim N(0, \Sigma_u \otimes I_n)$ e $\alpha = \text{vec}(\phi)$. A aproximação

usada na a priori de Minnesota é a substituição da matriz desconhecida Σ_u por uma estimativa $\hat{\Sigma}_u$ dela, assumindo que Σ_u é uma matriz diagonal. Dado isso, cada equação do VAR pode ser estimada independentemente, e pode-se definir $\hat{\sigma}_{ii} = s_i^2$, onde s_i^2 é a estimativa de MQO da variância do erro na i -ésima equação e $\hat{\sigma}_{ii}$ é o elemento da coluna i e linha i de $\hat{\Sigma}_u$.

Após substituir Σ_u por $\hat{\Sigma}_u$ é necessário apenas preocupar-se com a a priori para α , a qual assume-se ser

$$\alpha \sim N(\underline{\alpha}_{Min}, \underline{V}_{Min}). \quad (8.1)$$

Para a média a priori, $\underline{\alpha}_{Min}$, quando os dados estão em nível, $\underline{\alpha}_{Min} = 0_K$, $K = (1 + n \times p)$, exceto para os elementos da primeira defasagem da própria variável dependente em cada equação, que, por sua vez, são definidos igual a um. No caso de dados em primeira diferença ou variáveis sem tendência, $\underline{\alpha}_{Min} = 0_K$. A priori de Minnesota assume que a matriz de covariância a priori, \underline{V}_{Min} , é diagonal. Definindo \underline{V}_i como um bloco de \underline{V}_{Min} associado com K coeficientes na equação i e $\underline{V}_{i,jj}$ como seus elementos diagonais, então tradicionalmente a priori de Minnesota pode ser definida como

$$\underline{V}_{i,jj} = \begin{cases} \frac{a_1}{r^2} & \text{para os coeficientes da própria defasagem } r \\ & \text{para } r = 1, \dots, p \\ \frac{a_2}{r^2} \frac{\sigma_{ii}}{\sigma_{jj}} & \text{para coeficientes de defasagem} \\ & r \text{ da variável } j \text{ na equação de } i \\ a_3 \sigma_{ii} & \text{para coeficientes de variáveis exógenas.} \end{cases} \quad (8.2)$$

A forma de $\underline{V}_{i,jj}$ impõe o fato de que os hiperparâmetros a_1 , a_2 e a_3 da a priori de Minnesota são escolhidos, normalmente, de forma a colocar menos importância nas defasagens de outras variáveis e mais importância às próprias defasagens. A escolha exata dos valores de a_1 , a_2 e a_3 dependerá da aplicação empírica em questão.

Entre as grandes vantagens da a priori de Minnesota está a possibilidade de gerar inferência a posteriori em fórmula analítica, além de permitir maior flexibilidade para definir o conhecimento a priori. Usando a priori de Minnesota é possível mostrar que a função densidade de probabilidade a posteriori tem distribuição Normal dada por:

$$\alpha|y \sim N(\bar{\alpha}_{Min}, \bar{V}_{Min}) \quad (8.3)$$

onde

$$\bar{V}_{Min} = [\underline{V}_{Min}^{-1} + (\hat{\Sigma}_u^{-1} \otimes (Z'Z))]^{-1} \quad (8.4)$$

e

$$\bar{\alpha}_{Min} = \bar{V}_{Min} [\underline{V}_{Min}^{-1} \alpha_{Min} + (\hat{\Sigma}_u^{-1} \otimes Z)' y]. \quad (8.5)$$

Entretanto, a desvantagem dessa abordagem está no fato de que ela não fornece um tratamento Bayesiano completo ao parâmetro desconhecido Σ_u . Ao invés disso, simplesmente substitui-se Σ_u por $\hat{\Sigma}_u$, ignorando a incerteza a respeito desse parâmetro.

A expectativa preditiva de y_{T+h} é computada por substituição recursiva, assim, uma análise Bayesiana completa requer

$$E(y_{T+h}|Y_T) = E_{\hat{\Sigma}_u|Y_T}[E(y_{T+h}|Y_T, \hat{\Sigma}_u)], \quad (8.6)$$

isto é, a integral com respeito a posteriori de $\hat{\Sigma}_u$.

8.3 COMPARANDO PREVISÕES

Com o objetivo de comparar o desempenho, em termos de previsão, dos modelos adotou-se o procedimento proposto por Hansen, Lunde e Nason (2011), o *model confidence set* (MCS). Intuitivamente, o objetivo do MCS é determinar um conjunto M^* , que consiste nos “melhores” modelos de um conjunto de modelos, denotado por M_0 . O procedimento MCS fornece um conjunto de confiança de modelos, \hat{M}^* , que é um conjunto de modelos construído de maneira a conter os melhores modelos dado um nível de confiança, $1 - \alpha$. Os modelos em \hat{M}^* são avaliados usando informações amostrais sobre o desempenho relativo dos modelos em M_0 .

O ponto de partida para a aplicação do procedimento MCS é um conjunto finito de modelos de previsão M_0 . Os modelos em M_0 são avaliados em termos de uma função perda, assim, denota-se a perda associada ao modelo i no período t como $L_{i,t}$, em que $t = 1, \dots, n$. Define-se, então, o desempenho relativo como:

$$d_{i,j,t} \equiv L_{i,t} - L_{j,t} \quad \text{para todo } i, j \in M_0. \quad (8.7)$$

Assumindo que $\mu_{ij} \equiv E(d_{i,j,t})$ é finito e não depende de t para todo $i, j \in M_0$, portanto, ordena-se as alternativas em termos das perdas esperadas, assim, a alternativa i é preferida à alternativa j se $\mu_{ij} < 0$.

Então, o objetivo do MCS é identificar o conjunto M^* , tal que:

$$M^* = \{i \in M_0 : \mu_{i,j} \leq 0 \text{ para todo } j \in M_0\}. \quad (8.8)$$

Isso é feito através de uma sequência de testes de significância, onde modelos que são considerados significativamente inferiores aos outros elementos de M_0 são eliminados. As hipóteses são testadas da seguinte maneira:

$$H_{0,M} : \mu_{ij} = 0 \text{ para todo } i, j \in M, \quad (8.9)$$

onde $M \subset M_0$. A hipótese alternativa, $\mu_{ij} \neq 0$ para algum $i, j \in M$ é denotada por $H_{A,M}$.

O MCS baseia-se em uma regra de eliminação e_M e um teste de equivalência, δ_M . Usa-se então δ_M ao nível α para testar a $H_{0,M}$. Se a hipótese de equivalência for rejeitada, a regra de eliminação é utilizada para remover um modelo M que tem desempenho inferior. O procedimento repete-se até que a hipótese de equivalência não seja mais rejeitada, de forma que o MCS forneça o conjunto dos melhores modelos. O algoritmo do MCS pode ser descrito como

O MCS propõe a implementação do método bootstrap de forma a evitar a estimação de uma matriz de covariância dos erros de previsão de grande dimensão. O algoritmo a seguir descreve a implementação do bootstrap.

Algoritmo 3 Model Confidence Set

1. Inicializar o procedimento fazendo $M = M_0$.
2. Testar a equivalência dos modelos em M .
 - (a) Se o teste de equivalência não é rejeitado, $\hat{M}^* \equiv M$
 - (b) Se o teste de equivalência é rejeitado
 - i. Então definir

$$\bar{d}_i \equiv \frac{1}{m} \sum_{j \in M} \bar{d}_{ij}, \quad (8.10)$$

onde m denota o número de modelos em M . A estatística \bar{d}_i mede o desempenho do modelo i relativo à média de todos os modelos.

- ii. Determinar o “pior” modelo de M , definido como

$$i^* \equiv \arg \max_{i \in M} \frac{\bar{d}_i}{\sqrt{\hat{v}ar(\bar{d}_i)}} \quad (8.11)$$

onde $\hat{v}ar(\bar{d}_i)$ é uma estimativa de $var(\bar{d}_i)$.

- iii. Remover o modelo i^* de M , e repetir o passo 2.
-

Algoritmo 4 Implementação do Método Bootstrap

1. Computar as reamostras para $b = 1, \dots, B$.
2. Estimar a variância dos parâmetros como a seguir
 - (a) Definir as médias das re-amostras do bootstrap, $\bar{d}_{b,ij}^*$, e calcular a estimativa bootstrap de $var(\bar{d}_{ij})$ dada por

$$\hat{var}(\bar{d}_{ij}) \equiv \frac{1}{B} \sum_{b=1}^B (\bar{d}_{b,ij}^* - \bar{d}_{ij})^2, \quad \text{para } i, j \in M. \quad (8.12)$$

- (b) De forma similar, calcular $\bar{d}_{b,i}^* \equiv \frac{1}{m-1} \sum_{j \in M} \bar{d}_{b,ij}^*$ e a estimativa bootstrap de $var(\bar{d}_i)$ dada por

$$\hat{var}(\bar{d}_i) \equiv \frac{1}{B} \sum_{b=1}^B (\bar{d}_{b,i}^* - \bar{d}_i)^2, \quad \text{para } i \in M. \quad (8.13)$$

3. Calcular as estatísticas de interesse e obter suas distribuições bootstrap
 - (a) Calcular a estatística $T_R = \max_{i,j \in M} \frac{|\bar{d}_{ij}|}{\sqrt{\hat{var}(\bar{d}_{ij})}}$
 - (b) A distribuição bootstrap estimada de T_R , sob a hipótese nula, é dada por

$$T_{b,R}^* = \max_{i,j \in M} \frac{|\bar{d}_{b,ij}^* - \bar{d}_{ij}|}{\sqrt{\hat{var}(\bar{d}_{ij})}}. \quad (8.14)$$

- (c) Um p-valor do teste de equivalência é obtido calculando a taxa de $T_{b,R}^*$ que excede T_R , isto é,

$$\hat{p} - \text{valor} \equiv \frac{1}{B} \sum_{b=1}^B 1\{T_{b,R}^* > T_R\}. \quad (8.15)$$

8.4 DISTRIBUIÇÃO A PRIORI DE θ

As distribuições a priori refletem o conhecimento ou crença do pesquisador antes de observar os dados. Atribuir valores grandes aos desvios padrão significa que há pouca informação adicional aos dados. Por outro lado, desvios padrão pequenos indicam que há certa confiança de que o valor do parâmetro é um valor próximo à média a priori.

Para a implementação prática, as distribuições a priori foram determinadas com base em estudos anteriores para o Brasil, a saber: Silveira (2008) e Linardi (2016). Além disso, tomou-se como base alguns valores adotados no exercício de calibração realizado em Gali e Monacelli (2005).

Todos os resultados empíricos foram gerados a partir das distribuições a priori indicadas na Tabela 1. Os demais parâmetros têm seus valores calibrados indicados na Tabela 9, no apêndice B.

Tabela 1: Distribuição a Priori para os Parâmetros do Modelo DSGE

Nome	Domínio	Densidade	Média	DP
σ	\mathbb{R}^+	Gamma	1,0000	0,1000
φ	\mathbb{R}^+	Gamma	0,2500	0,1000
θ	[0,1)	Beta	0,5000	0,1500
ϕ_π	\mathbb{R}^+	Gamma	1,7500	0,1500
ϕ_y	\mathbb{R}^+	Gamma	0,5000	0,1000
ρ_a	[0,1)	Beta	0,5000	0,1500
ρ_y	[0,1)	Beta	0,5000	0,1500
ρ_z	[0,1)	Beta	0,5000	0,1500
γ^*	\mathbb{R}	Normal	0,1000	0,1500
q^*	\mathbb{R}	Normal	0,5000	0,1500
π^*	\mathbb{R}	Normal	-0,1000	0,1500
ρ	\mathbb{R}	Normal	1,0000	0,5000
σ_i	\mathbb{R}^+	Inv. Gamma	0,0100	4,0000
σ_a	\mathbb{R}^+	Inv. Gamma	0,0100	4,0000
σ_y	\mathbb{R}^+	Inv. Gamma	0,0100	4,0000
σ_z	\mathbb{R}^+	Inv. Gamma	0,0100	4,0000

Nota: DP indica desvio padrão.

Distribuições Beta com média 0.5 e desvio padrão igual à 0.15 foram selecionadas para os parâmetros que indicam a persistência dos choques tecnológico, de preferência e no produto do resto do mundo. Além disso, a mesma distribuição foi selecionada para o parâmetro de

rigidez de preços de Calvo.

Uma distribuição Gama foi adotada para os parâmetros que representam a elasticidade substituição do consumo e a elasticidade inversa da oferta de trabalho, com médias iguais a 1 e 0.25, respectivamente. Para os parâmetros que representam a resposta da autoridade monetária aos desvios da inflação e do hiato do produto também atribuiu-se uma distribuição Gama com médias a priori iguais à 1.75 e 0.5, respectivamente.

Para γ^* , q^* , π^* e ρ , os parâmetros das equações de medida, adotou-se distribuições Normal com médias a priori iguais à 0.1, 0.5, -0.1 e 1, respectivamente. Finalmente, distribuições Gama inversa foram usadas para a volatilidade dos choques.

8.5 ALGORITMO MCMC E DIAGNÓSTICO DE CONVERGÊNCIA

No exercício empírico realizado foram amostradas 50000 realizações do algoritmo RWMH sobre um *grid* contendo valores de λ , com ordem de defasagem de 1 até 3, sendo que 50% destas foram descartadas com o objetivo de garantir que a cadeia já tivesse convergido para a distribuição invariante antes de realizar inferências sobre a distribuição a posteriori. Ademais, a matriz de covariância da densidade candidata, $\hat{\Sigma}$, do algoritmo RWMH, foi parametrizada com $c = 0,3$ de forma a garantir uma taxa de aceitação de aproximadamente 27%.

Como forma de avaliar os resultados obtidos no algoritmo RWMH, adotou-se o diagnóstico de convergência de cadeias de Markov de Geweke (1992). Segundo Koop (2003), o diagnóstico é baseado na intuição de que se um número suficientemente grande de realizações foi obtido, a média estimada baseada na primeira metade das realizações deve ser essencialmente a mesma que aquela estimada baseada na segunda metade. De maneira geral, o que foi feito foi dividir as S realizações obtidas do algoritmo RWMH em uma amostra inicial, S_0 , descartada como *burn-in*, mantendo S_1 realizações, sendo essas divididas em 3 conjuntos: o primeiro conjunto de realizações S_A , o conjunto do meio S_B e o último conjunto S_C de realizações. Assim, $S_A = 0.1S_1$ e $S_C = 0.4S_1$. O fato de deixar de fora o conjunto S_B garante que o primeiro e último conjuntos de realizações sejam independentes um do outro. Seja \hat{g}_{S_A} e \hat{g}_{S_C} as médias estimadas usando as primeiras S_A replicações após o *burn-in* e as últimas S_C replicações, respectivamente. Definindo $\frac{\hat{\sigma}_A}{\sqrt{S_A}}$ e $\frac{\hat{\sigma}_C}{\sqrt{S_C}}$ os erros padrão numéricos dessas duas estimativas. Então, o

teorema do limite central diz que

$$DC \rightarrow N(0, 1) \quad (8.16)$$

onde DC é o diagnóstico de convergência, dado por

$$DC = \frac{\hat{g}_{S_A} - \hat{g}_{S_C}}{\frac{\hat{\sigma}_A}{\sqrt{S_A}} + \frac{\hat{\sigma}_C}{\sqrt{S_C}}}. \quad (8.17)$$

O diagnóstico de convergência, considerando $\lambda = 0.1$ e VAR(2), foi calculado e comparado com valores críticos de uma tabela estatística Normal padrão, conforme a Tabela 8, no apêndice B. Os resultados indicaram que todos os parâmetros convergiram, dado um nível de significância de 95%.

9 RESULTADOS

9.1 DETERMINAÇÃO DO MODELO

Como descrito na seção 7.4, a densidade marginal dos dados pode ser utilizada como critério de seleção para a determinação do valor ótimo do hiperparâmetro λ e da ordem de defasagem do VAR. As aproximações numéricas das densidades marginais dos dados foram obtidas através do estimador de média harmônica modificada de Geweke (1999) e os valores encontrados para diferentes λ , considerando a ordem de defasagem do VAR de 1 a 3, estão resumidos na Tabela 2.

Tabela 2: Densidade Marginal dos Dados

λ	Densidade Marginal dos Dados		
	VAR (1)	VAR (2)	VAR (3)
0,065	778,2221		
0,075	778,4458		
0,085	778,3661	852,7262	
0,095	778,1712	854,1735	
0,15	775,9776	854,8120	836,8801
0,25	770,7928	849,8689	837,1474
0,50	759,2241	833,5161	828,5080
0,75	749,6497	818,4322	816,2135
1	743,1546	805,1919	806,8621

Nota: Os valores de densidade marginal dos dados em destaque indicam a maior densidade marginal obtida para cada uma das ordens de defasagem.

Duas características importantes merecem a atenção com relação à densidade marginal dos dados. A primeira diz respeito ao fato que independente da ordem de defasagem escolhida, essa função sempre apresenta forma U invertida. Esse comportamento é uma evidência que a inclusão de informações a priori do DSGE são benéficas, entretanto, as densidades baixas para valores de λ altos também indicam que as restrições impostas pelo DSGE não são suportadas pelos dados. Em suma, essa característica sugere que a abordagem DSGE-VAR pode melhorar o ajustamento dos dados.

A segunda característica está ligado ao fato que o valor ótimo de λ , indicado por $\hat{\lambda}$, varia de acordo com a ordem de defasagem escolhida. Na Tabela 2 observa-se que o valor de $\hat{\lambda}$ aumenta sempre que há aumento da ordem de defasagem. A razão para isso pode ser atribuída ao fato que à medida que ocorre o aumento da ordem de defasagem há um aumento no número de parâmetros, reduzindo os graus de liberdade, e tornando a informação a priori mais importante, ou seja, aumentando o valor de $\hat{\lambda}$.

A escolha de um VAR(2) com $\lambda = 0.15$ corresponde a maximização da densidade marginal dos dados, portanto, todas as análises subsequentes serão feitas com base nesta escolha, agora denotada por $\hat{\lambda}$.

9.2 DISTRIBUIÇÃO A POSTERIORI DE θ

O procedimento proposto por Del Negro e Schorfheide (2004) também fornece as estimativas a posteriori para os parâmetros do modelo DSGE. A Tabela 3 traz as médias a posteriori, desvios padrão e intervalos de credibilidade de 90%, considerando o modelo ótimo, ou seja $\hat{\lambda}$, com base em toda a amostra (Jan/2003 a Dez/2016).

É importante perceber que as densidades a posteriori são, na grande maioria, menos difusas que as suas respectivas densidades a priori (ver Tabela 1). Esse resultado sugere que os dados contêm informações capazes de atualizar as distribuições a priori, embora a extensão dessas atualizações varie de parâmetro para parâmetro.

O parâmetro que representa a elasticidade substituição, σ , tem sua densidade a posteriori bem menos difusa que sua a priori, indicando contribuição dos dados. A média estimada igual a 0.9072 é sensivelmente inferior ao resultado obtido por Silveira (2008), entretanto, aproxima-se bastante do resultado de Linardi (2016). Esse resultado aponta maior sensibilidade do consumo às mudanças na taxa real de juros.

A média a posteriori da elasticidade inversa da oferta de trabalho, φ , mostrou-se relativamente inferior aos resultados obtidos por Linardi (2016) e Silveira (2008). Essa média a posteriori sugere uma resposta maior da oferta de mão-de-obra aos salários e é um resultado esperado em países emergentes, onde trabalhadores pobres são menos relutantes em trocar horas de lazer por horas de trabalho.

O parâmetro θ , que representa a rigidez de preços de Calvo, apresentou uma média a posteriori surpreendentemente baixa em relação à

Tabela 3: Distribuições a Posteriori dos Parâmetros

Nome	Posteriori			
	Média	DP	IC 90%	
σ	0,9072	0,0872	[0,7580	1,0418]
φ	0,3516	0,1282	[0,1371	0,5448]
θ	0,0179	0,0068	[0,0073	0,0275]
ϕ_π	1,7153	0,1441	[1,4673	1,9364]
ϕ_y	0,4934	0,0941	[0,3313	0,6369]
ρ_a	0,6320	0,1212	[0,4220	0,8181]
ρ_y	0,5929	0,1318	[0,3758	0,8077]
ρ_z	0,7110	0,1146	[0,5261	0,8918]
γ^*	0,0015	0,0024	[-0,0024	0,0054]
π^*	0,4935	0,0814	[0,3497	0,6176]
q^*	-0,0026	0,0089	[-0,0167	0,0126]
ρ	0,5769	0,1654	[0,2872	0,8246]
σ_a	0,0050	0,0008	[0,0037	0,0063]
σ_y	0,0282	0,0056	[0,0186	0,0370]
σ_z	0,0271	0,0129	[0,0104	0,0422]
σ_i	0,2385	0,0478	[0,1602	0,3138]

Nota: IC 90% significa intervalo de credibilidade de 90%. Os parâmetros σ_a , σ_y , σ_z e σ_i representam os desvios padrão dos choques de produtividade, no produto do resto do mundo, de preferências e monetário, respectivamente.

sua média a priori. Uma possível explicação para este resultado é o fato de que o Brasil é um país que historicamente possui taxas de inflação altas, o que ocasiona uma redução da rigidez de preços. Ademais, os dados parecem ser bastante informativos para este parâmetro, uma vez que a distribuição a posteriori mostrou-se menos difusa que a sua distribuição a priori.

As estimativas dos coeficientes da Regra de Taylor do Brasil para inflação e hiato do produto, ϕ_π e ϕ_y , parecem consistentes com os resultados obtidos em Linardi (2016) e Silveira (2008). Entretanto, a resposta da autoridade monetária aos desvios da inflação é sensivelmente mais alta do que os resultados observados na literatura, indicando que a política monetária reage fortemente às pressões inflacionárias. Por outro lado, a estimativa do parâmetro que indica a resposta da autoridade monetária aos desvios do hiato do produto, mostrou-se relativamente inferior aos resultados de Linardi (2016) e Silveira (2008), indicando que o banco central preocupa-se mais com a estabilização da inflação.

Finalmente, os parâmetros de persistência dos choques de produtividade, no produto do resto do mundo e de preferência, mostraram-se relativamente baixos frente aos resultados encontrados na literatura, indicando que os choques não têm, em geral, um alto grau de inércia. Ademais, suas distribuições a posteriori mostraram-se sensivelmente pouco menos difusas que suas distribuições a priori, sugerindo que os dados são pouco informativos aos conhecimentos do pesquisador.

9.3 PREVISÕES

As previsões para produto, inflação, taxa de juros e taxa de câmbio foram geradas com base em uma amostra de Jan/2003 a Dez/2013 para nove diferentes modelos, a saber: DSGE, BVAR com a priori de Minnesota (BVAR-Minn) e sete modelos DSGE-VAR com diferentes valores de λ , sob ordens de defasagem de 1 a 3. É importante ter em mente que quando $\lambda = 0$ o modelo DSGE-VAR colapsa em um BVAR com uma a priori difusa (BVAR-difusa). Para tal, computou-se previsões fora da amostra para 1, 3 e 6 passos a frente para o período de Jan/2014 a Dez/2016, sempre com base em uma amostra de tamanho fixo, isto é, *rolling window*.

O desempenho dos modelos foi comparado seguindo o procedimento MCS, descrito na seção 8.3, usando uma função perda definida como a média dos erros de previsão ao quadrado (MSE). O espaço de modelos inicial M_0 é composto por 25 modelos (BVAR com a priori

de Minnesota (BVAR-Minn) e sete modelos DSGE-VAR com diferentes valores de λ sob ordem de defasagem iguais a 1, 2 e 3 e o modelo DSGE). As Tabelas 4 até 6 exibem os ganhos (valores positivos) e as perdas (valores negativos) percentuais no MSE relativo ao modelo BVAR com a priori difusa sob ordem de defasagem igual a 1. Isto é, os valores positivos indicam uma melhora percentual em relação ao MSE obtido através do modelo *benchmark*. Já os valores negativos, por suas vez, representam uma deterioração percentual na capacidade preditiva, relativa ao BVAR com a priori difusa. Por construção, os ganhos para BVAR (1) com a priori difusa são zero. Portanto, as entradas nas linhas BVAR-difusa (1) e colunas MSE correspondem ao valor absoluto de MSE obtido a partir do modelo BVAR(1) com a priori difusa e, as linhas abaixo, correspondentes às colunas MSE, representam as melhoras (ou deteriorações) percentuais de cada um dos 24 modelos restantes relativos ao MSE do modelo *benchmark*.

Além disso, nas Tabelas 4 até 6 também contêm os \hat{p} -valores MCS correspondentes a cada um dos modelos, em que os valores em destaque indicam as previsões que pertencem ao $\hat{M}_{95\%}^*$. A interpretação de $\hat{M}_{95\%}^*$ é análoga aos intervalos de confiança da econometria clássica, isto é, $\hat{M}_{95\%}^*$ é um subconjunto aleatório que contém o subconjunto de “melhores” modelos dado uma probabilidade de 95%.

Tabela 4: Ganhos (perdas) percentuais no MSE relativo ao BVAR com a priori difusa e \hat{p} -valores MCS para previsões 1 passo a frente

Modelo	Defasagem	MSE	\hat{p} -valor	MSE	\hat{p} -valor	MSE	\hat{p} -valor	MSE	\hat{p} -valor
		PIB		Câmbio		Inflação		Juros	
BVAR-difusa	1	5e-05	0,61*	1e-03	0,59*	7e-02	0,23*	3e-02	0,03
DSGE-VAR, $\lambda = 0.25$	1	0,030	0,65*	0,014	0,64*	-0,015	0,23*	0,038	0,06*
DSGE-VAR, $\lambda = 0.5$	1	0,026	0,61*	-0,022	0,19*	-0,038	0,23*	0,059	0,10*
DSGE-VAR, $\lambda = 0.75$	1	0,039	0,70*	-0,036	0,19*	-0,088	0,23*	0,067	0,13*
DSGE-VAR, $\lambda = 1$	1	0,046	0,70*	-0,046	0,18*	-0,133	0,23*	0,058	0,10*
DSGE-VAR, $\lambda = 1.5$	1	0,053	0,70*	-0,065	0,18*	-0,223	0,23*	0,021	0,10*
DSGE-VAR, $\lambda = 2$	1	0,062	0,77*	-0,078	0,18*	-0,306	0,13*	-0,021	0,10*
BVAR-Minn	1	-0,067	0,61*	0,077	1,00*	0,039	0,88*	-0,282	0,03
BVAR-difusa	2	0,151	0,82*	-0,043	0,18*	-0,016	0,23*	0,237	0,30*
DSGE-VAR, $\lambda = 0.25$	2	0,149	0,77*	-0,034	0,18*	-0,020	0,23*	0,253	0,37*
DSGE-VAR, $\lambda = 0.5$	2	0,151	0,77*	-0,032	0,18*	-0,050	0,23*	0,277	0,37*
DSGE-VAR, $\lambda = 0.75$	2	0,150	0,77*	-0,069	0,18*	-0,077	0,23*	0,299	0,37*
DSGE-VAR, $\lambda = 1$	2	0,171	0,82*	-0,051	0,18*	-0,119	0,23*	0,304	0,37*
DSGE-VAR, $\lambda = 1.5$	2	0,174	1,00*	-0,069	0,18*	-0,201	0,13*	0,340	0,37*
DSGE-VAR, $\lambda = 2$	2	0,170	0,82*	-0,075	0,18*	-0,268	0,13*	0,345	0,37*
BVAR-Minn	2	0,040	0,65*	0,043	0,88*	0,044	0,88*	0,356	0,59*
BVAR-difusa	3	0,051	0,65*	-0,066	0,18*	-0,013	0,23*	0,349	0,37*
DSGE-VAR, $\lambda = 0.25$	3	0,034	0,61*	-0,047	0,19*	-0,003	0,23*	0,373	0,37*
DSGE-VAR, $\lambda = 0.5$	3	0,102	0,77*	0,053	0,88*	0,006	0,47*	0,378	0,52*
DSGE-VAR, $\lambda = 0.75$	3	0,086	0,73*	-0,046	0,18*	-0,045	0,23*	0,384	0,59*
DSGE-VAR, $\lambda = 1$	3	0,120	0,77*	-0,052	0,18*	-0,072	0,23*	0,390	0,59*
DSGE-VAR, $\lambda = 1.5$	3	0,136	0,82*	-0,058	0,18*	-0,130	0,23*	0,375	0,37*
DSGE-VAR, $\lambda = 2$	3	0,133	0,77*	-0,064	0,18*	-0,189	0,23*	0,405	0,59*
BVAR-Minn	3	0,037	0,61*	0,040	0,88*	0,068	1,00*	0,438	1,00*
DSGE		-0,490	0,00	-0,105	0,18*	-0,928	0,08*	-2,641	0,03

Nota: Entradas nas linhas BVAR-difusa com ordem de defasagem igual a 1 e colunas MSE correspondem ao MSE do modelo BVAR com a priori difusa. As entradas nas outras linhas correspondentes às colunas MSE são ganhos (valores positivos) ou perdas (valores negativos) percentuais em MSE dos 24 demais modelos relativos ao BVAR-difusa (1). As colunas \hat{p} -valor indicam o \hat{p} -valor MCS onde (*) indica que o modelo pertence à $\hat{M}_{95\%}^*$.

A Tabela 4 apresenta os resultados do procedimento MCS considerando um horizonte de previsão de um mês. Para a variável produto é possível notar que, independente da complexidade do modelo, é visível o bom desempenho dos modelos DSGE-VAR. Além disso, é importante perceber que em previsões de curto prazo a tendência é que modelos mais simples, como o modelo BVAR com a priori difusa, apresentem desempenho melhor em relação aos modelos mais complexos. No entanto, o que se observa para a variável produto são os modelos DSGE-VAR superando o *benchmark*. Esse resultado sugere um benefício da inclusão de informações a priori de cunho teórico em previsões de curto prazo. Finalmente, de acordo com o \hat{p} -valor MCS, o modelo que apresentou resultados mais precisos para produto foi o modelo DSGE-VAR com $\lambda = 1.5$ e ordem de defasagem 2.

Já para o câmbio nota-se que os modelos DSGE-VAR apresentam resultados menos acurados, com desempenho inferior ao *benchmark*, sugerindo que as restrições do DSGE distorcem as previsões para a essa variável. A única exceção é o modelo DSGE-VAR com $\lambda = 0.25$ e ordem de defasagem igual 1, cujo desempenho mostrou-se superior ao BVAR(1) com a priori difusa. Nota-se porém, que o modelo BVAR com a priori de Minnesota apresentou desempenho superior ao *benchmark*, sob todas as ordens de defasagem, sendo que o BVAR-Minn(1) foi, segundo o \hat{p} -valor MCS, o modelo que gerou previsões mais precisas. É importante perceber que, em câmbio, todos os modelos pertencem ao subconjunto $\hat{M}_{95\%}^*$ dado um nível de confiança de 95%

Considerando a variável inflação, assim como para câmbio, o desempenho dos modelos DSGE-VAR mostram-se menos precisos, em relação ao *benchmark*. A única exceção é o modelo DSGE-VAR com $\lambda = 0.5$ e ordem de defasagem igual 3, cujo desempenho foi superior ao BVAR(1) com a priori difusa, todavia, deve-se levar em conta que esse ganho em MSE trata-se de um valor ínfimo. O desempenho dos modelos BVAR com a priori de Minnesota mostraram-se relativamente superiores, para todas as ordens de defasagem. Especialmente, o modelo BVAR(3) com a priori de Minnesota mostrou-se o modelo mais preciso, de acordo com o procedimento MCS.

Para taxa de juros, nota-se, na Tabela 4, um desempenho significativo dos modelos DSGE-VAR. Quase todos os modelos DSGE-VAR, sob diferentes valores de λ e diferentes ordens de defasagem, apresentaram resultados mais precisos, em relação ao *benchmark*. A exceção se dá para o modelo DSGE-VAR(1) com $\lambda = 2$, sugerindo que sob um modelo menos complexo, uma imposição de forma mais restrita das informações a priori leva à distorções das previsões. Por outro lado, para

os modelos BVAR com a priori de Minnesota, quando considerado sob uma ordem de defasagem igual a 1, este apresentou resultado inferior ao BVAR(1) difuso, todavia, de acordo com o \hat{p} -valor MCS, BVAR-Minn(3) foi o modelo que apresentou resultados mais precisos, dentre o conjunto de modelos avaliados. É importante destacar que, ainda que o BVAR com a priori de Minnesota tenha apresentado os resultados mais precisos, DSGE-VAR mostrou desempenho competitivo, com ganho percentual em MSE similar.

É interessante perceber que o desempenho insatisfatório dos modelos DSGE relativo ao *benchmark*, no curto prazo, ilustra um ponto importante discutido por Del Negro e Schorfheide (2006) que modelos DSGE apresentam dificuldades em gerar previsões precisas para horizontes de curto e médio prazo.

Tabela 5: Ganhos (perdas) percentuais no MSE relativo ao BVAR com a priori difusa e \hat{p} -valores MCS para previsões 3 passos a frente

Modelo	Defasagem	MSE	\hat{p} -valor	MSE	\hat{p} -valor	MSE	\hat{p} -valor	MSE	\hat{p} -valor
		PIB		Câmbio		Inflação		Juros	
BVAR-difusa	1	1e-04	0,00	2e-03	0,23*	2e-01	0,00	5e-01	0,01
DSGE-VAR, $\lambda = 0.25$	1	0,526	0,00	0,258	0,23*	0,386	0,91*	0,539	0,07*
DSGE-VAR, $\lambda = 0.5$	1	0,501	0,00	0,266	0,31*	0,378	0,06*	0,529	0,14*
DSGE-VAR, $\lambda = 0.75$	1	0,494	0,00	0,266	0,31*	0,357	0,01	0,510	0,16*
DSGE-VAR, $\lambda = 1$	1	0,484	0,00	0,264	0,23*	0,346	0,00	0,497	0,16*
DSGE-VAR, $\lambda = 1.5$	1	0,469	0,00	0,265	0,31*	0,327	0,00	0,458	0,16*
DSGE-VAR, $\lambda = 2$	1	0,463	0,00	0,264	0,23*	0,316	0,00	0,429	0,16*
BVAR-Minn	1	0,476	0,00	0,272	0,38*	0,379	0,06*	0,304	0,04
BVAR-difusa	2	0,220	0,00	0,143	0,23*	0,251	0,00	0,050	0,01
DSGE-VAR, $\lambda = 0.25$	2	0,637	0,81*	0,290	0,86*	0,392	0,91*	0,695	0,16*
DSGE-VAR, $\lambda = 0.5$	2	0,640	1,00*	0,270	0,23*	0,386	0,06*	0,720	0,77*
DSGE-VAR, $\lambda = 0.75$	2	0,599	0,02	0,281	0,38*	0,393	1,00*	0,728	0,81*
DSGE-VAR, $\lambda = 1$	2	0,586	0,02	0,286	0,53*	0,365	0,00	0,722	0,77*
DSGE-VAR, $\lambda = 1.5$	2	0,558	0,00	0,287	0,86*	0,353	0,00	0,734	0,81*
DSGE-VAR, $\lambda = 2$	2	0,539	0,00	0,281	0,38*	0,339	0,00	0,704	0,67*
BVAR-Minn	2	0,546	0,00	0,281	0,53*	0,369	0,00	0,693	0,27*
BVAR-difusa	3	0,305	0,00	-0,449	0,07*	0,186	0,00	0,232	0,06*
DSGE-VAR, $\lambda = 0.25$	3	0,603	0,81*	0,304	1,00*	0,392	0,91*	0,703	0,67*
DSGE-VAR, $\lambda = 0.5$	3	0,604	0,81*	0,268	0,23*	0,387	0,78*	0,699	0,20*
DSGE-VAR, $\lambda = 0.75$	3	0,601	0,81*	0,280	0,38*	0,374	0,06*	0,713	0,77*
DSGE-VAR, $\lambda = 1$	3	0,591	0,02	0,286	0,53*	0,369	0,01	0,721	0,77*
DSGE-VAR, $\lambda = 1.5$	3	0,575	0,02	0,284	0,53*	0,363	0,00	0,707	0,27*
DSGE-VAR, $\lambda = 2$	3	0,563	0,00	0,281	0,53*	0,361	0,00	0,735	0,81*
BVAR-Minn	3	0,552	0,00	0,278	0,31*	0,371	0,06*	0,749	1,00*
DSGE		0,232	0,00	0,248	0,23*	0,256	0,00	0,093	0,14*

Nota: Entradas nas linhas BVAR-difusa com ordem de defasagem igual a 1 e colunas MSE correspondem ao MSE do modelo BVAR com a priori difusa. As entradas nas outras linhas correspondentes às colunas MSE são ganhos (valores positivos) ou perdas (valores negativos) percentuais em MSE dos 24 demais modelos relativos ao BVAR-difusa (1). As colunas \hat{p} -valores indicam o \hat{p} -valor MCS onde (*) indica que o modelo posterior é \hat{M}^* .

Aumentando o horizonte de previsão para 3 passos a frente, é possível perceber, na Tabela 5, uma notável melhora no desempenho dos modelos DSGE-VAR em relação aos resultados apresentados na Tabela 4. Isto é, para horizontes de previsão mais longos, DSGE-VAR apresenta-se expressivamente mais preciso em relação ao *benchmark*.

Para produto, segundo o \hat{p} -valor MCS, observa-se que o modelo DSGE-VAR com $\lambda = 0.5$ apresentou previsões mais acuradas que os demais modelos. Além disso, todos os demais modelos DSGE-VAR apresentaram significativos ganhos em MSE, relativo ao BVAR(1) com a priori difusa, sugerindo o bom desempenho dos modelos DSGE-VAR para horizontes de previsão de médio e longo prazo. A despeito do fato que o desempenho do modelo BVAR com a priori de Minnesota mostrou-se bastante semelhante ao DSGE-VAR, com ganhos expressivos em relação ao *benchmark*, em todas as ordens de defasagem, apenas modelos DSGE-VAR pertencem ao subconjunto $\hat{M}_{95\%}^*$ dado um nível de confiança de 95%

Para câmbio o aumento no horizonte de previsão também gerou sensível melhora na precisão das previsões geradas pelos modelos DSGE-VAR em relação ao BVAR(1) com a priori difusa. Assim como para produto, todos os modelos DSGE-VAR apresentaram expressiva melhora em MSE, em relação ao *benchmark*. Além disso, como observa-se na Tabela 5, o procedimento MCS indicou que para a variável câmbio, o modelo DSGE-VAR com $\lambda = 0.25$ e ordem de defasagem igual a 3 gerou previsões mais acuradas. Assim como para produto, em câmbio, o desempenho dos modelos BVAR com a priori de Minnesota mostra-se bastante semelhante ao desempenho dos modelos DSGE-VAR, onde ambos apresentam significativa melhora em relação ao BVAR com a priori difusa. Todavia, para câmbio, todos os modelos avaliados pertencem ao subconjunto $\hat{M}_{95\%}^*$ dado um nível de confiança de 95%, de acordo com o procedimento MCS.

Considerando a inflação, independente da ordem de defasagem escolhida, todos os modelos apresentaram desempenho superior ao *benchmark*. Mais ainda, o DSGE-VAR com $\lambda = 0.75$ e ordem de defasagem igual a 2, mostrou-se, de acordo com o MCS, o melhor modelo frente aos demais. Por outro lado, para os modelos BVAR com a priori de Minnesota, ainda que todos tenham apresentado expressivo ganho em MSE, em relação ao *benchmark*, apenas aqueles com ordem de defasagem igual a 1 e 3 pertencem ao subconjunto $\hat{M}_{95\%}^*$ com probabilidade de 95%.

Para a taxa de juros, sob um horizonte de previsão de 3 meses, o desempenho dos modelos DSGE-VAR mostrou-se expressivamente

superior ao BVAR com a priori difusa, sob todas as ordens de defasagem adotadas. Além disso, todos os modelos DSGE-VAR pertencem ao subconjunto $\hat{M}_{95\%}^*$ dado um nível de confiança de 95%. Apesar do fato que o procedimento MCS indicou que o BVAR com a priori de Minnesota com 3 defasagens é o modelo que gerou previsões mais acuradas, é importante notar que o desempenho dos DSGE-VAR, especialmente com 3 defasagens, mostrou-se bastante competitivo, com ganhos relativos em MSE muito similares ao BVAR com a priori de Minnesota.

Tabela 6: Ganhos (perdas) percentuais no MSE relativo ao BVAR com a priori difusa e \hat{p} -valores MCS para previsões 6 passos a frente

Modelo	Defasagem	MSE	\hat{p} -valor	MSE	\hat{p} -valor	MSE	\hat{p} -valor	MSE	\hat{p} -valor
		PIB		Câmbio		Inflação		Juros	
BVAR-difusa	1	1e-04	0,00	2e-03	0,04	2e-01	0,00	2e+00	0,00
DSGE-VAR, $\lambda = 0.25$	1	0,619	0,03	0,118	0,10*	0,207	0,08*	0,500	0,04
DSGE-VAR, $\lambda = 0.5$	1	0,613	0,02	0,133	0,63*	0,205	0,08*	0,499	0,04
DSGE-VAR, $\lambda = 0.75$	1	0,610	0,02	0,135	0,63*	0,201	0,08*	0,474	0,04
DSGE-VAR, $\lambda = 1$	1	0,606	0,02	0,137	0,67*	0,200	0,08*	0,467	0,04
DSGE-VAR, $\lambda = 1.5$	1	0,604	0,02	0,137	0,67*	0,198	0,08*	0,440	0,04
DSGE-VAR, $\lambda = 2$	1	0,602	0,02	0,136	0,63*	0,194	0,08*	0,420	0,04
BVAR-Minn	1	0,608	0,02	0,134	0,63*	0,188	0,02	0,333	0,02
BVAR-difusa	2	0,359	0,02	-0,401	0,04	0,069	0,00	0,334	0,00
DSGE-VAR, $\lambda = 0.25$	2	0,663	0,60*	0,119	0,10*	0,230	0,08*	0,625	0,05*
DSGE-VAR, $\lambda = 0.5$	2	0,694	1,00*	0,095	0,04	0,240	0,72*	0,660	0,87*
DSGE-VAR, $\lambda = 0.75$	2	0,652	0,60*	0,094	0,04	0,239	0,68*	0,664	1,00*
DSGE-VAR, $\lambda = 1$	2	0,631	0,02	0,115	0,04	0,228	0,08*	0,653	0,75*
DSGE-VAR, $\lambda = 1.5$	2	0,624	0,02	0,119	0,04	0,230	0,08*	0,652	0,86*
DSGE-VAR, $\lambda = 2$	2	0,612	0,02	0,124	0,10*	0,223	0,08*	0,606	0,45*
BVAR-Minn	2	0,622	0,02	0,107	0,04	0,201	0,00	0,585	0,04
BVAR-difusa	3	0,051	0,00	-0,135	0,04	0,049	0,00	-0,508	0,00
DSGE-VAR, $\lambda = 0.25$	3	0,687	0,80*	0,152	0,67*	0,233	0,08*	0,572	0,04
DSGE-VAR, $\lambda = 0.5$	3	0,639	0,03	0,179	1,00*	0,222	0,08*	0,587	0,05*
DSGE-VAR, $\lambda = 0.75$	3	0,658	0,60*	0,144	0,67*	0,177	0,02	0,591	0,05*
DSGE-VAR, $\lambda = 1$	3	0,640	0,03	0,138	0,63*	0,240	0,68*	0,631	0,65*
DSGE-VAR, $\lambda = 1.5$	3	0,638	0,02	0,131	0,26*	0,245	1,00*	0,621	0,05*
DSGE-VAR, $\lambda = 2$	3	0,629	0,02	0,130	0,10*	0,243	0,72*	0,658	0,87*
BVAR-Minn	3	0,630	0,02	0,110	0,04	0,208	0,08*	0,663	0,87*
DSGE		0,409	0,00	0,096	0,04	0,142	0,00	0,400	0,04

Nota: Entradas nas linhas BVAR-difusa com ordem de defasagem igual a 1 e colunas MSE correspondem ao MSE do modelo BVAR com a priori difusa. As entradas nas outras linhas correspondentes às colunas MSE são ganhos (valores positivos) ou perdas (valores negativos) percentuais em MSE dos 24 demais modelos relativos ao BVAR-difusa (1). As colunas \hat{p} -valor indicam o \hat{p} -valor MCS onde (*) indica que o modelo pertence à \hat{M}_{95}^* .

Sob um horizonte de previsão ainda mais longo, de 6 meses, a Tabela 6 explicita o bom desempenho dos modelos DSGE-VAR.

Para produto, de acordo com o procedimento MCS, o modelo DSGE-VAR com $\lambda = 0.5$ e 2 defasagens trouxe previsões mais precisas. Todos os modelos apresentaram melhoras percentuais em MSE relativo ao BVAR com a priori difusa. Especialmente no caso dos modelos DSGE-VAR e BVAR com a priori de Minnesota, a melhora na precisão das previsões, relativo ao *benchmark*, chega a ser superior a 60%. É importante perceber que o BVAR com a priori de Minnesota apresentou melhoras percentuais em MSE bastante semelhantes ao DSGE-VAR, no entanto, apenas modelos DSGE-VAR pertencem ao subconjunto $\hat{M}_{95\%}^*$ com probabilidade de 95%, sugerindo o benefício da inclusão de informações a priori que incorporam teoria econômica.

Os resultados da Tabela 6 para câmbio ilustram o bom desempenho dos modelos DSGE-VAR para previsões de longo prazo. Isso porque, ao se observar a Tabela 4, nota-se que no curto prazo os desempenhos dos modelos DSGE-VAR mostram-se relativamente inferiores ao *benchmark*, por outro lado, quando observa-se a Tabela 6 é possível perceber ganhos percentuais em MSE. Ademais, de acordo com o \hat{p} -valor MCS, DSGE-VAR(3) com $\lambda = 0.5$ foi considerado o modelo que gerou previsões mais precisas. Assim como para produto, em câmbio o desempenho do BVAR com a priori de Minnesota é bastante próximo do DSGE-VAR, ainda que o último tenha apresentado resultados mais precisos. Note que apenas BVAR-Minn(1) pertence ao subconjunto $\hat{M}_{95\%}^*$ dado nível de confiança de 95%.

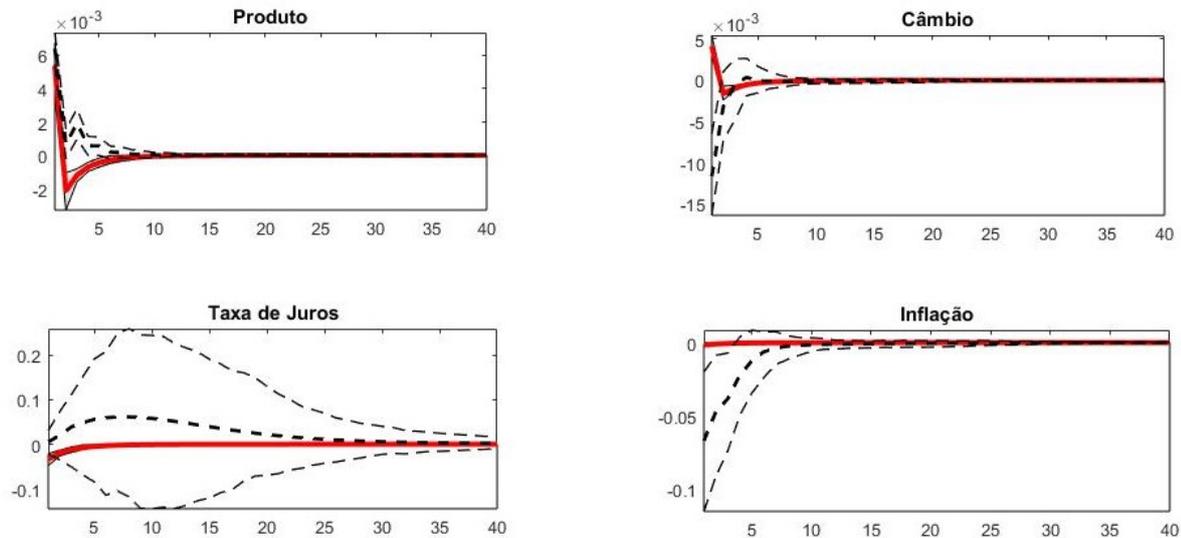
Para inflação, assim como as demais variáveis supracitadas, o modelo DSGE-VAR superou os demais no sentido que gerou previsões mais precisas. Segundo procedimento MCS, DSGE-VAR(3) com $\lambda = 1.5$ refere-se ao modelo com melhor desempenho. É importante notar que no caso da inflação, ainda que BVAR com a priori de Minnesota tenha apresentado ganhos percentuais significativos em relação ao *benchmark*, apenas o BVAR-Minn(3) pertence ao subconjunto $\hat{M}_{95\%}^*$ dado nível de confiança de 95%.

Por fim, na Tabela 6 observa-se que para taxa de juros há também uma significativa melhora percentual em MSE por parte de todos os modelos DSGE-VAR em relação ao *benchmark*. Especialmente no caso dessa variável, é importante notar que todos os modelos DSGE-VAR com ordens de defasagem iguais a 2 e 3 pertencem ao subconjunto $\hat{M}_{95\%}^*$ dado nível de confiança de 95%, enquanto que nenhum DSGE-VAR(1) pertence ao mesmo subconjunto. Ademais, de acordo com o \hat{p} -valor MCS, o DSGE-VAR(2) com $\lambda = 0.75$ mostrou-se o modelo com

melhor desempenho dentre os modelos avaliados. Além dos modelos DSGE-VAR supracitados, apenas o BVAR(3) com a priori de Minnesota pertence ao subconjunto $\hat{M}_{95\%}^*$ dado nível de confiança de 95%, sugerindo novamente que, para previsões de longo prazo, a incorporação de teoria através das informações a priori mostra-se benéfica quando o objetivo é gerar previsões.

9.4 FUNÇÕES IMPULSO-RESPOSTA

Com a estimação dos parâmetros dos modelos DSGE-VAR e DSGE em mãos, as funções impulso-resposta podem ser computadas. As Figuras 1, 2, 3 e 4 representam as respectivas funções impulso-resposta com relação aos choques tecnológico, monetário, de preferências e no produto do resto do mundo das variáveis de interesse: produto, inflação, câmbio e taxa de juros. Cada gráfico exibe a média das funções impulso-resposta do modelo DSGE-VAR com $\hat{\lambda}$, representado pelas linhas pontilhadas em preto, com espessura maior, com as linhas pontilhadas de menor espessura indicando os *highest posterior density intervals* (HPDI) de 90%. Por sua vez, as linhas contínuas, em vermelho, mais espessas, representam as médias das funções impulso-resposta do modelo DSGE, e as áreas em cinza representam os HPDI de 90%. Todas as estimativas são baseadas na amostra de Jan/03 até Dez/16.

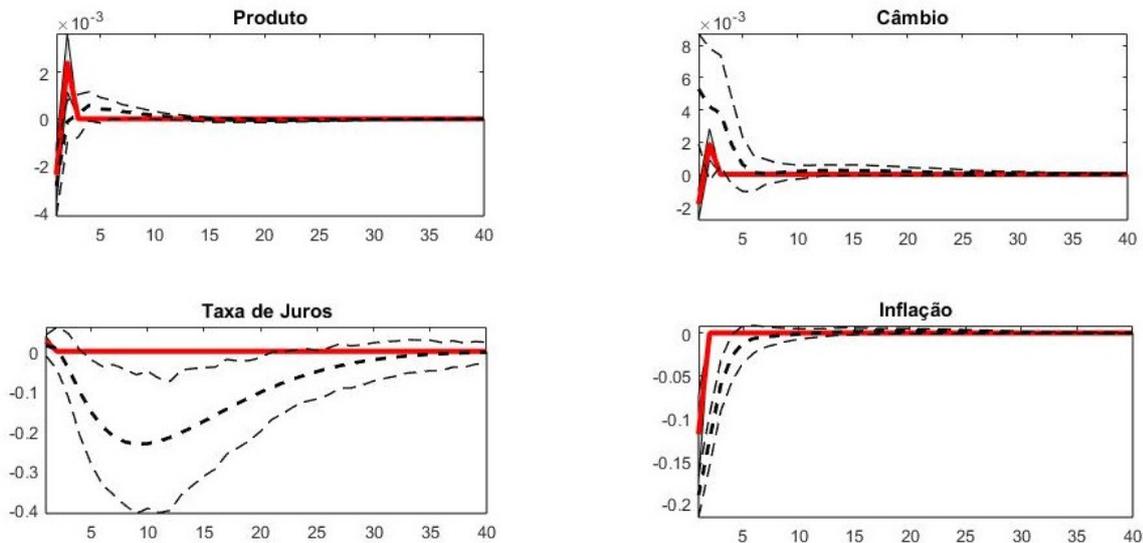


Nota: As linhas pontilhadas em preto, com espessura maior, representam as médias das funções impulso-resposta do modelo DSGE-VAR, onde as linhas pontilhadas de menor espessura representam os *highest posterior density intervals* (HPDI) de 90%. Já as linhas contínuas, em vermelho, mais espessas, representam as médias das funções impulso-resposta do modelo DSGE, e as áreas em cinza representam os HPDI de 90%. As funções impulso-resposta são baseadas na amostra de Jan/03 - Dez/16.

Figura 1: Funções Impulso-Resposta para um Choque Tecnológico

A Figura 1 mostra as funções impulso-resposta do modelo DSGE (linhas vermelhas) e as funções impulso resposta do modelo DSGE-VAR (linhas pontilhadas) com relação à um choque tecnológico para produto, câmbio, taxa de juros e inflação. Considerando primeiro as funções impulso-resposta obtidas da estimação do modelo DSGE, observa-se que um choque tecnológico gera, inicialmente, um crescimento no produto, seguido por um declínio ainda nos primeiros períodos. Além disso, percebe-se uma depreciação da moeda corrente e aumento da taxa de câmbio. A taxa de juros apresenta leve redução e a inflação parece permanecer estável.

Da mesma maneira, avaliando agora as funções impulso-resposta de um choque tecnológico obtido da estimação do modelo DSGE-VAR com $\hat{\lambda}$, observa-se aumento do produto doméstico sob um menor custo marginal, gerando pressão deflacionária. Em relação à taxa de juros, um resultado esperado seria uma reação da autoridade monetária reduzindo as taxas de juros com intuito de aliviar a pressão deflacionária. No entanto, o que se observa na Figura 1 é o oposto. Porém, deve-se levar em conta que esse resultado contém um alto grau de incerteza, dado que o HPDI mostra-se bastante amplo. Todavia, frente ao aumento nas taxas de juros, ocorreu apreciação da moeda doméstica e redução da taxa de câmbio.

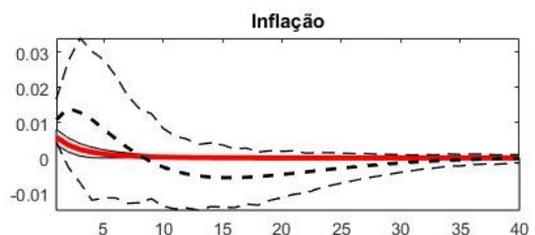
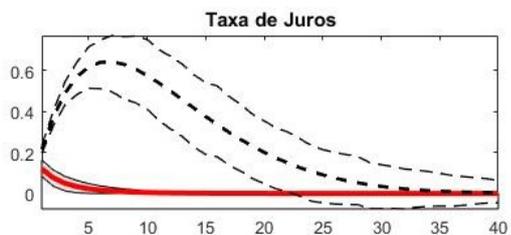
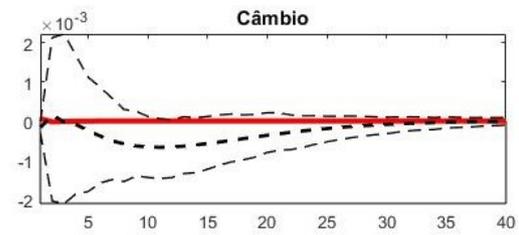
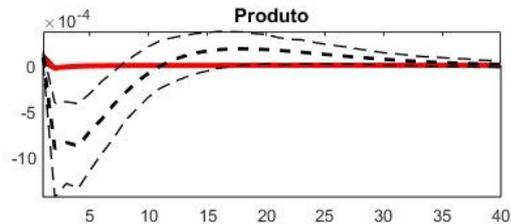


Nota: As linhas pontilhadas em preto, com espessura maior, representam as médias das funções impulso-resposta do modelo DSGE-VAR, onde as linhas pontilhadas de menor espessura representam os *highest posterior density intervals* (HPDI) de 90%. Já as linhas contínuas, em vermelho, mais espessas, representam as médias das funções impulso-resposta do modelo DSGE, e as áreas em cinza representam os HPDI de 90%. As funções impulso-resposta são baseadas na amostra de Jan/03 - Dez/16.

Figura 2: Funções Impulso-Resposta para um Choque Monetário

Similarmente, a Figura 2 exibe as funções impulso-resposta de um choque monetário exógeno sobre produto, taxa de câmbio, taxa de juros e inflação. Considerando primeiramente as funções impulso-resposta obtidas da estimação do modelo DSGE, observa-se que esse leva a um aumento da taxa de juros, gerando um declínio no produto e na inflação e apreciação da moeda. Sob rigidez de preços, a apreciação da moeda vai implicar em uma melhora nos termos de troca, que por sua vez, levam a queda da taxa real de câmbio. O choque não mostra-se muito persistente, uma vez que todas as variáveis voltam ao estado inicial após cinco períodos.

No caso do modelo DSGE-VAR a dinâmica que ocorre é a mesma, entretanto, um choque monetário exógeno leva a uma queda mais acentuada na taxa de juros, fazendo com que as demais variáveis mostrem-se mais sensíveis ao choque. Além disso, observa-se maior persistência do choque monetário, visto que o efeito do mesmo leva cerca de 10 períodos para dissipar-se.

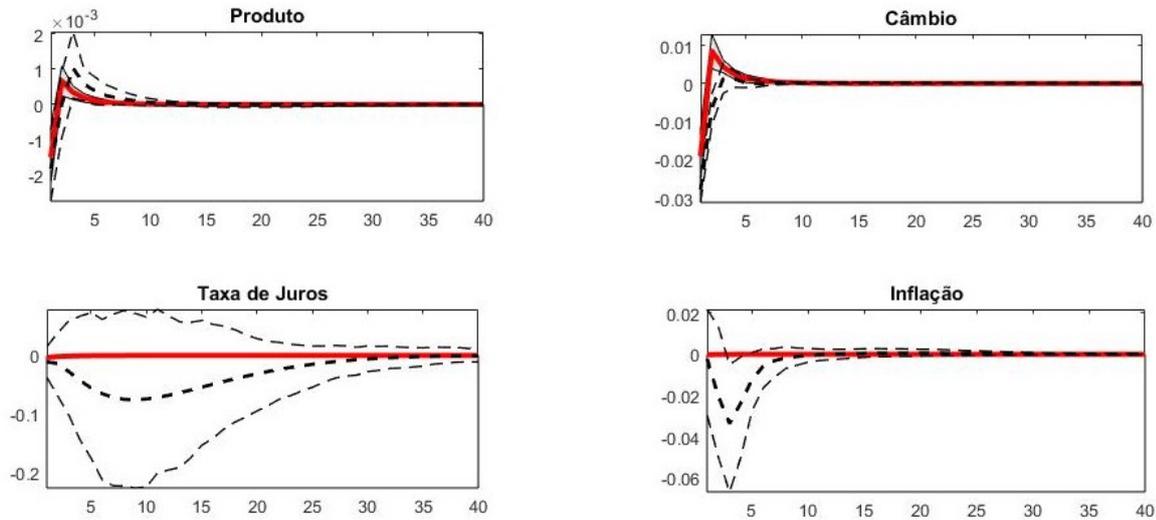


Nota: As linhas pontilhadas em preto, com espessura maior, representam as médias das funções impulso-resposta do modelo DSGE-VAR, onde as linhas pontilhadas de menor espessura representam os *highest posterior density intervals* (HPDI) de 90%. Já as linhas contínuas, em vermelho, mais espessas, representam as médias das funções impulso-resposta do modelo DSGE, e as áreas em cinza representam os HPDI de 90%. As funções impulso-resposta são baseadas na amostra de Jan/03 - Dez/16.

Figura 3: Funções Impulso-Resposta para um Choque de Preferências

A Figura 3 exibe as funções impulso-resposta para um choque de preferências. Observando-se primeiramente os resultados obtidos da estimação do modelo DSGE, nota-se um pequeno aumento no produto e aumento da demanda. Essa pressão na demanda gera aumento do custo marginal, que contribui para o aumento inflacionário e na taxa de juros. A taxa de câmbio, por sua vez, permanece estável.

Por outro lado, considerando o modelo DSGE-VAR, a dinâmica ocorre de maneira semelhante ao DSGE, porém de forma mais intensa. Frente à um choque de preferências, o produto eleva-se ligeiramente, gerando aumento inflacionário e nas taxas de juros. A pressão inflacionária, por suas vez, reduz o produto. Após cerca de dez períodos, ocorre deflação juntamente com declínio da taxa de juros, elevando o produto novamente. O choque dissipa-se após mais de 30 períodos, indicando persistência.



Nota: As linhas pontilhadas em preto, com espessura maior, representam as médias das funções impulso-resposta do modelo DSGE-VAR, onde as linhas pontilhadas de menor espessura representam os *highest posterior density intervals* (HPDI) de 90%. Já as linhas contínuas, em vermelho, mais espessas, representam as médias das funções impulso-resposta do modelo DSGE, e as áreas em cinza representam os HPDI de 90%. As funções impulso-resposta são baseadas na amostra de Jan/03 - Dez/16.

Figura 4: Funções Impulso-Resposta para um Choque no Produto do Resto do Mundo

Por fim, a Figura 4 apresenta as funções impulso-resposta para um choque no produto do resto do mundo. Considerando os resultados obtidos da estimação do DSGE, observa-se que um aumento no produto do resto do mundo gera ligeira queda no produto doméstico, que leva a uma queda na taxa de câmbio. O efeito do choque dissipa-se após cerca de 10 períodos.

Da mesma forma, ao avaliar-se as funções impulso-resposta obtidas da estimação do modelo DSGE-VAR com $\hat{\lambda}$, é possível perceber que esse captura de forma mais sensível os efeitos de curto prazo. O aumento no produto do resto do mundo gera declínio no produto doméstico, deflação e queda na taxa de juros. A redução dos preços domésticos frente aos preços estrangeiros gera uma apreciação da moeda doméstica e queda na taxa de câmbio. O efeito dissipa-se para produto, inflação e taxa de câmbio após cerca de 10 períodos. Para taxa de juros, o efeito mostra-se mais persistente, dissipando-se após cerca de 30 períodos.

10 CONCLUSÃO

Este trabalho procurou mostrar o benefício da utilização de modelos DSGE-VAR para gerar previsões, seguindo a abordagem proposta por Del Negro e Schorfheide (2004). O modelo teórico é baseado no DSGE de pequena economia aberta desenvolvido por Gali e Monacelli (2005). A base de dados é composta por variáveis macroeconômicas, de periodicidade mensal, da economia brasileira, como produto, inflação, taxa de câmbio e taxa de juros, de Jan/03 a Dez/16. A comparação do desempenho dos modelos DSGE-VAR foi realizada, através do procedimento MCS, em relação ao desempenho de um BVAR com a priori difusa, um BVAR com a conhecida priori de Minnesota e o modelo DSGE estimado independentemente. O desempenho dos modelos DSGE-VAR mostrou-se superior aos modelos BVAR com a priori difusa e DSGE. Já em relação ao modelo BVAR com a priori de Minnesota, DSGE-VAR mostrou-se competitivo e, em algumas situações, superior. Em resumo, informações a priori de modelos DSGE provaram ser úteis no sentido de melhorar o desempenho do VAR com a finalidade de gerar previsões, principalmente para previsões de médio e longo prazo. Apesar do modelo BVAR com a priori de Minnesota ter mostrado-se bastante competitivo e, em alguns momentos, superado o DSGE-VAR, deve-se levar em conta que o último permite reter a parte da estrutura econômica do modelo DSGE, podendo ser utilizado em situações onde modelos puramente estatísticos não são suficientes.

REFERÊNCIAS

- BLANCHARD, O. J.; KAHN, C. M. The solution of linear difference models under rational expectations. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 1305–1311, 1980.
- CALVO, G. A. Staggered prices in a utility-maximizing framework. **Journal of monetary Economics**, Elsevier, v. 12, n. 3, p. 383–398, 1983.
- CHIB, S.; GREENBERG, E. Understanding the metropolis-hastings algorithm. **The american statistician**, Taylor & Francis Group, v. 49, n. 4, p. 327–335, 1995.
- CHRISTIANO, L. J.; EICHENBAUM, M.; EVANS, C. L. Nominal rigidities and the dynamic effects of a shock to monetary policy. **Journal of political Economy**, The University of Chicago Press, v. 113, n. 1, p. 1–45, 2005.
- COSTA, H. C.; PORTO JUNIOR, S. d. S.; ALMEIDA, H. J. F. Dinâmica da inflação para uma economia sob metas de inflação: Evidências para o brasil usando modelos dsge-var. In: **45 Encontro Nacional de Economia**. [S.l.: s.n.], 2017.
- DE VROEY, M. Friedman and lucas on the phillips curve: from a disequilibrium to an equilibrium approach. **Eastern Economic Journal**, JSTOR, v. 27, n. 2, p. 127–148, 2001.
- DE VROEY, M.; MALGRANGE, P. The history of macroeconomics from keynes's general theory to the present. **IRES Discussion Papers**, v. 2011028, 2011.
- DEJONG, D.; INGRAM, B. F.; WHITEMAN, C. Analyzing vars with monetary business cycle model priors. In: **Proceedings of the American Statistical Association, Bayesian Statistics Section**. [S.l.: s.n.], 1993. v. 160, p. 69.
- DEL NEGRO, M.; SCHORFHEIDE, F. Priors from general equilibrium models for vars. **International Economic Review**, Wiley Online Library, v. 45, n. 2, p. 643–673, 2004.
- DEL NEGRO, M.; SCHORFHEIDE, F. How good is what you've got? dsge-var as a toolkit for evaluating dsge models. **Economic**

Review-Federal Reserve Bank of Atlanta, Federal Reserve Bank of Atlanta, v. 91, n. 2, p. 21, 2006.

DEL NEGRO, M.; SCHORFHEIDE, F. Monetary policy analysis with potentially misspecified models. **The American Economic Review**, American Economic Association, v. 99, n. 4, p. 1415–1450, 2009.

DIXIT, A. K.; STIGLITZ, J. E. Monopolistic competition and optimum product diversity. **The American Economic Review**, JSTOR, v. 67, n. 3, p. 297–308, 1977.

DOAN, T.; LITTERMAN, R.; SIMS, C. Forecasting and conditional projection using realistic prior distributions. **Econometric reviews**, Taylor & Francis, v. 3, n. 1, p. 1–100, 1984.

FERNANDEZ-VILLAVERDE, J.; RUBIO-RAMIREZ, J.; SARGENT, T. J. **A, B, C's (and D)'s for Understanding VARs**. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 2005.

FRIEDMAN, M. The role of monetary policy. **The American economic review**, American Economic Association, v. 58, n. 1, p. 1–17, 1968.

FRIEDMAN, M.; SCHWARTZ, A. J. **A monetary history of the United States, 1867-1960**. [S.l.]: Princeton University Press, 1963.

FRISCH, R. Propagation problems and impulse problems in dynamic economics. G. Allen & Unwin, 1933.

GALI, J.; MONACELLI, T. Monetary policy and exchange rate volatility in a small open economy. **The Review of Economic Studies**, Wiley-Blackwell, v. 72, n. 3, p. 707–734, 2005.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: **IN BAYESIAN STATISTICS**. [S.l.]: University Press, 1992. p. 169–193.

GEWEKE, J. Using simulation methods for bayesian econometric models: inference, development, and communication. **Econometric reviews**, Taylor & Francis, v. 18, n. 1, p. 1–73, 1999.

GIACOMINI, R. The relationship between dsge and var models. In: **VAR Models in Macroeconomics—New Developments and**

Applications: Essays in Honor of Christopher A. Sims. [S.l.]: Emerald Group Publishing Limited, 2013. p. 1–25.

GOODFRIEND, M.; KING, R. G. The new neoclassical synthesis and the role of monetary policy. **NBER macroeconomics annual**, MIT Press, v. 12, p. 231–283, 1997.

GUPTA, R.; STEINBACH, R. A dsge-var model for forecasting key south african macroeconomic variables. **Economic Modelling**, Elsevier, v. 33, p. 19–33, 2013.

HANSEN, P. R.; LUNDE, A.; NASON, J. M. The model confidence set. **Econometrica**, Wiley Online Library, v. 79, n. 2, p. 453–497, 2011.

HERBST, E. P.; SCHORFHEIDE, F. **Bayesian Estimation of DSGE Models.** [S.l.]: Princeton University Press, 2015.

INGRAM, B. F.; WHITEMAN, C. H. Supplanting the minnesota prior: Forecasting macroeconomic time series using real business cycle model priors. **Journal of Monetary Economics**, Elsevier, v. 34, n. 3, p. 497–510, 1994.

KADIYALA, K. R.; KARLSSON, S. Forecasting with generalized bayesian vector auto regressions. **Journal of Forecasting**, Wiley Online Library, v. 12, n. 3-4, p. 365–378, 1993.

KLEIN, P. Using the generalized schur form to solve a multivariate linear rational expectations model. **Journal of Economic Dynamics and Control**, Elsevier, v. 24, n. 10, p. 1405–1423, 2000.

KOOP, G. **Bayesian Econometrics.** [S.l.]: John Wiley, Chichester, 2003.

KOOP, G.; KOROBILIS, D. Bayesian multivariate time series methods for empirical macroeconomics. **Foundations and Trends® in Econometrics**, Now Publishers, Inc., v. 3, n. 4, p. 267–358, 2010.

KYDLAND, F. E.; PRESCOTT, E. C. Time to build and aggregate fluctuations. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 1345–1370, 1982.

LEES, K.; MATHESON, T.; SMITH, C. Open economy forecasting with a dsge-var: Head to head with the rbzn published forecasts. **International Journal of Forecasting**, Elsevier, v. 27, n. 2, p. 512–528, 2011.

LINARDI, F. **Assessing the Fit of a Small Open-economy DSGE Model for the Brazilian Economy.** [S.l.], 2016.

LITTERMAN, R. B. **Techniques of forecasting using vector autoregressions.** [S.l.], 1979.

LITTERMAN, R. B. Forecasting with bayesian vector autoregressions: five years of experience. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 4, n. 1, p. 25–38, 1986.

LONG, J. B.; PLOSSER, C. I. Real business cycles. **Journal of political Economy**, The University of Chicago Press, v. 91, n. 1, p. 39–69, 1983.

LUCAS, R. E. Expectations and the neutrality of money. **Journal of economic theory**, Elsevier, v. 4, n. 2, p. 103–124, 1972.

LUCAS, R. E. Econometric policy evaluation: A critique. In: ELSEVIER. **Carnegie-Rochester conference series on public policy.** [S.l.], 1976. v. 1, p. 19–46.

LUCAS, R. E. Understanding business cycles. In: ELSEVIER. **Carnegie-Rochester conference series on public policy.** [S.l.], 1977. v. 5, p. 7–29.

LUCAS, R. E. **Business Cycle Theory.** [S.l.]: Cambridge, MIT Press, 1981.

LUCAS, R. E.; SARGENT, T. After keynesian macroeconomics. **Rational expectations and econometric practice**, University of Minnesota Press Minneapolis, v. 1, p. 295–319, 1981.

MUTH, J. F. Rational expectations and the theory of price movements. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 315–335, 1961.

NELSON, C. R.; PLOSSER, C. R. Trends and random walks in macroeconomic time series: some evidence and implications. **Journal of monetary economics**, Elsevier, v. 10, n. 2, p. 139–162, 1982.

SCHORFHEIDE, F. Loss function-based evaluation of dsge models. **Journal of Applied Econometrics**, Wiley Online Library, v. 15, n. 6, p. 645–670, 2000.

SILVEIRA, M. A. C. d. Using a bayesian approach to estimate and compare new keynesian dsge models for the brazilian economy: the role for endogenous persistence. **Revista brasileira de Economia**, SciELO Brasil, v. 62, n. 3, p. 333–357, 2008.

SIMS, C. A. Macroeconomics and reality. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 1–48, 1980.

SIMS, C. A. Solving linear rational expectations models. **Computational economics**, Springer, v. 20, n. 1, p. 1–20, 2002.

SMETS, F.; WOUTERS, R. An estimated dynamic stochastic general equilibrium model of the euro area. **Journal of the European economic association**, Wiley Online Library, v. 1, n. 5, p. 1123–1175, 2003.

SNOWDON, B.; VANE, H. R. **Modern macroeconomics: its origins, development and current state**. [S.l.]: Edward Elgar Publishing, 2005.

TAYLOR, J. B. Discretion versus policy rules in practice. In: ELSEVIER. **Carnegie-Rochester conference series on public policy**. [S.l.], 1993. v. 39, p. 195–214.

THEIL, H.; GOLDBERGER, A. S. On pure and mixed statistical estimation in economics. **International Economic Review**, JSTOR, v. 2, n. 1, p. 65–78, 1961.

THOMPSON, P. A.; MILLER, R. B. Sampling the future: a bayesian approach to forecasting from univariate time series models. **Journal of Business & Economic Statistics**, Taylor & Francis Group, v. 4, n. 4, p. 427–436, 1986.

WATANABE, T. **The application of DSGE-VAR model to macroeconomic data in Japan**. [S.l.], 2007.

WOODFORD, M. Revolution and evolution in twentieth-century macroeconomics. In: **Library of Congress, Washington, June**. [S.l.: s.n.], 1999. p. 203–210.

APÊNDICE A - DEMONSTRAÇÕES

CONSTANTE DE INTEGRAÇÃO

A normalização do fator $\mathbf{c}(\boldsymbol{\theta})$ para a densidade a priori condicional dos parâmetros do VAR $\mathbf{p}(\boldsymbol{\phi}, \boldsymbol{\Sigma}_u | \boldsymbol{\theta})$ na Equação (7.24) é dado por

$$\begin{aligned} \mathbf{c}(\boldsymbol{\theta}) &= (2\pi)^{-\frac{nk}{2}} |\boldsymbol{\lambda T \Gamma_{xx}^*}|^{\frac{n}{2}} \\ &\times 2^{-\frac{n(\lambda T - k)}{2}} \pi^{-\frac{n(n-1)}{4}} |\boldsymbol{\lambda T \Sigma_u^*}|^{\frac{\lambda T - k}{2}} \\ &\times \prod_{i=1}^n \Gamma[(\lambda T - k + 1 - i)/2]^{-1} \end{aligned} \quad (\text{A.1})$$

onde $\Gamma[\cdot]$ denota a função gama.

VEROSSIMILHANÇA MARGINAL DE $\boldsymbol{\theta}$

A densidade marginal dos dados, dada pela Equação (7.32) pode ser escrita como

$$p_{\lambda}(Y) = \int p(Y | \boldsymbol{\phi}, \boldsymbol{\Sigma}_u, \boldsymbol{\theta}) p(\boldsymbol{\phi}, \boldsymbol{\Sigma}_u | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d(\boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \boldsymbol{\phi}), \quad (\text{A.2})$$

porém, note que a a priori dos parâmetros do VAR dado $\boldsymbol{\theta}$ tem forma Normal-Wishart-Invertida, então a função de verossimilhança marginal de $\boldsymbol{\theta}$ pode ser calculada analiticamente

$$\begin{aligned} &\int p(Y | \boldsymbol{\phi}, \boldsymbol{\Sigma}_u) p(\boldsymbol{\phi}, \boldsymbol{\Sigma}_u | \boldsymbol{\theta}) d(\boldsymbol{\phi}, \boldsymbol{\Sigma}_u) \\ &= \frac{p(Y | \boldsymbol{\phi}, \boldsymbol{\Sigma}_u) p(\boldsymbol{\phi}, \boldsymbol{\Sigma}_u | \boldsymbol{\theta})}{p(\boldsymbol{\phi}, \boldsymbol{\Sigma}_u | Y)} \end{aligned} \quad (\text{A.3})$$

O segundo termo da igualdade pode ser obtido a partir das constantes de normalização das distribuições Normal-Wishart-Invertida, en-

tão

$$\begin{aligned}
p(Y|\theta) &= \frac{p(Y|\phi, \Sigma_u)p(\phi, \Sigma_u|\theta)}{p(\phi, \Sigma_u|Y)} \\
&= \frac{|\lambda T \Gamma_{xx}^*(\theta) + X'X|^{-\frac{n}{2}} |(\lambda + 1)T \tilde{\Sigma}_u(\theta)|^{-\frac{(\lambda+1)T-k}{2}}}{|\lambda T \Gamma_{xx}^*(\theta)|^{-\frac{n}{2}} |\lambda T \Sigma_u^*(\theta)|^{-\frac{\lambda T-k}{2}}} \\
&\quad \times \frac{(2\pi)^{-nT/2} 2^{\frac{n((\lambda+1)T-k)}{2}}}{2^{\frac{n(\lambda T-k)}{2}}} \\
&\quad \times \frac{\prod_{i=1}^n \Gamma[(\lambda + 1)T - k + 1 - i]/2!}{\prod_{i=1}^n \Gamma[(\lambda T - k + 1 - i)/2]}. \tag{A.4}
\end{aligned}$$

DENSIDADE A POSTERIORI

A densidade a posteriori é proporcional ao produto da verossimilhança e da priori. Utilizando as expressões da verossimilhança em (7.12) e da priori em (7.24), tem-se:

$$\begin{aligned}
p(\phi, \Sigma_u|Y, \theta) &\propto |\Sigma_u|^{-\frac{T}{2}} \\
&\quad \times \exp\left\{-\frac{1}{2}Tr[S\Sigma_u^{-1}]\right\} \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(\phi - \hat{\phi})'X'X(\phi - \hat{\phi})]\right\} \\
&\quad \times |\Sigma_u|^{-\frac{\lambda T+n+1}{2}} \exp\left\{-\frac{1}{2}Tr[\lambda T \Sigma_u^* \Sigma_u^{-1}]\right\} \\
&\quad \times \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(\phi - \phi^*)'\lambda T \Gamma_{xx}^*(\phi - \phi^*)]\right\}. \tag{A.5}
\end{aligned}$$

Trabalhando com os termos, tem-se

$$\begin{aligned}
p(\phi, \Sigma_u|Y, \theta) &\propto |\Sigma_u|^{-\frac{(\lambda+1)T+n+1}{2}} \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1} \right. \\
&\quad \times \underbrace{S + \lambda T \Sigma_u^* + (\phi - \hat{\phi})'X'X(\phi - \hat{\phi})}_{\kappa} \\
&\quad \left. + \underbrace{(\phi - \phi^*)'\lambda T \Gamma_{xx}^*(\phi - \phi^*)}_{\kappa}\right\}. \tag{A.6}
\end{aligned}$$

Resolvendo apenas para o termo κ :

$$\begin{aligned}
 \kappa &= S + \lambda T \Sigma_u^* + \phi' X' X \phi - 2\phi' X' X \hat{\phi} + \hat{\phi}' X' X \hat{\phi} \\
 &\quad + \phi' \lambda T \Gamma_{xx}^* \phi - 2\phi' \lambda T \Gamma_{xx}^* \phi^* + \phi^* \lambda T \Gamma_{xx}^* \phi^* \\
 &= S + \lambda T \Sigma_u^* + \hat{\phi}' X' X \hat{\phi} + \phi^* \lambda T \Gamma_{xx}^* \phi^* \\
 &\quad + \phi(X' X + \lambda T \Gamma_{xx}^*) \phi - 2\phi'(X' X \hat{\phi} + \lambda T \Gamma_{xx}^* \phi^*).
 \end{aligned}$$

Então, definindo $V^{-1} = (X' X + \lambda T \Gamma_{xx}^*)$ e $\tilde{\phi} = (\lambda T \Gamma_{xx}^* + X' X)^{-1}(X' X \hat{\phi} + \lambda T \Gamma_{xx}^* \phi^*)$, obtém-se

$$\begin{aligned}
 \kappa &= S + \lambda T \Sigma_u^* + \hat{\phi}' X' X \hat{\phi} + \phi^* \lambda T \Gamma_{xx}^* \phi^* \\
 &\quad + \phi V^{-1} \phi - 2\phi' V^{-1} \tilde{\phi}.
 \end{aligned} \tag{A.7}$$

Somando e subtraíndo $\tilde{\phi} V^{-1} \tilde{\phi}$ em (A.7), tem-se

$$\begin{aligned}
 \kappa &= S + \lambda T \Sigma_u^* + \hat{\phi}' X' X \hat{\phi} + \phi^* \lambda T \Gamma_{xx}^* \phi^* \\
 &\quad + \phi V^{-1} \phi - 2\phi' V^{-1} \tilde{\phi} + \tilde{\phi} V^{-1} \tilde{\phi} - \tilde{\phi} V^{-1} \tilde{\phi} \\
 &= \underbrace{S + \lambda T \Sigma_u^* + \hat{\phi}' X' X \hat{\phi} + \phi^* \lambda T \Gamma_{xx}^* \phi^* - \tilde{\phi} V^{-1} \tilde{\phi}}_{\kappa \kappa} \\
 &\quad + (\phi - \tilde{\phi})' V^{-1} (\phi - \tilde{\phi}).
 \end{aligned}$$

Trabalhando apenas com o termo $\kappa \kappa$, lembrando das definições de $S = Y' Y - \hat{\phi}' X' X \hat{\phi}$ e de $\lambda T \Sigma_u^* = \lambda T \Gamma_{yy}^* - \phi^* \lambda T \Gamma_{xx}^* \phi^*$, então

$$\kappa \kappa = (Y' Y + \lambda T \Gamma_{yy}^*) - \tilde{\phi} V^{-1} \tilde{\phi}.$$

Utilizando as definições de V^{-1} e $\tilde{\phi}$

$$\begin{aligned}
 \kappa \kappa &= (Y' Y + \lambda T \Gamma_{yy}^*) \\
 &\quad - (\lambda T \Gamma_{yx}^* + Y' X)(\lambda T \Gamma_{xx}^* + X' X)^{-1}(\lambda T \Gamma_{xy}^* + X' Y) \\
 &= (1 + \lambda) T \tilde{\Sigma}_u.
 \end{aligned}$$

Portanto,

$$\kappa = (1 + \lambda) T \tilde{\Sigma}_u + (\phi - \tilde{\phi})' V^{-1} (\phi - \tilde{\phi}). \tag{A.8}$$

Substituindo (A.8) em (A.6), tem-se:

$$\begin{aligned}
 p(\phi, \Sigma_u | Y, \theta) &= \chi |\Sigma_u|^{-\frac{(1+\lambda)T+n+1}{2}} \\
 &\quad \times \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(1+\lambda)T\tilde{\Sigma}_u]\right\} \\
 &\quad \times \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}\right. \\
 &\quad \left. \times (\phi - \tilde{\phi})'(\lambda T\Gamma_{xx}^* + X'X)(\phi - \tilde{\phi})]\right\}. \quad (A.9)
 \end{aligned}$$

Portanto, a posteriori é proporcional ao produto de uma densidade Normal para $\phi|\Sigma_u$ vezes uma densidade Wishart-Invertida para Σ_u , condicional a θ

$$\Sigma_u | Y, \theta \sim Wi((\lambda + 1)T\tilde{\Sigma}_u, (1 + \lambda)T - k, n) \quad (A.10)$$

$$\phi | \Sigma_u, Y, \theta \sim N(\tilde{\phi}, \Sigma_u \otimes (\lambda T\Gamma_{xx}^* + X'X)^{-1}). \quad (A.11)$$

Onde χ é a constante de integração, dada por

$$\begin{aligned}
 \chi &= (2\pi)^{-\frac{nk}{2}} |\lambda T\Gamma_{xx}^* + X'X|^{\frac{n}{2}} |(\lambda + 1)T\tilde{\Sigma}_u|^{\frac{(1+\lambda)T-k}{2}} \\
 &\quad \times 2^{-\frac{n((1+\lambda)T-k)}{2}} \pi^{-\frac{n(n-1)}{4}} \prod_{i=1}^n \Gamma[(\lambda + 1)T - k + 1 - i]/2]^{-1}. \quad (A.12)
 \end{aligned}$$

DENSIDADE PREDITIVA

Para realizar previsões um passo a frente, ou seja $h = 1$, a verossimilhança dos novos dados pode ser escrita como:

$$\begin{aligned}
 p(Y_{T+h} | \phi, \Sigma_u, \theta) &= (2\pi)^{-\frac{nh}{2}} |\Sigma_u|^{-\frac{h}{2}} \\
 &\quad \times \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(Y_{T+h} - \bar{X}\phi)'(Y_{T+h} - \bar{X}\phi)]\right\}. \quad (A.13)
 \end{aligned}$$

A densidade preditiva é, então, resultado da multiplicação entre

a verossimilhança dos novos dados e da densidade a posteriori (7.29),

$$\begin{aligned}
 p(\mathbf{Y}_{T+h}|\mathbf{Y}) &= \int \int \int |\Sigma_u|^{-\frac{(\lambda+1)T+n+h+1}{2}} \chi \\
 &\quad \times \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(\lambda+1)T\tilde{\Sigma}_u]\right\} \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}\right. \\
 &\quad \times \underbrace{(\phi - \tilde{\phi})'(\lambda T\Gamma_{xx}^* + \mathbf{X}'\mathbf{X})(\phi - \tilde{\phi})}_{\Upsilon} \\
 &\quad \left. + \underbrace{(\mathbf{Y}_{T+h} - \bar{\mathbf{X}}\phi)'(\mathbf{Y}_{T+h} - \bar{\mathbf{X}}\phi)}_{\Upsilon}\right\} d\phi d\Sigma_u d\theta. \quad (\text{A.14})
 \end{aligned}$$

onde χ é a constante de integração, definida como

$$\begin{aligned}
 \chi &= (2\pi)^{-\frac{nh}{2}} (2\pi)^{-\frac{nk}{2}} |\lambda T\Gamma_{xx}^* + \mathbf{X}'\mathbf{X}|^{\frac{n}{2}} 2^{-\frac{n((\lambda+1)T-k)}{2}} \pi^{-\frac{n(n-1)}{4}} \\
 &\quad \times |(\lambda+1)T\tilde{\Sigma}_u|^{\frac{(1+\lambda)T-k}{2}} \prod_{i=1}^n \Gamma\left[\frac{(\lambda+1)T-k+1-i}{2}\right]^{-1}
 \end{aligned} \quad (\text{A.15})$$

Trabalhando apenas com o termo agora denominado Υ e definindo $\mathbf{V}^{-1} = (\lambda T\Gamma_{xx}^* + \mathbf{X}'\mathbf{X})$:

$$\begin{aligned}
 \Upsilon &= \phi' \underbrace{(\mathbf{V}^{-1} + \bar{\mathbf{X}}'\bar{\mathbf{X}})}_A \phi - \phi' \underbrace{(\bar{\mathbf{X}}\mathbf{Y}_{T+h} + \mathbf{V}^{-1}\tilde{\phi})}_b \\
 &\quad - \underbrace{(\mathbf{Y}'_{T+h}\bar{\mathbf{X}} + \tilde{\phi}'\mathbf{V}^{-1})\phi + \mathbf{Y}'_{T+h}\mathbf{Y}_{T+h} + \tilde{\phi}'\mathbf{V}^{-1}\tilde{\phi}}_r
 \end{aligned} \quad (\text{A.16})$$

$$\Upsilon = \phi' \mathbf{A}\phi - \phi' \mathbf{b} - \mathbf{b}'\phi + r \quad (\text{A.17})$$

Somando e subtraindo $\mathbf{b}'\mathbf{A}^{-1}\mathbf{b}$ chega-se em

$$\begin{aligned}
 \Upsilon &= \phi' \mathbf{A}\phi - \phi' \mathbf{b} - \mathbf{b}'\phi + \mathbf{b}'\mathbf{A}^{-1}\mathbf{b} + \underbrace{r - \mathbf{b}'\mathbf{A}^{-1}\mathbf{b}}_c \\
 &= \phi' \mathbf{A}\phi - \phi' \mathbf{A}\mathbf{A}^{-1}\mathbf{b} - \mathbf{b}'\mathbf{A}\mathbf{A}^{-1}\phi + \mathbf{b}'\mathbf{A}^{-1}\mathbf{A}\mathbf{A}^{-1}\mathbf{b} + c \quad (\text{A.18})
 \end{aligned}$$

Chamando $\bar{\Sigma} = \mathbf{A}^{-1}$ e $\bar{\mu} = \mathbf{A}^{-1}\mathbf{b}$, assim

$$\begin{aligned}
 \Upsilon &= \phi'\bar{\Sigma}^{-1}\phi - \phi'\bar{\Sigma}^{-1}\bar{\mu} - \bar{\mu}'\bar{\Sigma}^{-1}\phi + \bar{\mu}'\bar{\Sigma}^{-1}\bar{\mu} + c \\
 &= (\phi - \bar{\mu})'\bar{\Sigma}^{-1}(\phi - \bar{\mu}) + c \quad (\text{A.19})
 \end{aligned}$$

Substituindo (A.19) em (A.14)

$$\begin{aligned}
p(Y_{T+h}|Y) &= \int \int |\Sigma_u|^{-\frac{(\lambda+1)T+n+h+1}{2}} \chi \\
&\times \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}((\lambda+1)T\tilde{\Sigma}_u + c)]\right\} \\
&\times \underbrace{\left[\int \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(\phi - \bar{\mu})'\bar{\Sigma}^{-1}(\phi - \bar{\mu})]\right\}d\phi\right]}_{\text{Núcleo de uma Normal para } \phi} d\Sigma_u d\theta
\end{aligned} \tag{A.20}$$

Sabendo que $\int N(\text{vec}(\bar{\mu}), \Sigma_u \otimes \bar{\Sigma}) = \mathbf{1}$, ou seja, o resultado da integral do núcleo é o inverso da constante de uma densidade, é possível simplificar a integral em relação a ϕ em (A.20)

$$\int \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}(\phi - \bar{\mu})'\bar{\Sigma}^{-1}(\phi - \bar{\mu})]\right\}d\phi = (2\pi)^{\frac{nk}{2}} |\bar{\Sigma}|^{-\frac{n}{2}} |\Sigma_u|^{\frac{k}{2}} \tag{A.21}$$

Substituindo (A.21) em (A.20)

$$\begin{aligned}
p(Y_{T+h}|Y) &= \int \int (2\pi)^{nk/2} |\bar{\Sigma}|^{-n/2} |\Sigma_u|^{-\frac{(\lambda+1)T-k+n+h+1}{2}} \chi \\
&\times \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}((\lambda+1)T\tilde{\Sigma}_u + c)]\right\} d\Sigma_u d\theta
\end{aligned} \tag{A.22}$$

Definindo $G = (\lambda+1)T\tilde{\Sigma}_u + c$

$$\begin{aligned}
p(Y_{T+h}|Y) &= \int (2\pi)^{\frac{nk}{2}} |\bar{\Sigma}|^{-\frac{n}{2}} \chi \\
&\times \underbrace{\left[\int |\Sigma_u|^{-\frac{(\lambda+1)T-k+n+h+1}{2}} \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}G]\right\}d\Sigma_u\right]}_{\text{Núcleo da Wishart-Invertida}} d\theta
\end{aligned} \tag{A.23}$$

Usando novamente o resultado de que a integral do núcleo é o

inverso da constante, obtém-se

$$\begin{aligned}
 & \left[\int |\Sigma_u|^{-\frac{(\lambda+1)T-k+n+h+1}{2}} \exp\left\{-\frac{1}{2}Tr[\Sigma_u^{-1}G]\right\} d\Sigma_u \right] \\
 &= 2^{\frac{n((\lambda+1)T-k+h)}{2}} \pi^{\frac{n(n-1)}{4}} |G|^{-\frac{(\lambda+1)T-k+h}{2}} \\
 &\times \prod_{i=1}^n \Gamma\left[\frac{(\lambda+1)T-k+h+1-i}{2}\right] \quad (A.24)
 \end{aligned}$$

Substituindo (A.24) em (A.23)

$$\begin{aligned}
 p(Y_{T+h}|Y) &= \int (2\pi)^{\frac{nk}{2}} |\bar{\Sigma}|^{-\frac{n}{2}} \chi^2 \frac{n((\lambda+1)T-k+h)}{2} \pi^{\frac{n(n-1)}{4}} \\
 &\times |G|^{-\frac{(\lambda+1)T-k+h}{2}} \prod_{i=1}^n \Gamma\left[\frac{(\lambda+1)T-k+h+1-i}{2}\right] d\theta \quad (A.25)
 \end{aligned}$$

Substituindo (A.15) em (A.25)

$$\begin{aligned}
 p(Y_{T+h}|Y) &= \int (2\pi)^{\frac{nk}{2}} (2\pi)^{-\frac{nh}{2}} (2\pi)^{-\frac{nk}{2}} \\
 &\times 2^{-\frac{n((\lambda+1)T-k)}{2}} \pi^{-\frac{n(n-1)}{4}} 2^{\frac{n((\lambda+1)T-k+h)}{2}} \pi^{\frac{n(n-1)}{4}} \\
 &\times |\lambda T \Gamma_{xx}^* + X'X|^{n/2} |\bar{\Sigma}|^{-n/2} \\
 &\times |(\lambda+1)T\tilde{\Sigma}_u|^{-\frac{(1+\lambda)T-k}{2}} |G|^{-\frac{(\lambda+1)T-k+h}{2}} \\
 &\times \prod_{i=1}^n \Gamma\left[\frac{(\lambda+1)T-k+1-i}{2}\right]^{-1} \\
 &\times \prod_{i=1}^n \Gamma\left[\frac{(\lambda+1)T-k+h+1-i}{2}\right] d\theta. \quad (A.26)
 \end{aligned}$$

Organizando os termos comuns é possível obter:

$$\begin{aligned}
 p(Y_{T+h}|Y) &= \int \pi^{-\frac{nh}{2}} \frac{\overbrace{|\lambda T \Gamma_{xx}^* + X'X|^{\frac{n}{2}}|(\lambda+1)T\tilde{\Sigma}_u|^{\frac{(1+\lambda)T-k}{2}}}}{|\bar{\Sigma}|^{\frac{n}{2}} |G|^{\frac{(\lambda+1)T-k+h}{2}}} \\
 &\times \prod_{i=1}^n \frac{((\lambda+1)T-k+h+1-i)/2}{((\lambda+1)T-k+1-i)/2} d\theta \quad (A.27)
 \end{aligned}$$

A equação (A.27) não depende mais dos parâmetros do VAR, apenas dos parâmetros do DSGE, portanto, os demais passos são apenas manipulações matemáticas com o objetivo de obter a mesma forma da densidade da distribuição variada t .

Trabalhando apenas com o termo F e lembrando que $V^{-1} \equiv \lambda T \Gamma_{xx}^* + X'X$ e $\bar{\Sigma}^{-1} \equiv V^{-1} + \bar{X}'\bar{X}$, tem-se:

$$\begin{aligned}
 F &= |V^{-1}|^{n/2} |(V^{-1} + \bar{X}'\bar{X})|^{-n/2} \\
 &= |V|^{-n/2} |(V^{-1} + \bar{X}'\bar{X})|^{-n/2} \\
 &= |V(V^{-1} + \bar{X}'\bar{X})|^{-n/2} \\
 &= |I + V\bar{X}'\bar{X}|^{-n/2} \\
 &= |I + \bar{X}V\bar{X}'|^{-n/2}.
 \end{aligned} \tag{A.28}$$

Substituindo (A.28) em (A.27), obtem-se:

$$\begin{aligned}
 p(Y_{T+h}|Y) &= \int \pi^{-\frac{nh}{2}} |I + \bar{X}V\bar{X}'|^{-\frac{n}{2}} |(\lambda + 1)T\tilde{\Sigma}_u|^{\frac{(1+\lambda)T-k}{2}} \\
 &\quad \times \underbrace{|G|^{-\frac{(\lambda+1)T-k+h}{2}}}_{K} \prod_{i=1}^n \frac{((\lambda + 1)T - k + h + 1 - i)/2}{((\lambda + 1)T - k + 1 - i)/2} d\theta
 \end{aligned} \tag{A.29}$$

Agora, trabalhando apenas com o termo K , sabendo que $G = (\lambda + 1)T\tilde{\Sigma}_u + c$, onde :

$$\begin{aligned}
 c &= r - b'A^{-1}b \\
 &= Y'_{t+h}Y_{t+h} + \tilde{\phi}'V^{-1}\tilde{\phi} - b'A^{-1}b \\
 &= Y'_{t+h}Y_{t+h} + \tilde{\phi}'V^{-1}\tilde{\phi} \\
 &\quad - (\bar{X}'Y_{T+h} + V^{-1}\tilde{\phi}')(V^{-1} + \bar{X}'\bar{X})^{-1}(\bar{X}'Y_{T+h} + V^{-1}\tilde{\phi})
 \end{aligned}$$

em que r , b e A^{-1} foram definidos em (A.16).

Então,

$$\begin{aligned}
 G &= (\lambda + 1)T\tilde{\Sigma}_u + Y'_{t+h}Y_{t+h} + \tilde{\phi}'V^{-1}\tilde{\phi} \\
 &\quad - (\bar{X}'Y_{T+h} + V^{-1}\tilde{\phi}')(V^{-1} + \bar{X}'\bar{X})^{-1}(\bar{X}'Y_{T+h} + V^{-1}\tilde{\phi})
 \end{aligned} \tag{A.30}$$

Através de manipulações matemáticas é possível chegar em

$$\begin{aligned}
 G &= (\lambda + 1)T\tilde{\Sigma}_u + Y'_{T+h} \overbrace{(I - \bar{X}(V^{-1} + \bar{X}'\bar{X})^{-1}\bar{X}')^I} Y_{T+h} \\
 &+ \underbrace{\tilde{\phi}'(V^{-1} - V^{-1}(V^{-1} + \bar{X}'\bar{X})^{-1}V^{-1})\tilde{\phi}}_{II} \\
 &- 2 \underbrace{\tilde{\phi}V^{-1}(V^{-1} + \bar{X}'\bar{X})^{-1}\bar{X}' Y_{T+h}}_{III}
 \end{aligned} \tag{A.31}$$

onde termos destacados em (A.31) podem ser reescritos da seguinte maneira:

Trabalhando apenas com o termo *I*, este será definido com base no resultado de Searle (1982) que estabelece que $(I + AB)^{-1} = I - A(I + BA)^{-1}B$. Definindo $A = \bar{X}V$ e $B = \bar{X}'$, tem-se:

$$I - \bar{X}(V^{-1} + \bar{X}'\bar{X})^{-1}\bar{X}' = (I + \bar{X}V\bar{X}')^{-1}. \tag{A.32}$$

Trabalhando apenas com o termo *II*, utilizando o resultado de Searle (1982) para matrizes dado por $A - A(A + B)^{-1}A = B - B(A + B)^{-1}B$, considerando $A = V^{-1}$ e $B = \bar{X}'\bar{X}$.

$$\begin{aligned}
 &(V^{-1} - V^{-1}(V^{-1} + \bar{X}'\bar{X})^{-1}V^{-1}) \\
 &= \bar{X}'\bar{X} - \bar{X}'\bar{X}(V^{-1} + \bar{X}'\bar{X})^{-1}\bar{X}'\bar{X} \\
 &= \bar{X}'(\bar{X} - \bar{X}(V^{-1} + \bar{X}'\bar{X})^{-1}\bar{X}) \\
 &= \bar{X}'(I - \bar{X}(V^{-1} + \bar{X}'\bar{X})^{-1})\bar{X} \\
 &= \bar{X}'(I + \bar{X}V\bar{X}')^{-1}\bar{X}.
 \end{aligned} \tag{A.33}$$

Agora, trabalhando apenas com o termo *III*, tem-se:

$$\begin{aligned}
 V^{-1}(V^{-1} + \bar{X}'\bar{X})^{-1}\bar{X}' &= V^{-1}(V^{-1} + \bar{X}'\bar{X}VV^{-1})^{-1}\bar{X}' \\
 &= V^{-1}((I + \bar{X}'\bar{X}V)V^{-1})^{-1} \\
 &= V^{-1}(V^{-1})^{-1}(I + \bar{X}'\bar{X}V)^{-1}\bar{X}' \\
 &= (I + \bar{X}'\bar{X}V)^{-1}\bar{X}' \\
 &= \bar{X}'(I + \bar{X}V\bar{X}')^{-1}.
 \end{aligned} \tag{A.34}$$

Por fim, substituindo os termos *I*, *II* e *III* na equação (A.31):

$$\begin{aligned}
 G &= (\lambda + 1)T\tilde{\Sigma}_u + Y'_{T+h}(I + \bar{X}V\bar{X}')^{-1}Y_{T+h} \\
 &+ \tilde{\phi}'\bar{X}'(I + \bar{X}V\bar{X}')^{-1}\bar{X}\tilde{\phi} - \tilde{Z}'\bar{X}'(I + \bar{X}V\bar{X}')^{-1}Y_{T+h} \\
 &= (\lambda + 1)T\tilde{\Sigma}_u + (Y_{T+h} - \bar{X}\tilde{\phi})'(I + \bar{X}V\bar{X}')^{-1}(Y_{T+h} - \bar{X}\tilde{\phi}).
 \end{aligned} \tag{A.35}$$

Substituindo (A.35) em (A.29), tem-se:

$$\begin{aligned}
 p(Y_{T+h}|Y) &= \int \pi^{-\frac{nh}{2}} |I + \bar{X}V\bar{X}'|^{-\frac{n}{2}} |(\lambda + 1)T\tilde{\Sigma}_u|^{\frac{(1+\lambda)T-k}{2}} \\
 &\times |(\lambda + 1)T\tilde{\Sigma}_u \\
 &+ (Y_{T+h} - \bar{X}\tilde{\phi})'(I + \bar{X}V\bar{X}')^{-1}(Y_{T+h} - \bar{X}\tilde{\phi})|^{-\frac{(\lambda+1)T-k+h}{2}} \\
 &\times \prod_{i=1}^n \frac{((\lambda + 1)T - k + h + 1 - i)/2}{((\lambda + 1)T - k + 1 - i)/2} d\theta.
 \end{aligned} \tag{A.36}$$

Por fim, manipulações matemáticas levam a forma final da densidade preditiva a posteriori:

$$\begin{aligned}
 p(Y_{T+h}|Y) &= \int \pi^{-\frac{nh}{2}} |I + \bar{X}V\bar{X}'|^{-\frac{n}{2}} |(\lambda + 1)T\tilde{\Sigma}_u|^{-\frac{h}{2}} \\
 &\times |I + (\lambda + 1)T\tilde{\Sigma}_u^{-1} \\
 &\times (Y_{T+h} - \bar{X}\tilde{\phi})'(I + \bar{X}V\bar{X}')^{-1}(Y_{T+h} - \bar{X}\tilde{\phi})|^{-\frac{(\lambda+1)T-k+h}{2}} \\
 &\times \prod_{i=1}^n \frac{((\lambda + 1)T - k + h + 1 - i)/2}{((\lambda + 1)T - k + 1 - i)/2} d\theta.
 \end{aligned} \tag{A.37}$$

onde $Y_{T+h} \in \mathfrak{R}^{h \times n}$, $(\bar{X}\tilde{\phi}) \in \mathfrak{R}^{h \times n}$, $(I + \bar{X}V\bar{X}')^{-1}(h \times h) > 0$, $((\lambda + 1)T\tilde{\Sigma}_u)(n \times n) > 0$ e $(\lambda + 1)T - k > 0$.

Diz então que $Y_{T+h} \sim Mt_{hn}(\bar{X}\tilde{\phi}, (I + \bar{X}V\bar{X}')^{-1}, (\lambda + 1)T\tilde{\Sigma}_u, (\lambda + 1)T - k)$. Entretanto, é importante perceber que quando $h = 1$, (A.37) reduz-se a uma distribuição t multivariada. Mais especificamente, tomando $\bar{X}\tilde{\phi} = \nu'$ e $(I + \bar{X}V\bar{X}') = \sigma$, é fácil perceber que $Y'_{T+1} \sim t_n((\lambda + 1)T - k, \sigma, \nu, (\lambda + 1)T\tilde{\Sigma}_u)$.

A expectativa preditiva de Y_{T+1} é, então, computada da seguinte maneira:

$$E(Y_{T+1}|Y, \phi, \Sigma_u, \theta) = (\bar{X}\tilde{\phi})' \tag{A.38}$$

APÊNDICE B – TABELAS E RESULTADOS

Tabela 7: Dados

Dados	Período	Fonte	Transformações
Índice de Ativ. Econ. do Banco Central (IBC-Br) - com ajuste sazonal	Jan/03 - Dez/16	Banco Central	Primeira diferença do logaritmo
Índice nac. de preços ao consumidor amplo (IPCA) - Var. % mensal	Jan/03 - Dez/16	IBGE	Nenhum
Índice da taxa de câmbio efetiva real (IPCA) - Jun/1994=100	Jan/03 - Dez/16	Banco Central	Primeira diferença do logaritmo
Taxa de juros - Selic acumulada no mês (%a.a.)	Jan/03 - Dez/16	Banco Central	Nenhum

Tabela 8: P-valor para a hipótese nula de convergência da cadeia de Markov

VAR(2)	
Nome	$\lambda = \mathbf{0}, 15$
σ	0,2413
φ	0,0622
θ	0,3680
ϕ_π	0,0999
ϕ_y	0,2331
ρ_a	0,4896
ρ_y	0,4657
ρ_z	0,1641
γ^*	0,4476
π^*	0,4019
q^*	0,4639
ρ	0,2822
σ_i	0,3505
σ_a	0,1594
σ_y	0,2338
σ_z	0,2224

Tabela 9: Parâmetros Calibrados

Parâmetro	Média	Descrição
α	0,12	Grau de abertura da economia
η	2	Substituição entre os bens domésticos e importados
γ	1	Substituição entre os diferentes bens produzidos em diferentes países estrangeiros.